Identification of Novel Interacting Partners of the Pre-mRNA Processing Factor 31

A Thesis Submitted to the University College London
for the Degree of Doctor in Philosophy

Francesca Fiocco

MSc

Department of Molecular Genetics
Institute of Ophthalmology
University College London
2009

Declaration

I, Francesca Fiocco confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed
Francesca Fiocco

Date

Abstract

Mutations in *PRPF31* (RP11 locus) cause autosomal dominant retinitis pigmentosa (adRP), an inherited disorder of the retina characterised by degeneration of rod photoreceptors. *PRPF31* is ubiquitously expressed and encodes a splicing factor involved in the spliceosome assembly. Interestingly, the phenotype caused by mutations in *PRPF31* is restricted to the retina.

This study aims to identify protein interactants of PRPF31, to contribute to a better understanding of the disease mechanism of RP11.

Three potential interacting candidates were isolated through the yeast two-hybrid technique. Two of them (MAGI3, TRAK2) were confirmed to interact with PRPF31 using co-immunoprecipitation experiments after overexpression of the proteins; co-immunoprecipitation of the endogenous proteins was not carried out. *MAGI3* and *TRAK2* expression in human retina was proven by RT-PCR, however they are not retina-specific. *MAGI3* and *TRAK2* human retina transcripts did not appear to present major differences to the transcripts deposited in the Ensembl database.

Tagged proteins were overexpressed in epithelial cells, and co-localisation between PRPF31 and each partner was mainly nuclear. In epithelial cells, both endogenous PRPF31 and MAGI3 localised to intercellular junctions but did not co-localise; however, after stress treatments, the proteins appeared to co-localise to the nucleus. In neuroblastoma cells, endogenous PRPF31 and TRAK2 did not co-localise. Immunohistochemistry experiments identified PRPF31 to be present largely in the inner segment and in the outer plexiform layer of murine retina.

AdRP patients' DNA was screened to detect mutations in *MAGI3* that could be linked to retinal degeneration, but none was identified.

The results are not conclusive of the molecular mechanism of RP11, but instead provide new grounds for research on PRPF31. These insights into the type of interacting partners and the localisations observed within the cell and the retina may cast new light into the function(s) of PRPF31, which may (or may not) link to retinal degeneration.

Acknowledgments

I would like to thank first and foremost my supervisor, Professor Shomi Bhattacharya, for giving me the opportunity to achieve such a high plain, and for believing in me and in my work: I really appreciate your constant support. I would also like to extend my appreciation to Professor Karl Matter, for his constructive criticism on my work and for his teaching and guidance, and also for letting me use his laboratory, reagents and protocols!

I also thank Dr Cecilia Maubaret, for her supervision and advice in the initial times; and also Dr Christina Chakarova, for her valuable suggestions and for providing me with materials.

I would also like to warmly thank everyone in the group: Naushin, Mai, Amna, Ciara, Kinga, Sancy, Brotati, Giovanna, Vanita, Beverley and Quincy. Without your cheerfulness, these years would have been very difficult!

A special thank to Amna and Naushin for proofreading my PhD manuscript, particularly to Amna for correcting it and reading it several times, and for her precision and ability to read my mind!

I would also like to thank all the people that have collaborated on this project: Stef Letteboer, Ronald Roepman and Frans Cremers, from Nijmegen; Olga Makarova and Genia Makarov from Leicester, for their advice and expertise.

Finally, I would really like to thank past and present people at the Institute: Alison and Virginia for their support, Margherita for her friendship as a scientist and a flatmate, and Peter, Wayne, Matt, Yanai, Louise, Nele, Livia, Juan, Emma. And I would like to thank all the RetNet fellows for their company during the various meetings!

Special thanks to my father and sister who have supported me unconditionally, encouraging me all the time throughout my years of studying. Your presence and love was always appreciated!

List of Publications

Fiocco F, Maubaret C, Roepman R, Cremers FPM, Matter K and Bhattacharya SS. Investigation and Characterization of PRPF31 Interacting Proteins. Poster presentation at the European Vision Summit (EVS), 2007.

Fiocco F, Maubaret C, Roepman R, Cremers FPM, Matter K and Bhattacharya SS. Investigation and Characterization of PRPF31 Interacting Proteins in Retina. Poster presentation at the Association of Research in Vision and Ophthalmology (ARVO) annual meeting, 2008.

Table of Contents

List of Tables and Figures

Abbreviations

A	Adenine
aa	Amino acids
ABCR	ATP-binding cassette transporter, retina specific
AD	Activation domain
adRP	Autosomal dominant retinitis pigmentosa
Arr	Arrestin
arRP	Autosomal recessive retinitis pigmentosa
ATP	Adenosine triphosphate
b/bp	Base/base pair
BD	Binding domain
Brr2/BRR2	Pre-mRNA-splicing helicase BRR2
C	Cytosine
Ca^{2+}	Calcium ions
CB	Cajal body
CC	Connecting cilium
CDK	Cyclin-dependent kinase
cDNA	Complementary deoxyribonucleic acid
cGMP	Cyclic guanosine monophosphate
CRALBP	Cellular retinal-binding protein
Da	Daltons
ddNTP	Dideoxynucleotide
DEAD-box	Aspartic acid/glutamic acid/alanine/aspartic acid
DMEM	Dulbecco's modified Eagle's medium
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
EDC	Exon definition complex
ESE	Exonic splicing enhancer
ESS	Exonic splicing silencer
G	Guanine
GCAP	Guanylate cyclase activating protein
GCL	Ganglion cell layer
gDNA	Genomic DNA
GDP	Guanosine diphosphate
GK	Guanylate kinase domain

GPCR	G-protein coupled receptor
GTP	Guanosine triphosphate
H	Histidin
HAP	Huntingtin-associated protein
HEK293	Human embryonic kidney cell line
hnRNP	Heterogeneous ribonucleoprotein
ILM	Internal limiting membrane
IMPDH1	Inosine-5'-monophosphate dehydrogenase 1
INL	Inner nuclear layer
IP	Immunoprecipitation
IPL	Inner plexiform layer
IRBP	Interphotoreceptor retinoid-binding protein
IS	Inner segment
K^+	Potassium ions
L	Leucin
LRAT	Lecithin:retinol acyl transferase
MAGI3	Membrane-associated guanylate kinase with inverted orientation 3
MAGUK	Membrane-associated guanylate kinase
MDCK	Madin Darby canine kidney cell line
mRNA	Messenger ribonucleic acid
Na^+	Sodium ions
$NADP^+$/NADPH	Nicotinamide adenine dinucleotide phosphate
NLS	Nuclear localisation signal
NMD	Nonsense mediated decay
NOP	snoRNA binding domain
NOSIC	Central domain in Nop56/SIK1-like proteins
OD	Optical density
OLM	Outer limiting membrane
ONL	Outer nuclear layer
OPL	Outer plexiform layer
OS	Outer segment
PAGE	Polyacrylamide gel electrophoresis
PAP1	Pim-1-associated protein
PCR	Polymerase chain reaction
PDE	Phosphodiesterase
PDZ	**P**ostsynaptic density protein-95, **d**isk-large tumor suppressor protein, **z**onula occludens-1

PMSF	Phenylmethylsulfonyl fluoride
PRPF	Pre-mRNA processing factor
Prp28p	Pre-mRNA-splicing ATP-dependent RNA helicase PRP28
Prp8	Pre-mRNA processing factor 8
RAL	Retinal
RDH	Retinol dehydrogenase
RE	Retinyl ester
Rec	Recoverin
RHO	Rhodopsin
RK	Rhodopsin kinase
RNA	Ribonucleic acid
ROL	Retinol
RP	Retinitis pigmentosa
RPE	Retinal pigment epithelium
rpm	Revolutions per minute
RT-PCR	Reverse transcription polymerase chain reaction
SDS	Sodium dodecyl sulphate
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
snRNP	Small nuclear ribonucleoprotein
SPLICE	Swift PCR for ligating *in vitro* constructed exons
SR (protein)	Serine/arginine-rich (protein)
T	Thymine
TF	Transcription factor
TOPORS	Topoisomerase I binding RS domain protein
TRAK2	Trafficking protein, kinesin-binding 2
U	Uracil
UAS	Upstream activation sequence
UTR	Untranslated region
UV	Ultra violet
U2AF	U2 auxiliary factor
W	Tryptophan
WT	Wild type
xlRP	X-linked retinitis pigmentosa
X-α-Gal	5-bromo-4-chloro-3-indolyl α-D-galactopyranoside
X-β-Gal	5-bromo-4-chloro-3-indolyl β-D-galactopyranoside
ZNF143	Zinc finger protein 143

1. Introduction

1.1 The Eye

The eye is the sensory organ responsible for the detection of light, hence allowing an individual to perceive the surrounding environment. This sensory system detects external light stimuli, originating from various sources, and converts them into electric signals. These are sent along the optic nerve and converge to specific regions of the brain where they are further processed into the images we ultimately perceive. This complex system allows visual perception by means of a precise cellular organisation and a fine-tuned electrical net.

The human eye has a spherical shape, being 22-24 mm in diameter (Figure 1.1). The cornea forms the clear part of the visible area of the eye, and it protrudes from the main globe; the rest of the eye is formed by the sclera, which is a white and opaque membrane partially visible in the frontal part of the eye. Three layers compose the eye: 1) the external layer, formed by the cornea (transparent) and the sclera; the latter is mainly composed of collagen which protects the inner parts and helps maintain the spherical structure; 2) the choroid: a highly vascularised tissue which provides oxygen and nutrients; and 3) the retina: the innermost layer formed by neurons, which is the photosensitive component of the eye involved in detecting and part processing visual information. In the cavity formed by the three layers (the posterior chamber) lies the vitreous humour, a thick gel-like substance which helps maintain the shape of the eye. It is separated from the frontal part by the lens. The space between the lens and the cornea (known as the anterior chamber) is filled with aqueous humour; the iris lies in front of the lens, and regulates the aperture of the pupil (Bron *et al.*, 1997).

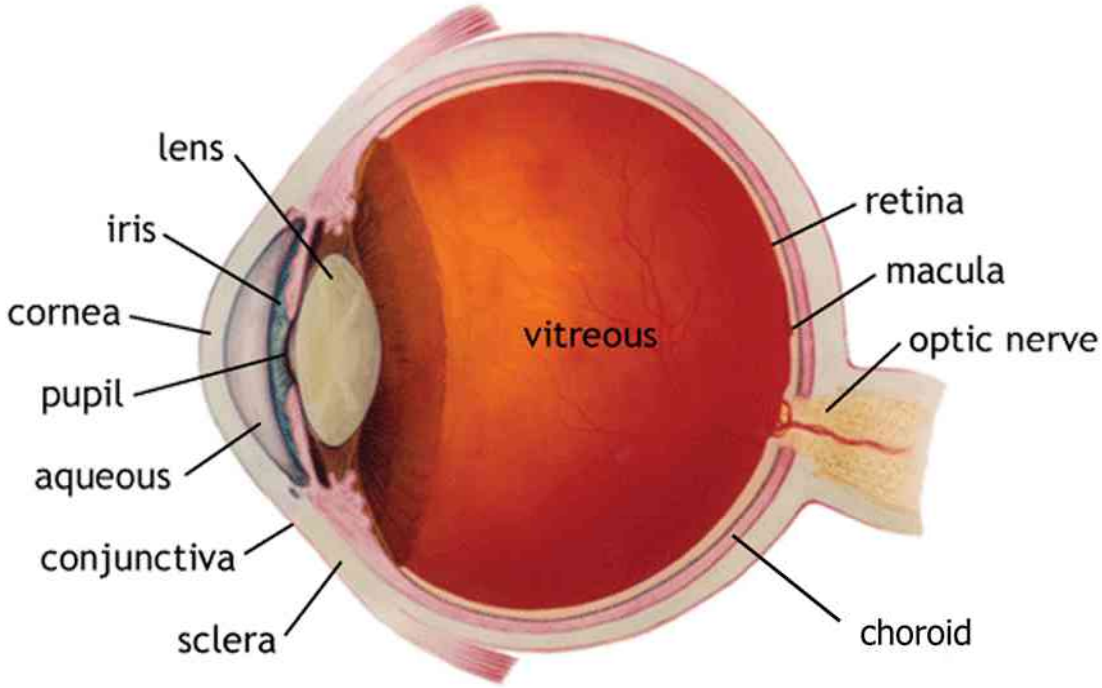


Figure 1.1. Anatomy of the human eye.
Anteroposterior section of the bulb. Two chambers are distinguishable, the anterior and the posterior chambers, followed by the vitreous body and last by a series of membranes (present also on the sides). On the outside the anterior chamber is delimited by the cornea, on the inside it is limited by the lens and the iris. The posterior chamber is comprised between the iris and the formations which constrain the lens. At the back is the vitrous body. The membranes which encase the bulb are the retina, the choroid and the sclera. (Image adapted from http://www.blundelloptometry.com).

1.2 The Retina

The human retina is a thin transparent tissue located at the back of the eye, between the vitreous humour and the retinal pigment epithelium. It is responsible for detecting photons of light and converting them to electrical impulses that are sent along the optic nerve to the brain. The human retina is approximately 0.5 mm thick, it is part of the central nervous system and composed of six different types of neural cells: photoreceptor cells, horizontal cells, bipolar cells, amacrine cells, interplexiform cells and ganglion cells. These cells are organised in a neuronal network which detects the visual stimuli, allowing for high sensitivity (rod photoreceptor pathway) and high precision (cone photoreceptor pathway). Along with neurons, other cells are present, such as Müller cells and epithelial cells.

1.2.1 Organisation of the Retina

The retina can be divided into several layers (Figure 1.2), including (following a ray of light, from the front to the back):

- internal limiting membrane (ILM);
- nerve fibre layer (NFL);
- ganglion cell layer (GCL);
- inner plexiform layer (IPL);
- inner nuclear layer (INL);
- outer plexiform layer (OPL);
- outer nuclear layer (ONL);
- outer limiting membrane (OLM)
- inner and outer segments of photoreceptors (IS and OS);
- retinal pigment epithelium (RPE).

A ray of light has to cross several layers of cells before reaching the photoreceptors, the light-sensitive cells in the penultimate layer of the retina. The retina itself is separated from the vitreous humour by the internal limiting membrane (ILM). Behind the ILM are the ganglion cell fibres, which project axons to the brain via the optic nerve relaying the gathered visual information. The INL contains the nuclei of bipolar, horizontal,

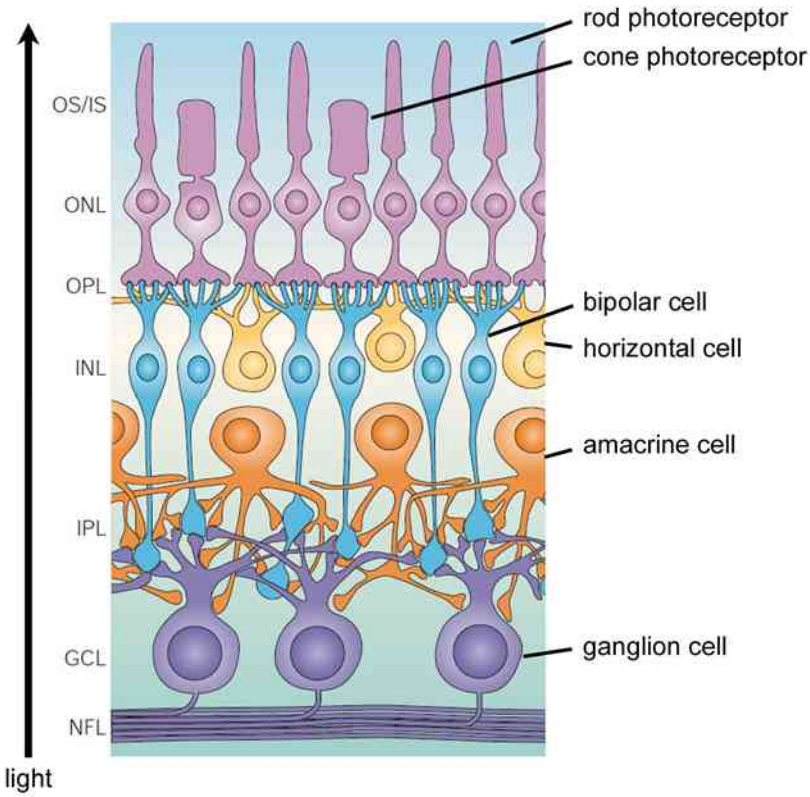


Figure 1.2. Structure of the human retina.
Schematic representation of the layers composing the mammalian retina (adapted from Wässle, 2004). (OS: outer segment; IS: inner segment; ONL: outer nuclear layer; OPL: outer plexiform layer; INL: inner nuclear layer; IPL: inner plexiform layer; GCL: ganglion cell layer; NFL: nerve fibre layer).

amacrine and interplexiform cells; bipolar cells contact photoreceptors and transmit the signal to ganglion cells; the remaining cells help regulate the transduction of the signal; Müller glia cells are mainly present in the INL, but their processes spread across the adjacent layers: they play a role in the metabolism of the retina, and in its development, as they help orienting, disposing and displacing the developing neurons. The outer nuclear layer (ONL) contains the soma and nuclei of photoreceptors cells, with their outer segment extending outwards towards the retinal pigment epithelium (RPE). The RPE is a single layer of epithelial cells that lies at the back of the retina, adjacent to the photoreceptors, next to the choroid. The cells are polarised and tightly packed and are involved in crucial functions such as nourishing and protecting the retinal environment. The RPE facilitates transport and exchange of nutrients, oxygen and by-products between the retina and the choroid. It contacts and engulfs the tips of photoreceptor OS, providing not only a structural support but is also involved in the renewal of the OS by phagocytosing the photoreceptor tips and metabolising their content (Figure 1.3). Consequently, it plays a role in the generation of the visual chromophore (11-*cis*-retinal, see section 1.3.2). Finally, the melanin granules present throughout the RPE help absorb the scattered light, preventing inner reflection and any damage due to phototoxicity.

When a photon of light is absorbed by the photoreceptor cell, the stimulus is conveyed to bipolar cells first, and then to ganglion cells which eventually convey the information via the optic nerve to the brain. Hence one recognises a transverse organisation among the retinal layers: a receptor can activate more than one bipolar cell, and one single bipolar cell can activate more than one ganglion cell. The signal is amplified as it is spread horizontally, while it moves vertically towards the optic nerve. The amplification is tuned by the neighbouring horizontal cells, together with amacrine and interplexiform cells, all of them playing a modulatory role within the system.

1.2.2 Photoreceptor Cells

Photoreceptor cells are responsible for the processing of the visual stimuli into electric signals. They constitute the penultimate layer of the retina, lined by RPE on the outermost side and making contact with bipolar cells on the other.

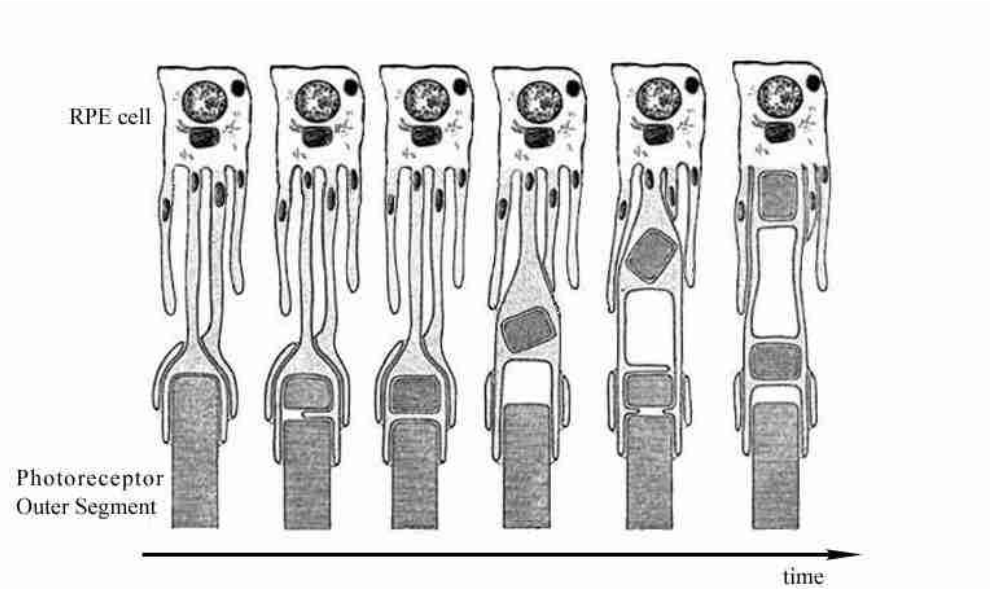


Figure 1.3. Photoreceptor disc shedding.
The tips of the photoreceptor outer segments are phagocytosed by the RPE cells and metabolised. Image adapted from webvision.med.utah.edu

There are two types of photoreceptors: rod cells and cone cells. They have different photopigment, phototransduction and chromophore regeneration dynamics, explaining why rods are responsible for vision in dim lighting conditions (scotopic vision, *i.e.* at night) and cones for colour or bright light conditions (photopic vision).

Both rods and cones share a basic structure comprising four distinct sections: the axon terminal, the nuclear region (cell body), the inner segment (IS) and the outer segment (OS) (Figure 1.4). The axon terminal synapses upon bipolar cells in the OPL; the cell body contains the nucleus and is found in the ONL. The IS is rich in mitochondria, Golgi bodies and ribosomes, being where metabolic reactions take place; the OS is the site of the phototransduction and it is rich in visual pigment. The rod OS is cylindrical and composed of a stack of flattened membranous discs; they originate from the plasma membrane, and eventually are pinched off, and then lay free one on top of the other within the plasma membrane. Conversely, the cone OS tapers off and its discs remain attached to the plasma membrane along the entire length of the outer segment. The discs are replaced on a daily basis, as the older ones at the tip of the OS are phagocytosed by the RPE in a process known as disc shedding. The OS is entirely renewed within ten days, as new discs constantly originate at the base of the OS. The protein synthesis mechanisms in the IS have to be highly efficient in order to produce the large amount of proteins required for the renewal (in rods, more than 90% of the OS proteins are rhodopsin; Palczewski, 2006). The IS and OS are connected by a nonmotile connecting cilium, which facilitates the trafficking of proteins (and other metabolites) between the two segments. The synaptic termini of both cone and rod are connected with the bipolar cells, which in turn are connected to the ganglion cells. The signal generated by the light travels through these three types of cells.

Photoreceptors show a diverse distribution throughout the human retina: generally cones are numerous in the macula, an area approximately at the centre of the retina. In the macula is the fovea, a virtually rod-free pit involved in central vision; given its high abundance of cones, the macula is responsible for high acuity vision. Towards the peripheral retina, cones decrease gradually in favour of rods, which become more abundant. Of all the photoreceptors, ~92 million are rods and ~4.6 million are cones (Bron *et al.*, 1997).

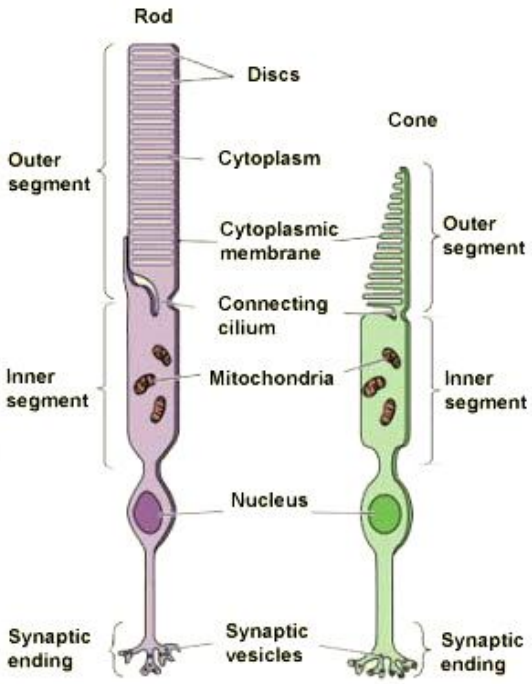


Figure 1.4. Rod and cone photoreceptors.
Rods and cones are structurally comparable, both showing a partition in four sections (axon terminal or synaptic ending, the nuclear region, the inner segment and the outer segment). Cones are shorter than rods, although wider in diameter: the synaptic terminus and the nuclear region are larger, and the outer segment is shorter. Image adapted from thebrain.mcgill.ca.

The light signal is detected by the visual pigment found within the disc membrane of the OS of rod and cone photoreceptors; it is formed by a protein moiety (opsin) and a chromophore (11-*cis*-retinal). Rods possess only one type of opsin, rhodopsin, with its absorption peak at 498 nm. Cones, however, contain different photo pigments: there are three different opsins, with peak absorptions at different wavelengths: short-wavelength opsin (420 nm – blue cones), medium-wavelength opsin (530 nm – green cones) and long-wavelength opsin (560 nm – red cones). Similarly, according to the opsin present in the discs, cones are categorised as short-wavelength, medium-wavelength or long-wavelength cones, respectively. Since rods detect only one wavelength, they provide monochromatic vision, whereas cones provide colour vision. The two photoreceptor types have different kinetics of signal transduction: rods have a longer recovery time than cones, showing more sensitivity to light and are therefore involved in vision in low lighting conditions. Cones show a faster recovery, allowing vision in bright light; they are less sensitive and discriminate between different light intensities, and the interaction between the responses of the different cone types determines colour, depth and intensity.

1.3 Physiology of the Retina

The visual pigments, or light-receptive proteins, are the opsins found on the membranous discs of the OSs: rhodopsin in rods and cone-opsin(s) in cones. When a photon of light is captured by a visual pigment such as rhodopsin, it initiates the phototransduction cascade, which converts the stimulus into an electric response. Opsins belong to the G-protein coupled receptor (GPCR) family; these proteins are characterised by seven transmembrane helices, whose spacial arrangement gives the receptor an elliptical shape (Figure 1.5). The N-terminal part is extracellular (it is located in the inner space of the disc), whereas the C-terminus is cytosolic.

Rhodopsin and cone-opsin pigments consist of a protein moiety, the opsin, and a chromophore, 11-*cis*-retinal, which is embedded within the transmembrane domains, and covalently bound to a lysine residue. Opsin itself does not absorb light; the light-sensitive molecule is the chromophore. 11-*cis*-retinal is non active, and its “bent” conformation allows the molecule to fit stably into the opsin.

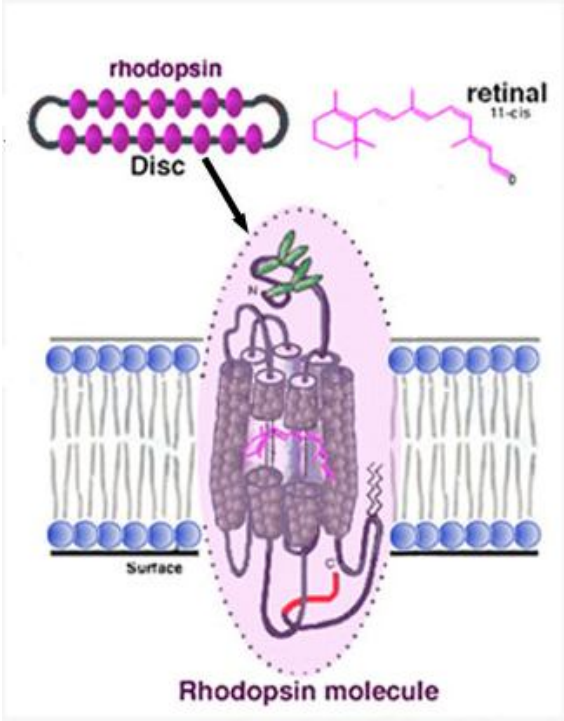


Figure 1.5. The rhodopsin molecule.
Rhodopsin molecules (purple balls on top left) are embedded within the rod disc membrane (Blue and grey layer in main image). Each molecule of rhodopsin has seven transmembrane domains, with the chromophore retinal buried within them. Image adapted from webvision.med.utah.edu

When the visual pigment absorbs a photon of light, the photosensitive molecule isomerises to all-*trans*-retinal, whose "straight" conformation acts as a spring and allows the molecule to break free from the opsin. This way the photo-activated receptor sets off the G-protein cascade, leading to the phototransduction.

1.3.1 Phototransduction

In the dark, photoreceptors are depolarised, and maintain a steady concentration of Na^+ and K^+. The high concentration of cGMP keeps cGMP-gated Na^+-channels open, allowing for an inward flux of Na^+, whereas non-gated channels allow for an outward K^+ flux, keeping the cell depolarised (Figure 1.6); these channels are located on the OS plasma membrane. This state of depolarisation allows an influx of Ca^{2+}. Light causes hyperpolarisation of the cell, because the Na^+ flux is blocked and the Ca^{2+} concentration decreases. The hyperpolarised state results in the transmission of a signal to the optic nerve, through the layers of the retina.

When the photon of light hits rhodopsin (Figure 1.6 A), the 11-*cis* retinal undergoes a conformational change to all-*trans*-retinal: the receptor switches to its activated form (R*), catalysing the exchange of GDP to GTP in the subunit α of the G-protein transducin (G); the enzymatically active G*α binds to the inhibitory subunits of phoshodiesterase (E), which is then able to catalyse the hydrolysis of cGMP (cG in Figure 1.6 A). The resulting decrease in cytoplasmic concentration of cGMP provokes the closure of cGMP-gated channels, therefore blocking the inward flux of Ca^{2+} and Na^+.

The decreased Ca^{2+} concentration triggers three mechanisms of recovery: one is regulated by recoverin (Rec, Figure 1.6 B). Recoverin has an affinity for rhodopsin kinase (RK) at normal Ca^{2+} concentration, but when the Ca^{2+} concentration drops, it frees the RK, allowing the latter to phosphorylate rhodopsin (R* to R-P); R-P is a substrate for arrestin (Arr) which binds and quenches rhodopsin activity. Another way is by inactivation of transducin (G, Figure 1.6 C), which is possible by the hydrolysis of GTP to GDP, leading to the inactivation of phosphodiesterase (E). A third mechanism is controlled by a guanylyl cyclase (GC, Figure 1.6 A):

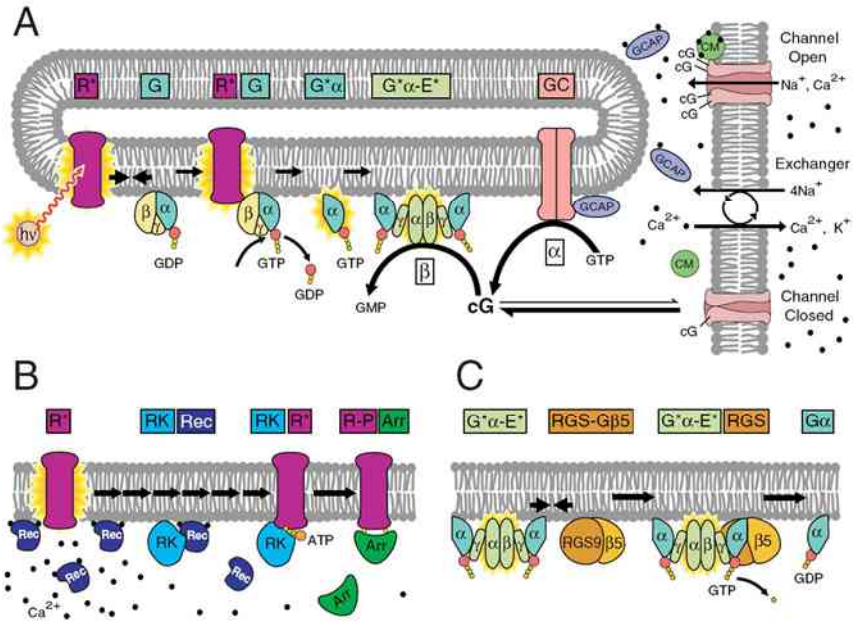


Figure 1.6. Schematic view of phototransduction in vertebrate photoreceptor cells.
A: the activation of the cascade and the quenching mediated by guanylyl cyclase; B and C: quenching mediated by recoverin and by inactivation of transducin respectively. Image from Pugh & Lamb 2000.

as Ca^{2+} concentration drops, the guanylyl cyclase activating protein (GCAP) releases its Ca^{2+} and binds to GC increasing its cyclase activity, raising the cGMP cytoplasmic concentration, hence restoring the inward fluxes of Ca^{2+} and Na^{+} (Pugh & Lamb, 2000).

1.3.2 The Visual Cycle

In rod cells, once the chromophore has isomerised to all-*trans*-retinal, it can only return to 11-*cis*-retinal through an enzyme pathway called the visual cycle (Figure 1.7). Following photon absorption, rhodopsin is activated, and after its inactivation, all-*trans*-retinal dissociates and diffuses into the bilayer of the OS discs. It is transported into the cytoplasm by the ATP-binding cassette transporter ABCA4, and here it is reduced to all-*trans*-retinol (vitamin A) by a retinol dehydrogenase (RDH), likely to be in a NADPH-dependent manner. Following this, the all-*trans*-retinol reaches the extracellular space and eventually enters the RPE cells. All-*trans*-retinol also enters when the shed OS are phagocytosed by the RPE cells, and by endocytosis of interphotoreceptor retinoid-binding protein (IRBP). All-*trans*-retinol can also be taken up from the choroid, as it circulates in the blood stream.

In the RPE, all-*trans*-retinol is esterified by lecithin:retinol acyl transferase (LRAT) and the ester is later converted to 11-*cis*-retinol by isomerase RPE65; the product binds to cellular retinal-binding protein (CRALBP). 11-*cis*-retinol is finally oxidised to 11-*cis*-retinal by 11-*cis*-retinol dehydrogenase (RDH5 and RDH11). The chromophore is then transported from the RPE to the OS to regenerate the visual pigments. Alternatively, 11-*cis*-retinol can be esterified by LRAT to 11-*cis*- retinyl ester, and stored (Figure 1.7; as reviewed by Travis *et al.*, 2007).

Recent studies have revealed an alternate cone-cycle (Figure 1.8) to regenerate the chromophore in cone-dominated retina. This pathway does not depend on RPE but relies on Müller cells, and could explain the faster time of recovery of cone photoreceptors. Cone cells can regenerate the visual pigments by taking up both 11-*cis*-retinal and 11-*cis*-retinol, since a $NADP^{+}$/NADPH-dependent 11-*cis*-RDH activity was discovered, making them independent from RPE as a source of 11-*cis*-retinal, as opposed to rods. The cone-visual cycle resembles the rod-visual cycle: all-*trans*-retinal is reduced to all-*trans*-retinol by NADPH-dependent all-*trans*-RDH in cone OS;

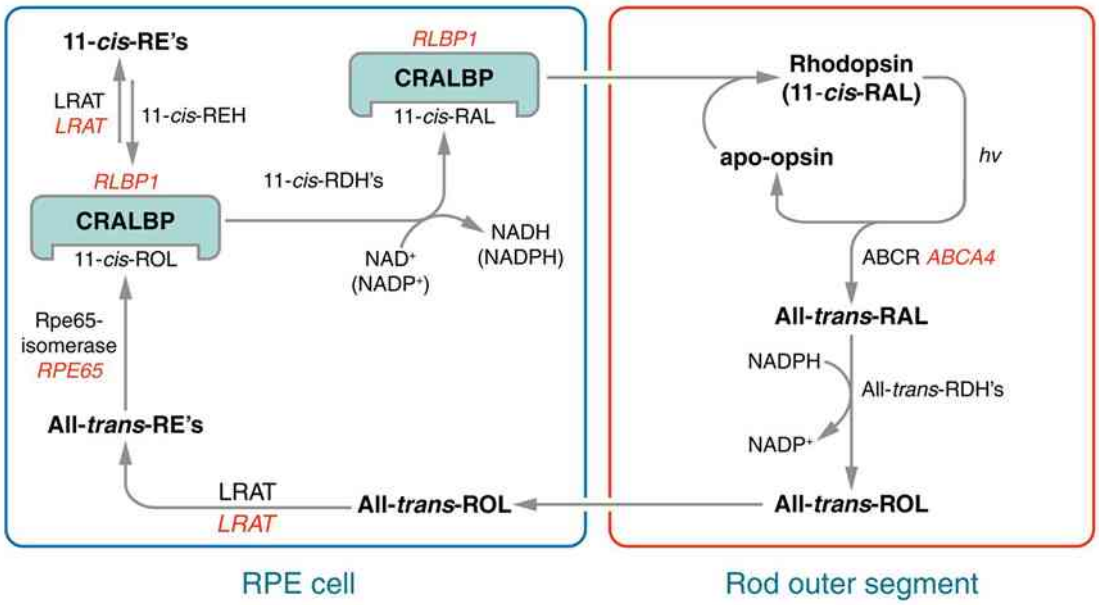


Figure 1.7. The visual cycle.
The genes mutated in RP and involved in this pathway are shown in red. Image adapted from Travis et al., 2007. (RAL: retinal; ROL: retinol; RE: retinyl ester; RPE: retinal pigment epithelium).

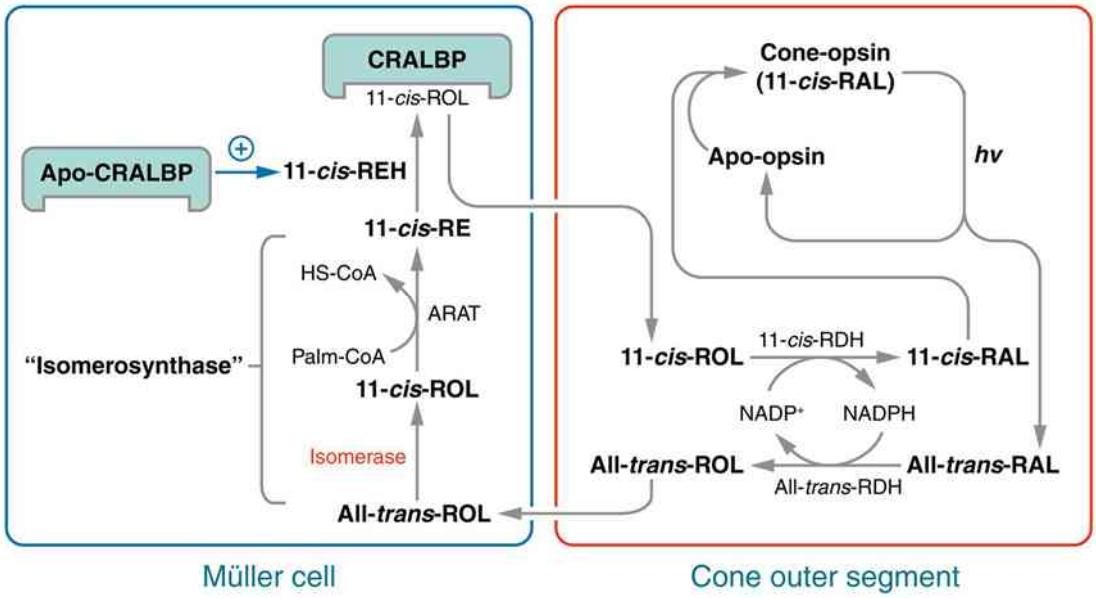


Figure 1.8. The cone-visual cycle.
Alternative cycle for regeneration of cone pigments. Image taken from Travis et al., 2007. (RAL: retinal; ROL: retinol; RE: retinyl ester).

but unlike in rods, all-*trans*-retinol enters the Müller cell and is isomerised to 11-*cis*-retinol which re-enter the cone OS and it is converted to 11-*cis*-retinal by the $NADP^+$-dependent 11-*cis*-RDH (Figure 1.8) (As reviewed by Muniz *et al.*, 2007, Travis *et al.*, 2007).

1.4 Retinitis Pigmentosa

1.4.1 Overview

Retinitis pigmentosa (RP; MIM#268000) is an inherited disorder of the retina, characterised by progressive degeneration of rod photoreceptors followed by cone degeneration, eventually resulting in severe visual impairment to loss of sight. Originally, the term "retinitis pigmentosa" was given to this type of disease believing that it was caused by retinal inflammation. However, now, it is well known that retinitis pigmentosa primary etiology is not inflammation but photoreceptor cell death.

Inherited retinal degenerations are a group of disorders characterised by deterioration of the retina, causing loss of neuronal cells. These disorders range from the most common retinitis pigmentosa to Leber congenital amaurosis (LCA), when rods and cones are dead/non-functional at birth (Francois, 1968); macular degeneration, when only the rods and the cones of the macula degenerate, with preservation of the peripheral retina (Green & Enger, 1993); cone degeneration, when all three cone types die leaving rods alive and functional (To *et al.*, 1998), leading to rod monochromacy in which vision is mediated by rods only; hereditary optic atrophy, characterised by a reduced number of retinal ganglion cells; stationary night blindness, characterised by a full complement of rods and cones, however rods are not as sensitive to dim light and therefore show longer adaptation times (Dryja, 2000). Other types of inherited retinal degenerations affect the macula, like the Best disease, also known as vitelliform macular dystrophy, which is caused by the deposition of lipofuscin-like material in the subretinal space, damaging irreparably the macula (White *et al.*, 2000); or the Doyne's honeycomb retinal dystrophy, characterised by drusen deposits (white-yellow deposits) in the macular area and the optic nerve head (Gregory *et al.*, 1996).

Due to rod cells degeneration, RP initially manifests as difficulties in dark adaptation to night blindness, and with gradual loss of peripheral vision; RP is diagnosed through

testing of visual field loss and photoreceptor activity (electroretinogram). RP is highly heterogeneous, both phenotypically and genetically: more than 40 genes associated with RP have been identified (Daiger *et al.*, 2007), with autosomal dominant, autosomal recessive and X-linked modes of inheritance (see section 1.4.3); affected individuals may experience loss of vision very early on in life, whereas other patients do not manifest the disease until their 40s or 50s, or never become legally blind, retaining some sight late into their life. Although most of the cases are represented by non-syndromic RP, RP can manifest together with other non-ocular symptoms, such as deafness (Usher syndrome; Boughman *et al.*, 1983), or cognitive impairment, obesity, polydactyly, hypogenitalism and renal disease (Bardet-Biedl syndrome; Tieder *et al.*, 1982; Beales *et al.*, 1999), or nephronophthisis (Senior-Loken syndrome, Saunier *et al.*, 2005); many other syndromes which associate with RP are reviewed in Hamel, 2006.

1.4.2 Symptoms of Retinitis Pigmentosa

Many of the patients suffering from RP experience two key symptoms: night blindness and progressive loss of visual field (tunnel vision, Figure 1.9) both due to early loss of rod cells in the peripheral retina. With progression of the disease, cone cells can be also affected and therefore degenerate, disturbing day vision and central visual acuity. Fundus examinations often reveal other symptoms of RP, which include: attenuation of retinal vessels, pallor of the optic nerve, and perivascular pigmentation in the form of bone-spicule deposits due to the migration of RPE cells into the neural retina (Li *et al.*, 1995), in response to photoreceptor death (Figure 1.10). If none of the described symptoms is yet detectable, it is possible to diagnose RP via electroretinograms, which measure and record the electrical response of the retina to flashes of light.

Histologically, RP leads to thinning of the outer nuclear layer, as photoreceptors die, whereas the inner nuclear layer and the ganglion cell layer remain intact (Figure 1.11).

In most cases, patients only suffer retinal degeneration (non-syndromic RP), but there are cases in which RP is manifested together with other non-ocular symptoms (syndromic RP), with an autosomal recessive mode of inheritance. There are several syndromes (30) and they account for the 20-30% of RP cases: Usher syndrome is the most common, accounting for 10-20% of RP cases (Hartong *et al.*, 2006), it is

Figure 1.9. Retinitis pigmentosa disease progression: tunnel vision.
Peripheral vision is gradually lost in RP, resulting in tunnel vision, as only the central vision is left functioning. On the left, a representation of normal vision (unaffected patient). On the right what an RP patient might see. Image adapted from http://en.wikipedia.org/wiki/Tunnel_vision.

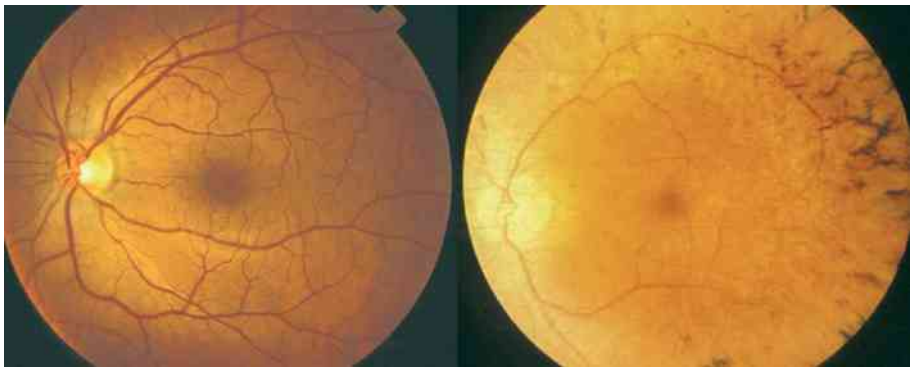


Figure 1.10. Comparison between a normal fundus, left, and a RP patient fundus, right.
In RP patients, waxy colour of the optic disc, attenuation of blood vessels and peripheral pigmentation are noticeable on the affected patient's fundus. (Image from Hartong *et al.*, 2006)

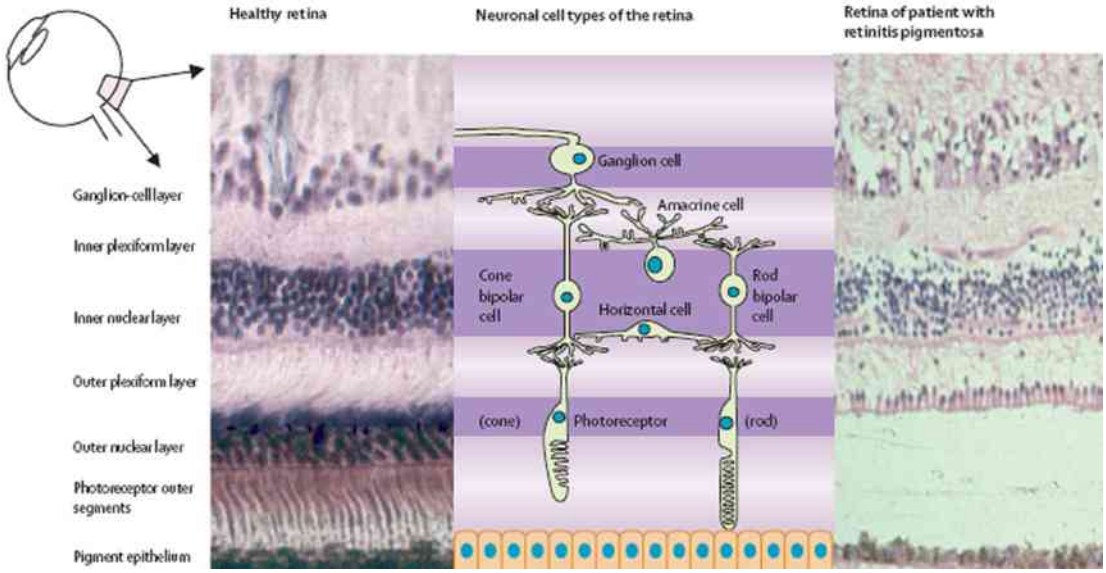


Figure 1.11. Histology of healthy- (left) and RP-retina (right).
There is a noticeable thinning of the photoreceptor nuclear layer in the RP retina, due to photoreceptor cell death. In contrast, bipolar cells and ganglion cells nuclei remain relatively intact. (Image from Hartong *et al.*, 2006).

characterised by retinal disorder (RP) and hearing loss (Boughman *et al.*, 1983); Bardet-Biedl syndrome is another major form of syndromic RP (5-6% of RP cases; Hartong *et al.*, 2006), and patients display severe symptoms such as obesity, cognitive impairment, polydactyly, hypogenitalism and renal disease alongside RP (Tieder *et al.*, 1982; Beales *et al.*, 1999); the Senior-Loken syndrome, characterised by renal disease (Saunier *et al.*, 2005).

1.4.3 Genetics of Retinitis Pigmentosa

RP affects one in 4000 individuals worldwide, hence accounting for more than a million cases (Hartong *et al.*, 2006). Most forms of RP are monogenic, and present a typical inheritance pattern: autosomal recessive RP (arRP) represents the ~30% of the cases, autosomal dominant RP (adRP) accounts for the ~20% of the cases and finally X-linked RP (xlRP) represents the ~15% of the cases (Daiger *et al.*, 2007). A small proportion of cases show different modes of inheritance, such as mitochondrial inheritance, digenic diallelic and digenic triallelic inheritance (Rivolta *et al.*, 2002). Approximately 30% of RP cases have no family history (Daiger *et al.*, 2007). Generally, xlRP is developed at an early age and usually shows the most severe phenotypes; arRP presents a milder phenotype than xlRP; finally, adRP shows the latest onsets and a milder phenotype (Hamel, 2006). However, among the affected for each pattern of inheritance, there may be phenotypical heterogeneity, due to different genetic background and environmental conditions, along with a different molecular cause.

To date, 53 loci associated with RP have been identified (http://www.sph.uth.tmc.edu/Retnet/sum-dis.htm#A-genes, accessed 17/09/2009), and more than 40 genes have been identified (Daiger *et al.*, 2007). Most of the genes encode for proteins implicated in photoreceptor and RPE biology, with a currently small but growing portion encoding ubiquitously expressed proteins. Table 1.1 provides an overview of the majority of the RP genes, categorised according to their known/presumed function.

Table 1.1 Retinitis Pigmentosa Genes.

Retinitis Pigmentosa Genes	Inheritance	Reference
Phototransduction Cascade		
RHO - Rhodopsin	Dominant,recessive	Hargrave 2001
PDE6A - Rod cGMP-specific phosphodiesterase subunit α	Recessive	Koutalos et al 1995
PDE6B - Rod cGMP-specific phosphodiesterase subunit β	Recessive	Koutalos et al 1995
CNGA1 - cGMP-gated cation channel alpha-1	Recessive	Dhallan et al 1992
CNGB1 - cGMP -gated cation channel beta-1	Recessive	Korschen et al 1999
SAG - Arrestin	Recessive	Palczewski et al 1998
Visual Cycle		
ABCA4 - Retinal-specific ATP-binding cassette transporter	Recessive	Weng et al 1999
RLBP1 - Retinaldehyde-binding protein 1	Recessive	Saari et al 2001
RPE65 - Visual cycle isomerase	Recessive	Xue et al 2004
LRAT - Lecithin retinol acyltransferase	Recessive	Sue et al 2004
RGR - RPE-retinal G protein-coupled receptor	Recessive	Chen et al 2001
Structural, cytoskeletal, intracellular trafficking		
RDS - Peripherin-2	Dominant, digenic	Travis et al 1991
ROM1 - Rod outer segment membrane protein 1	Dominant, digenic	Clarke et al 2000
FSCN2 - Fascin-2	Dominant	Saishin et al 2000
TULP1 - Tubby-related protein 1	Recessive	Xi et al 2005
CRB1 - Crumbs homolog 1	Dominant	Pellikka et al 2002
RP1 - Oxygen-regulated protein 1	Dominant,recessive	Liu et al 2004
TOPORS - Topoisomerase I binding, arginine/serine-rich	Dominant	Chakarova et al 2007
RPGR - Retinitis pigmentosa GTPase regulator	X-linked	Hong et al 2003
EYS - Protein eyes shut homolog	Recessive	Abd El-Aziz et al 2008
Signaling, cell-cell interaction, synaptic interaction		
SEMA4A - Semaphorin-4A	Dominant	Rice et al 2004
USH2A - Usherin	Recessive	Reiners et al 2005
RP2 - Protein XRP2	X-linked	Chapple et al 2003
NRL - Neural retina-specific leucine zipper protein	Dominant	Swaroop et al 1992
Splicing factors		
PRPF3 - Pre-mRNA processing factor 3	Dominant	Wang et al 1997
PRPF8 - Pre-mRNA processing factor 8	Dominant	Umen et al 1995
PRPF31 - Pre-mRNA processing factor 31	Dominant	Zhou et al 2002
PAP1 - Pim-1-associated protein	Dominant	Maita et al 2004
ASCC3L1 - RNA helicase BRR2	Dominant	Li et al 2009
Other		
CERKL - Ceramide kinase-like protein	Recessive	Bornancin et al 2005
MERTK - Proto-oncogene tyrosine-protein kinase MER	Recessive	Vollrath et al 2001
NR2E3 - Retinal nuclear receptor transcription factor	Dominant,recessive	Kobayashi et al 1999
CA4 - Carbonic anhydrase 4	Dominant	Yang et al 2005
IMPDH1 - Inosine-monophosphate dehydrogenase 1	Dominant	Bowne et al 2002

The table summarises most of the disease genes, their related product functions and mode of inheritance.

1.4.4 Autosomal Dominant Retinitis Pigmentosa

Autosomal dominant retinitis pigmentosa (adRP) is caused by mutations in 16 known genes, and to date one locus (RP33) for which the gene may be *ASCC3L1* (Li *et al.*, 2009; http://www.sph.uth.tmc.edu/Retnet/sum-dis.htm#A-genes, accessed 17/09/2009).

The gene for rhodopsin (*RHO*) was the first gene to be found associated with RP. It encodes the photopigment rhodopsin which plays a pivotal role in the phototransduction cascade as previously discussed. Mutations in rhodopsin account for the majority of adRP cases (25%; Hartong *et al.*, 2006): over 100 mutations have been found in the gene, all leading to aberrant protein products, with one or few amino acids deleted or changed (Wang *et al.*, 2005).

Mutations in retinal-specific genes (e.g. photoreceptor- or RPE-specific genes) can understandably result in retinal degeneration. However there are a group of genes whose products are ubiquitously expressed, and yet mutations in them are only known to cause non-syndromic RP. For adRP, these include *PRPF31* (Vithana *et al.*, 2001), *PRPF3* (Chakarova *et al.*, 2002), *PRPF8* (McKie *et al.*, 2001), *PAP1* (Keen *et al.*, 2002), *ASCC3L1* (Li *et al.*, 2009), *IMPDH1* (Bowne *et al.*, 2002), and *TOPORS* (Chakarova *et al.*, 2007). IMPDH1 is a dehydrogenase, and catalyses the synthesis of guanine nucleotides. TOPORS is a nuclear protein with ubiquitin and SUMO-1 E3 ligase activity which localises to the connecting cilium of mouse rod photoreceptor cells (Dr C. Chakarova personal communication). The five remaining ubiquitously expressed genes are all involved in pre-mRNA splicing, among them is *PRPF31*, which is the second most common gene known to be mutated in adRP (5% of adRP cases, Hartong *et al.*, 2006). Its disease mechanism is not understood yet, and before discussing its role in the pathogenesis of RP, the following section will provide an overview of splicing. To note, other ubiquitously expressed genes are found also in the arRP category – e.g. *MERTK* (Graham *et al.*, 1994) or *USH2A* (Eudy *et al.*, 1998) - and xlRP category – e.g. *RPGR* (Kirschner *et al.*, 1999) or *RP2* (Schwahn *et al.*, 1998) - .

1.5 RNA Splicing

The vast majority of eukaryotic genes contain regions which code for a protein (exons) and regions that do not code (introns). A protein is therefore the product of the translation of the coding regions only. In order for this to occur, the non-coding regions, introns, must be removed. After the transcription of a gene, a primary transcript (pre-mRNA) is produced, having the same organisation of the gene as the DNA, containing both coding and non-coding regions. To generate a single continuous molecule of joined exons, the pre-mRNA faces various modifications such as 5' capping and the 3' polyadenylation, and also the removal of the intronic regions. The critical process responsible for intron removal and the joining of exons is known as splicing. Hence, splicing is defined as a post-transcriptional genetic modification: it occurs naturally in eukaryotic cells, and takes place in the nucleus, along with the other RNA modifications. Splicing leads to a mature transcript (mRNA), which is then ready for downstream processes like the translation into a protein, allowing the gene to be expressed.

1.5.1 Biology of Splicing

Splicing is carried out by a multi-megadalton ribonucleoprotein enzyme, the spliceosome: it is a large complex of proteins and small nuclear RNA (snRNA) molecules. Five snRNAs are involved in splicing: U1, U2, U4, U5, U6. Each of them associates with several different proteins, forming ribonucleoprotein particles (snRNPs). The proteins are either involved in the splicing reaction, take part in the assembly and the structure of the complex, or else help connections between the snRNPs. Many other non-snRNP proteins, generally termed "splicing factors", also take part in the splicing assemblage, and more than 200 have been identified through mass spectrometry (Liu *et al.*, 2006). From the number and type of elements implicated, it is comprehensible that a correct and functional spliceosome relies on RNA-RNA interactions, as well as RNA-protein and protein-protein interactions, building evidence for the high sophistication of the entire spliceosomal molecular machine.

Introns show defined features in order to be recognised from exons, allowing for correct excision. In fact, the identification of the splice sites is mediated by highly conserved

sequences found at the intron boundaries only. The 5' end is marked by a conserved dinucleotide, GU; the 3' end is characterised by three elements: a branch site nucleotide A, a polypyrimidine sequence and a final AG dinucleotide. The GU-AG sequences define the intron margins.

The splicing reaction can be divided in two major steps: first, the recognition of the consensus sequences, and second, the cleavage and ligation of the substrate RNA. The reaction consists of a first cut at the 5' end of the exon-intron junction, with the release of the exon as a linear molecule (Figure 1.12); the intron bends on itself forming a lariat, linking its 5' end to the target base A (branch site). Next, a cut at the 3' end releases the intron and the two exons are ligated together; finally the intron is linearised and degraded. Introns are not necessarily spliced sequentially from the 3' to the 5', but are spliced depending on the accessibility of the recognition sequences.

The splicing complex forms *de novo* at each round of splicing. Figure 1.13 illustrates the splicing dynamics, including the snRNPs. At the beginning is complex E, when U1 snRNP base-pairs to 5' splice site GU, U2AF (U2 auxiliary factor) recognises the 3' polypirimidine tract and 3'-terminal AG, and associates to it, SF1 binds to the branch site sequence. These associations bridge the intron and bring the flanking exons close to each other. Complex E evolves to A as soon as U2 snRNP is recuited, and its snRNA base-pairs with the branch point sequence in an ATP-dependent fashion. Subsequently, again through an ATP-dependent way, U4/U6•U5 tri-snRNP joins to form the large spliceosome (complex B). Finally, molecular rearrangements of complex B, including the detachment of U1 and U4 snRNPs, allow the formation of the catalytic complex C, which performs the transesterification reactions, resulting in the ligation of the exons and the excision of the intron.

Along with a major class of introns which show the features described above, a class of minor introns has been identified: they are spliced by a different spliceosome apparatus, termed the minor spliceosome. Minor introns rarely display the common GU-AG boundary, and they also have a different branch site sequence (Figure 1.14). The set of snRNPs involved is also different: U11, U12, U4/atac, U6/atac, which cover the same role as the major U1, U2, U4 and U6; U5 is shared with both spliceosomes.

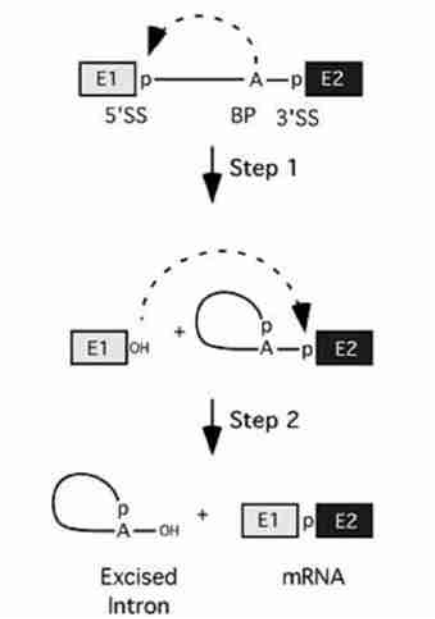


Figure 1.12. Two-step splicing reaction.
The branch site A is first recognised by U2 snRNP (not depicted), with the intron sequence forming a lariat on itself; secondly, exon 1 (E1) is joined to exon 2 (E2). *p* is for phosphate groups. Image from Will & Luhrmann 2005.

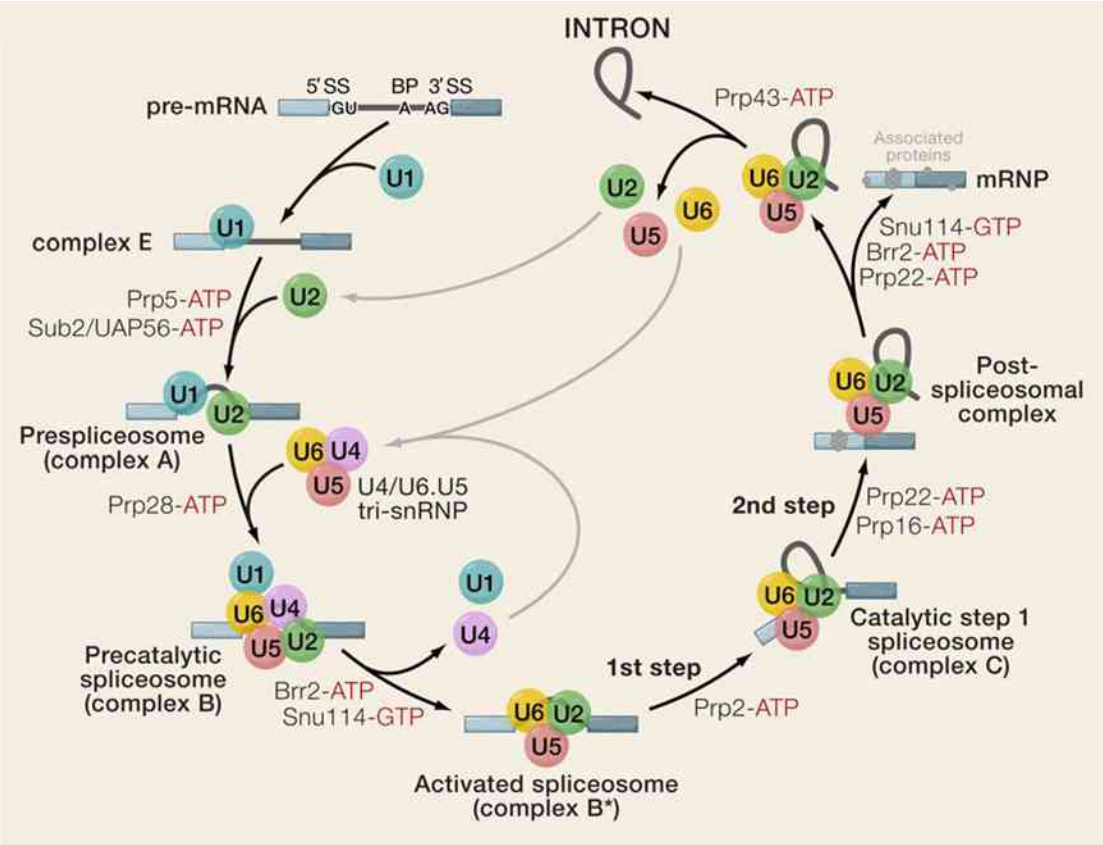


Figure 1.13. Splicesome assembly.
The U1 snRNP base pairs with the 5' splice site and U2AF associates with the 3' end region (complex E). U2AF (not shown) recruits U2 snRNP which base pairs with the branchpoint sequence forming the complex A. This is followed by the formation of the large spliceosome, the complex B, with the addition of U4/U6•U5 tri-snRNP. Subsequent rearrangements, such as the disassociation of U1 and U4 (complex B*), and transesterification reactions (complex C) result in an intron lariat and the spliced mRNA. SS: splice site; BP: branch point. Image adapted from Wahl *et al.*, 2009.
The distinction between introns and exons is achieved through the exon definition complexes (EDCs; not shown), which are a set of SR proteins bound along the length of the exon, defining it; the intron is recognised through its conserved sequences (GU-AG).

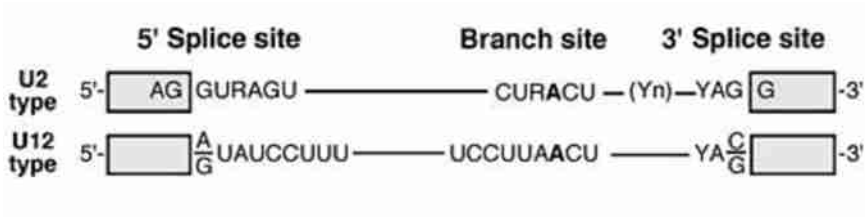


Figure 1.14. U2 and U12 type introns.
U2 type introns present the conserved sequences AG/GURAGU at the 5' splice site, CURACU at the branch site, a polypyrimidine tract (Yn) and the YAG/G sequence at the 3' splice site ("/" denotes an exon-intron boundary, R is a purine, Y is a pyrimidine). U12 type introns have different and longer recognition sequences: A(G)UAUCCUUU at the 5' site, UCCUUAACU at the branch site and YAC(G) at the 3' site; U12 type introns lack a polypirimidine tract. Bold letters represent the branch point. Image from Will & Luhrmann 2005.

This minor complex is named U12-type spliceosome, whereas the major apparatus described before, is named U2-type spliceosome (as reviewed by Will & Luhrmann 2005).

Major and minor spliceosomes appear to share most of the proteins composing them, and only seven are specific to the minor: essentially, they are part of the U11•U12 di-snRNP, hence they could be involved in the recognition of the splicing site. Regarding U11 and U12, they form a di-snRNP, which does not bind the pre-mRNA alone as U1 and U2, but together after association.

Compared to the dynamics of the major counterpart, in the minor spliceosome assembly, the complex E does not exist: the spliceosomal machinery assembles straight into complex A. Moreover, U11 and U4atac snRNPs do not appear to take part in the catalytic complex C. A distinct feature of the U12 spliceosome is its lower rate of splicing compared to the U2 equivalent (Patel *et al.*, 2002): because of that, it has been proposed that minor splicing is involved in post-transcriptional modulation of gene expression. Also, its presence is ~100-fold less than the U2 complex, and minor introns account for less than 1% of all introns (Will & Luhrmann 2005).

The various snRNPs must be assembled beforehand: this process starts and finishes in the nucleus (Figure 1.15). snRNPs are assembled and gathered in the Cajal bodies (CBs), which are non-membranous structures present within the nucleus. The initial step for the snRNP biogenesis is the transcription of the snRNA, which happens in the nucleus; from here, all the snRNAs except U6 are transported to the cytoplasm, where they are modified and proteins are attached; finally they return to the nucleus (Figure 1.15). U6 snRNA does not shuttle to the cytoplasm and back, as all its modifications happen in the nucleus. Once the snRNA is modified, it is targeted to the CB and subjected to assembly with several proteins to form the complete snRNP particle (Figure 1.15). It is likely that the minor snRNP particles also follow the same pathway, since the same set of proteins is shared (although this may not be true for U11 and U12, as they have several dedicated splicing factors together with the common ones; Stanek & Neugebauer, 2006).

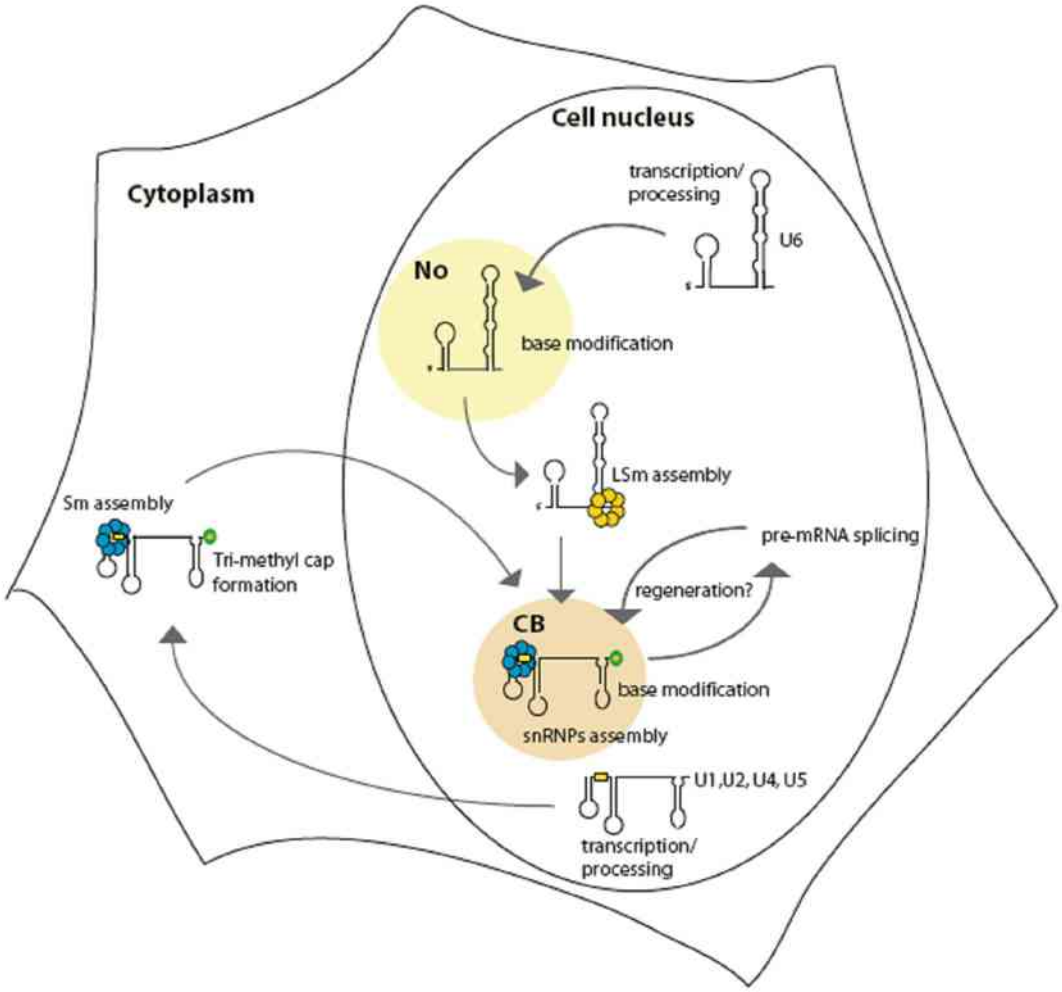


Figure 1.15. snRNP biogenesis.
After transcription and processing, the U1, U2, U4, U5 snRNAs are shuttled to the cytoplasm, where they undergo hypermethylation of the 5' end (green circle) and an addition of Sm ring (blue circles). These modifications allow the import of the U snRNAs to the cajal body (CB) in nucleus where they undergo base modifications and are assembled as snRNPs. U6 snRNA follows a different pathway: after its transcription and processing, a methyl cap is added; it is then shuttled to the nucleolus (No) where it undergoes base modification, and eventually, an LSm ring (yellow circles) is added and U6 is targeted to the cajal body, to complete snRNP assembly. Image from Stanek & Neugebauer 2006.

1.5.2 Splicing Regulation

The splicing mechanism can be regulated to achieve either inclusion or exclusion of exons in the mature transcript, a process known as alternative splicing. Alternative splicing occurs for a large number of genes, and it is a principal mode by which organisms regulate the protein output, to create proteome diversity.

Patterns of alternative splicing are illustrated in Figure 1.16: changes may affect the 5' end, with alternative promoters; or the 3' end with different poly-A sites; or else, alternative 5' or 3' splice sites may be used; and also cassette exons, mutually exclusive exons, and introns that can be retained and become part of the final coding sequence.

These regions are selected and then either retained or excised from the pre-mRNA. There are auxiliary regions, which may be intronic or exonic, which either promote the recruitment of the splicing machinery (splicing enhancers) to retain the sequence, or can cause exon skipping (splicing silencers) by disrupting the spliceosome. In the first instance, the SR set of proteins bind to exonic splicing enhancers (ESEs), favoring the exon inclusion, whereas a set of hnRNPs bind to exonic splicing silencers (ESSs), and repress exon inclusion (Black, 2003). However, SR proteins are thought to constantly bind exonic sequences, as they interact with U1 snRNP and U2AF, helping their recruitment to the pre-mRNA. To regulate exon exclusion, as SR proteins are already bound normally and their association with the ribonucleic sequence is not very strong, neighboring proteins can alter their association, shifting the splicing dynamics to exon exclusion. Hence, according to the protein species present in a given cell, there may be profound differences in the splicing pattern of the same gene. SR proteins are also involved in promoting the association between U6 snRNA and the 5' splice site, hence they are likely to modulate this event as well, leading to a different outcome. On the other hand, hnRNPs are thought to provide a steric block, preventing snRNPs particles U1 and U2 from accessing the pre-mRNA (House & Lynch 2008).

Other type of regulations include the modifications of the splicing factors constituting the snRNPs: for example, yeast Prp28, a DEAD-box helicase required to release U1 from the spliceosome, must be phosphorylated to be recruited to the splicoesome (Mathew *et al.*, 2008). If Prp28 is not present, U1 snRNA cannot be exchanged with the

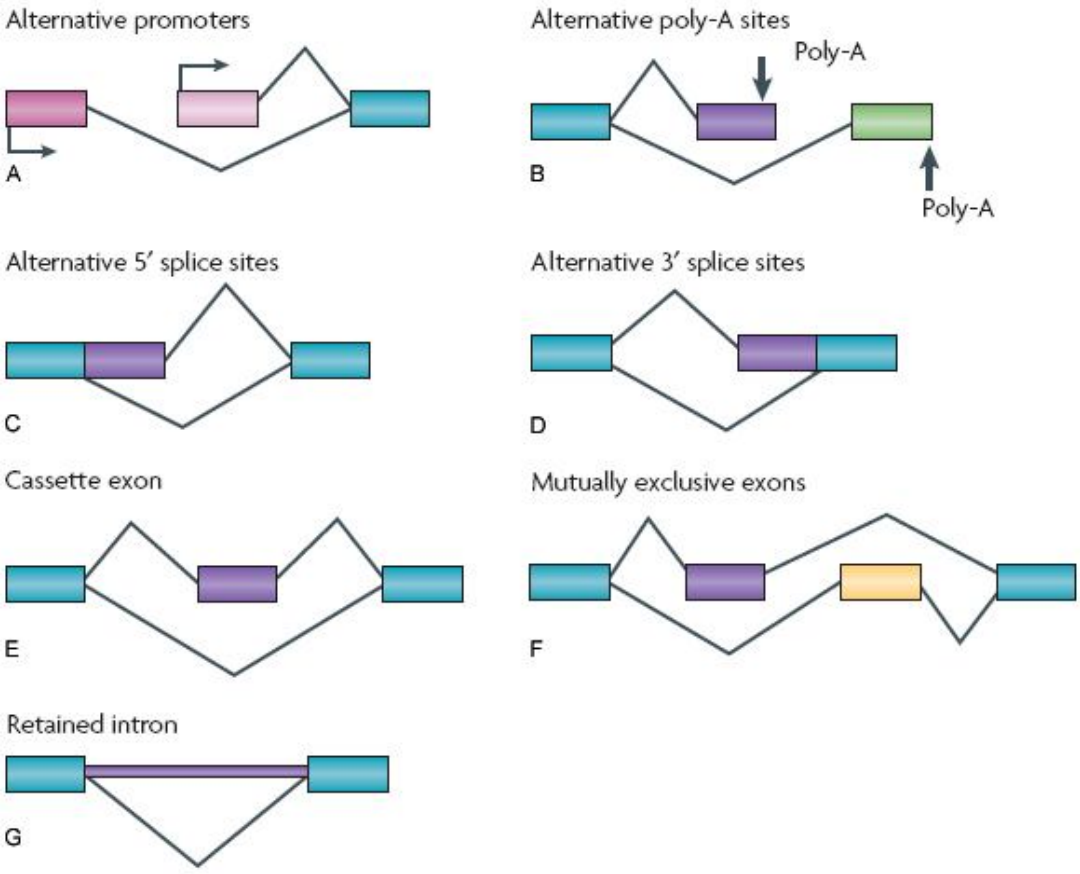


Figure 1.16. Different patterns of alternative splicing.
A, alternative promoters generate exons (light pink and dark pink) which are joined to a common downstream exon. B, alternative poly-A sites generate alternative exons which are joined to a common upstream exon. Alternative 5' and 3' splice sites (C and D respectively) lead to exon shortening/extension. E, cassette exons can be included or excluded. F, mutually exclusive exons are particular cassette exons, whose inclusion depends whether or not the other has been included or skipped, as they are not both retained. G, a retained intron is generated when the intronic sequence is not excised and becomes part of the mRNA. (Image adapted from Li *et al.*, 2007)

U6 snRNA at 5' splice site in the event that the 5'splice site differs from the consensus (House & Lynch 2008).

It is plausible that these and other forms of regulation happen during the first moments of spliceosome formation, *i.e.* at the pre-spliceosomal stage during complex E and A arrangement. There are also changes that may happen during the catalytic step. Bellare and colleagues (2008) elegantly showed that ubiquitination of Prp8 modulates Brr2 activity (in yeast), downregulating its action of unwinding U4/U6.

This evidence strongly indicates that the spliceosome not only is a large structure, but it is also highly dynamic and responds to numerous changes in the cellular environment, whether they are due to action of neighboring proteins, post-translational modifications occurring to the proteins composing it, or alterations affecting the substrate. It is expected that many other implications will be revealed in the future that will help elucidate further the splicing mechanism.

1.6 Splicing Factors and Retinitis Pigmentosa

As described in the previous section 1.4.3, the majority of genes known to be mutated in RP encode for products which play a crucial role in retina physiology. Nevertheless, some of them are ubiquitously expressed in a wide range of tissues, still causing a retina-specific phenotype. Among them there are the splicing factor genes *PRPF3*, *PRPF8*, *PRPF31*, *PAP1* and *ASCC3L1*. All gene products are involved in splicing, a genetic modification mechanism occurring in every cell, and they assist the assembly of the spliceosome. As described previously, the spliceosome is complex machinery, therefore, it is likely that these mutant proteins affect the whole mechanism but due to the physiology of the retina, they are disadvantageous only for that tissue. The challenge is to uncover their roles in the rod photoreceptor cells that cause this specific phenotype. The hypotheses that can be put forward to explain RP are 1) haploinsufficiency, that is the expression of only one wild type allele of the gene is not sufficient to fulfil the metabolic requirements; 2) a retina-specific (splicing) factor interacting with the aforementioned gene products; 3) the splicing factors mutated in RP may have a different, unknown role in retina, whose function is disrupted if a mutation occurs in the splicing factor gene. The first hypothesis also opens to the speculation that a lower

amount of PRPF31 may lead to less splicing of a gene particularly important for retina, and very susceptible to any splicing alteration such that it causes retinal degeneration.
As the research work included in this project focuses on *PRPF31*, its characteristics are described first, followed briefly by each of the other splicing factor genes.

1.6.1 PRPF31 (Pre-mRNA Processing Factor 31)

Mutations in splicing factor genes account for ~12% of adRP cases, 5% of which are due to *PRPF31*, second only to the rhodopsin gene (30% of adRP cases; Hartong *et al.*, 2006). The RP11 locus was first identified in our laboratory in a large English family, known as ADRP5 (Al-Maghtheh *et al.*, 1994). Later in 2001, Vithana and colleagues discovered *PRPF31* as the gene responsible for RP11. *PRPF31* maps to chromosome 19q13.4, spanning approximately 18 kb of DNA (comprising 14 exons), and encoding a 499 amino acid long protein (~61 kDa). cDNA expression profiles showed that the splicing factor gene is expressed in a wide range of tissues (Vithana *et al.*, 2001). To date 42 mutations have been reported and include missense substitutions, splice-site mutations, deletions and insertions (see following Table 1.2; Frio *et al.*, 2009; Rio Frio *et al.*, 2008b and references herein). Most of these mutations generate an unstable pre-mRNA, which fails to be processed to mature mRNA through nonsense mediated decay (NMD); the remaining mutations result in production of truncated protein.
A characteristic clinical feature of *PRPF31* is the non penetrance of the symptoms showed by carriers of the pathogenic allele. This form of incomplete penetrance was firstly reported to be linked to *RP11* locus by Al-Mahghteh and colleagues in 1994, and was later confirmed by Vithana and colleagues (2001) at the time *PRPF31* was identified. These asymptomatic obligate carriers are difficult to identify since they never develop a significant phenotype, but may have affected parents and affected offsprings (Vithana *et al.*, 2001). The incomplete penetrance may be due to wild type isoalleles, inherited from the non-affected parents; *PRPF31* isoalleles then could be differentially expressed and not produce any appreciable phenotype, but when inherited together with a pathogenic allele may modulate its pathogenic effects.
Past investigations (Vithana *et al.*, 2003, Abu-Safieh *et al.*, 2006, Rivolta *et al.*, 2006, Rio Frio *et al.*, 2008b) have provided data which suggests that haploinsufficiency is the

Table 1.2. Known mutations in the *PRPF31* gene.

n	Exon / Intron	Mutation
1	15 kb usptream *PRPF31*	44.8 kb deletion
2	Exon 1 / Intron 1	110bp del+39bp intron1, ins 640bp
3	Intron 1	IVS1+1G>T
4	Exon 2	c.79G>T
5	Exon 2 / Intron 2	IVS2+1G>A
6	Intron 2	IVS2+1delG
7	Exon 3	c.220C>T
8	Exon 4	c.319C>G
9	Exon 4 to exon 8	4843 bp deletion
10	Exon 4 to exon 13	11.3 kb deletion
11	Intron 4	c.323-2A>G (IVS-2A>G)
12	Exon 5	c.330-342del
13	Exon 5	c.390delC
14	Exon 5	c.412C>A
15	Intron 5	IVS5-1G>A
16	Exon 6 / Intron 6	c.522-527del, IVS6+1 to +10del
17	Intron 6	IVS6+1G>T
18	Intron 6	c.527+3A>G (IVS6+3A>G)
19	Intron 6	IVS6-3 to -45del
20	Intron 6	c.528-1G>A
21	Exon 7	c.421G>T
22	Exon 7	c.580-581dup33bp
23	Exon 7	c.581C>A
24	Exon 7	c.636delG
25	Exon 7	c.646G>C
26	Exon 8	c.758_767del
27	Exon 8	c.732-737delins20bp
28	Exon 8	c.769-770insA
29	Exon 8	c.785delT
30	Exon 8	c.828-829delCA
31	Intron 8	IVS8+1G>C
32	Intron 8	c.856-2A>G (IVS8-2A>G)
33	Exon 9	c.877_910del
34	Exon 10	c.973G>T
35	Exon 10 / Intron 10	c.1049_IVS10+20delins
36	Exon 10 / Intron 10	IVS10+1G>A
37	Exon 11	c.1115-1125del
38	Exon 11	c.1142delG
39	Exon 11 / Intron 11	c.1146+2T>C
40	Exon 12	c.1155to1159delins
41	Intron 13	c.1374+654C>G
42	Upstream exon 14	30 kb deletion

likely cause of RP11, as the large majority of mutations in *PRPF31* result in null alleles. Hence, a non-mutant isoallele which allows for a higher expression may cover up the lack of expression due to the pathogenic allele. Recently, Rio Frio and coworkers (2008a) have investigated possible genetic elements determining the variable expression of *PRPF31*: they hypothesised two *trans*-acting elements: 1) a currently unidentified product within a 8.2 Mb region on chromosome 14q (identified through linkage analysis), supposed to regulate the expression of both *PRPF31* alleles; 2) *RP11*-associated isoalleles could act *in cis* and *in trans*, increasing the levels of pre-mRNA from both alleles.

PRPF31 codes for a protein involved in the assembly of the U4/U6•U5 tri-snRNP, contributing to the processing of pre-mRNA molecules into mature RNA. Initial studies reported that PRPF31 shares 25% identity and 60% similarity with Prp31p, a pre-mRNA splicing factor from *S. cerevisiae* (Makarova *et al.*, 2002), hence suggesting that the two proteins are orthologues. Yeast Prp31p is associated with the tri-snRNP U4/U6•U5 (Weidenhammer *et al.*, 1997), and is also involved in the assembly of the tri-snRNP (Makarova *et al.*, 2002). PRPF31 can be divided into three regions which show similarity with the following domains, from N-terminus to C-terminus (Figure 1.17): 1) a NOSIC domain (aa 92-144 in PRPF31), which is usually found in Nop56/SIK1-like proteins; 2) a NOP domain (aa 186-334 in PRPF31) which shows RNP binding activity; 3) a the C-terminal Nuclear Localisation Signal sequence (aa 351-364; Deery *et al.*, 2002).

Two roles have been assigned to PRPF31: one is to bridge U4 and U6 together with PRPF3 (Nottrott *et al.*, 2002), the other is to tether U5 to the U4/U6 di-snRNP, by interacting with the U5-associated protein PRPF6.

In fact, Makarova and co-workers (2002) showed that immunodepletion of PRPF31 in HeLa cell nuclear extracts stalled the spliceosome complex, and the mature complex B was not formed, halting the assemblage at complex A. Moreover, subsequent addition of PRPF31 restored the spliceosomal assembly and therefore the splicing activity. It was postulated that PRPF31 acted as a bridge between the U5 snRNP and the U4/U6 di-snRNP: in fact, PRPF31 interacts with PRPF6, which associates with U5 snRNP, allowing PRPF31 to join the di-snRNP to the U5 snRNP. Liu and colleagues (2006)

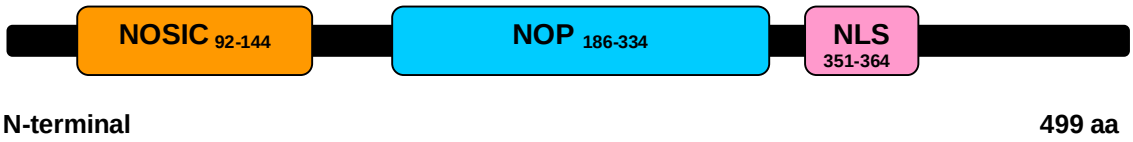


Figure 1.17. PRPF31 protein domain structure.
The NOSIC (central domain in Nop56/SIK1-like proteins), NOP (snoRNA binding domain), and NLS (nuclear localisation signal) domains are indicated along the sequence, including their aminoacidic position.

demonstrated that the NOP domain is involved in RNP recognition and binding, but not in RNA binding, showing that PRPF31 associates with a snRNP, and not with snRNA alone, *i.e.* the surface contacting the RNP interacts with 15.5K protein and RNA, but it fails to bind one of them alone (Liu *et al.*, 2006).

1.6.2 The U4/U6•U5 tri-snRNP

A core component of the U2-type spliceosome is the so called tri-snRNP, formed by three snRNPs: U4, U5 and U6. The human tri-snRNP is a protein-rich particle, comprising at least 30 different proteins (Liu *et al.*, 2006).

Before joining the spliceosomal complex B, the mature tri-snRNP must form: both U4 and U6 RNAs are base-paired forming two intermolecular RNA helices, thus the two snRNP are tightly associated, structuring the U4/U6 di-snRNP. Later this unit is joined by U5, forming the U4/U6•U5 tri-snRNP (Makarova *et al.*, 2002). The formation of the tri-snRNP particle occurs in the Cajal Bodies (CBs): mono-snRNPs are shuttled to CBs, where first the association between U4 and U6 snRNP occurs. Subsequently the di-snRNP is joined by the U5 snRNP, and the mature particle is then ready to be released from the CBs (Stanek & Neugebauer, 2006).

Liu and colleagues (Liu *et al.*, 2006) reported several interactions between the proteins associated to the tri-snRNP, drawing a network and cross-links among them (Figure 1.18). Among the proteins binding specifically to the U4/U6 particle there are PRPF31, PRPF3 and PAP1, whereas PRPF8 and PRPF6 are specific to U5. U4 and U6 are bridged by PRPF31 and PRPF3 (Nottrott *et al.*, 2002); little is known about interactions between U4/U6 and U5 snRNPs, apart from the unique interaction found between the U4/U6-specific PRPF31 and the U5-associated PRPF6 (Makarova *et al.*, 2002). Liu and colleagues (2006) also showed, through yeast two-hybrid and *in vitro* binding assays, that PRPF6 interacts with some portions of PRPF8 and with PRPF3, and also with itself, suggesting that PRPF6 can dimerise. Therefore PRPF6 may be a key player in linking the di-snRNP with the U5 snRNP.

Immunoprecipitation studies demonstrated an interaction between PAP1 and PRPF3 (Maita *et al.*, 2005); however PAP1 was mostly located free from any complex, and bound only in little amount with U4/U6 and U4/U6•U5, whereas PRPF3 predominantly

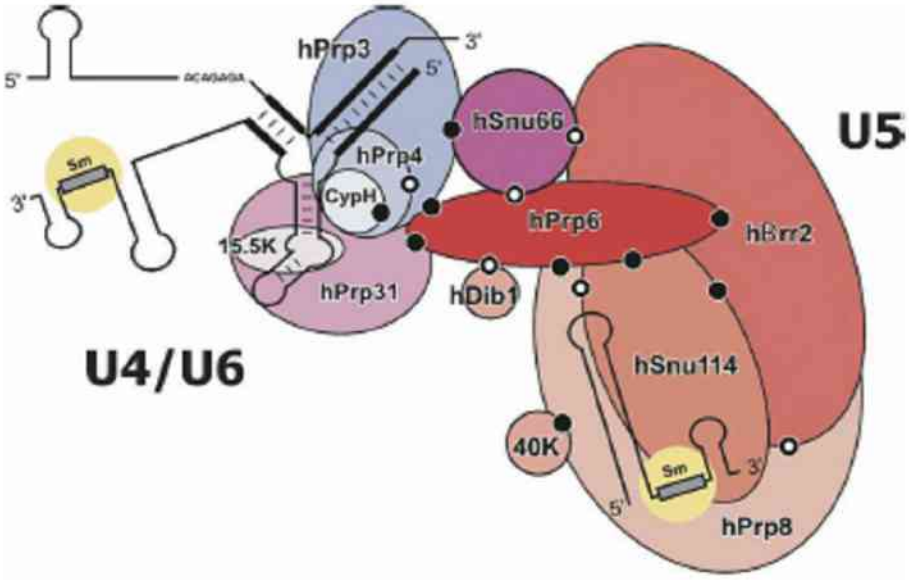


Figure 1.18. The human tri-snRNP particle.
Black dots represent observed interactions between the protein species (black and white dots represent interactions observed in yeast othologue proteins). For a reference, see Liu *et al.*, 2006, from which the image was taken.

associated with U4/U6 and poorly with U4/U6•U5. Hence, PAP1 should be associated with U4/U6, although it may be regarded as a non essential unit of the spliceosome, likely to be a modulator of splicing.

So far no RNA-RNA interactions have been reported between U4/U6 and U5, only RNA-protein links, suggesting that the whole tri-snRNP assembly is coordinated via protein-protein interaction. This network remains to be fully ascertained.

1.6.3 Other Splicing Factors Implicated in adRP

Four other splicing factors genes have been identified to cause adRP. These are *PRPF3* (Chakarova *et al.*, 2002), *PRPF8* (McKie *et al.*, 2001), *PAP1* (Keen *et al.*, 2002) and *ASCC3L1* (Li *et al.*, 2009). Like *PRPF31*, these genes are widely expressed in many tissues.

PRPF3 comprises 16 exons and encodes for a 683 aa-long protein, of approximately 77 kDa (Ensembl Genome Browser, December 2008). Along with PRPF31, PRPF3 splicing factor is involved in the assembly and the correct functioning of the tri-snRNP; in fact, PRPF3 is known to interact with PAP1 (Maita *et al.*, 2005), PRPF4, U4/U6 snRNA, as well as U5-associated PRPF6 (Gonzalez-Santos *et al.*, 2008). All these associations are mediated by its C-terminal region, which is highly conserved among numerous species. To date only three mutations have been found, which lead to three amino acid changes, respectively P493S, T494M (Chakarova *et al.*, 2002) and the more recent A498D (Gamundi *et al.*, 2008). All changes are located within the C-terminal part of the protein. It is suggested that the C-terminal region is responsible for a crucial function, which may have a key responsibility in spliceosome assembly (Mordes *et al.*, 2007). Gonzalez-Santos and colleagues proposed that the latter mutation affects the splicing, as a T to M change alters the phosphorylation state of the protein, likely modulating the splicing by destabilising the spliceosome complex. The other two mutations may disrupt the recognition motif which leads to the phosphorylation of the threonine 494.

PRPF8 is the splicing factor gene responsible for RP13. *PRPF8* comprises 43 exons and encodes a large protein of 2335 residues, of an approximate molecular weight of 220 kDa (Ensembl Genome Browser, December 2008). Various point mutations lead to single amino acid changes, and one mutation is a 6bp deletion with an 11bp insertion; all

mutations have been found in exon 42. PRPF8 is a component of the U5 snRNP and of the U4/U6•U5 tri-snRNP, and also interacts with the pre-mRNA molecule. PRPF8 is present in different complexes in the spliceosome dynamics, as it is found associated with the complex containing reaction intermediates, or the excised intron; PRPF8 also associates with BRR2, an unwindase implicated in the unwinding of the base-pairing between U4 and U6 snRNAs (Grainger & Beggs 2005). Hence, as PRPF8 harmonises several associations between spliceosomal proteins and snRNAs, it is thought to adjust the coordination of the spliceosome dynamics. Given that the mutations are clustered in exon 42, the corresponding protein region should be the core centre for organising the network of associations of PRPF8, thus preventing the protein to perform its function.

RP9 is caused by mutations in *PAP1* gene. *PAP1* comprises six exons and encodes a small protein of 181 residues, of an approximate molecular weight of 21 kDa (Ensembl Genome Browser, December 2008). To date, only two point mutations have been identified, both altering the protein sequence (H137L and D170G; Keen *et al.*, 2002). PAP1 protein is also involved in the tri-snRNP formation, as reported by Maita and colleagues (2005), and it interacts with the C-terminal region of PRPF3. However, none of the two mutations affect the binding with PRPF3. The mechanism of RP9 remains to be elucidated with further investigations.

Finally, recently Li and colleagues (2009) reported a single base pair change in the *ASCC3L1* gene, resulting in an amino acid change (R1090L) in the protein sequence. *ASCC3L1* comprises 45 exons and encodes the BRR2 protein, a 200 kDa RNA helicase involved in the unwinding of U4/U6 preceding the activation of the spliceosome.

1.7 Aim of the Project

The aim of this project is to investigate the molecular mechanisms underlying the retinitis pigmentosa 11 (RP11) locus. RP11 is caused by mutations in the *PRPF31* gene, which encodes a ubiquitously expressed splicing factor; however, the mutant phenotype is only apparent in the eye with no other symptoms. The hypotheses to explain RP11 could be 1) haploinsufficieny, or 2) a PRPF31 retina-specific partner, or 3) that PRPF31 has a different role in retina. The main focus of this project is to test the second

hypothesis, aiming to the identification of retina-specific interacting partners of the PRPF31 splicing factor, elucidating its function, and perhaps shedding light on the pathogenic mechanism(s) of RP11.

Since investigations on splicing and its protein network are conducted mainly in *S.cerevisiae* and do not allow for tissue specificity, this approach may provide a different perspective other than yeast. This rationale follows the steps of the finding of RPGRIP as an interacting partner of the ubiquitous RPGR: *RPGR* is involved in xlRP and *RPGRIP* was later discovered as involved in Leber congenital amaurosis (Roepman *et al.*, 2000a). Therefore it may be that also the ubiquitous PRPF31, like RPGR, may have a tissue-specific partner: in fact, the RPGR-interactig protein RPGRIP is expressed in testis and retina only (Boylan & Wright, 2000).

The yeast two-hybrid screen is then to be followed up by confirmation of these interacting partners using different methodologies, in addition to studies regarding PRPF31 localisation in both cell lines and retinal tissue to better understand the role of PRPF31 in the retina and retinal degenerations.

2. Materials and Methods

2.1 Materials

All materials and reagents were purchased from Analar, Applied Biosystems, BD Biosciences, Bio-Rad, Clontech, Eurogentec, Fisher Scientific, GE Healthcare, Gibco BRL, Jackson Immunoresearch Labs, Invitrogen, Merck, National Diagnostics, Nunclon, Promega, Sigma, or VWR. All the reagents for cell culture were purchased from Invitrogen.

2.2 Reagents and Buffers

Table 2.1. Summary of all solutions, buffers and gels used

Reagent	Composition
Agarose Gel Electrophoresis	
TAE buffer 1x	2M Tris-Acetate, 0.05M EDTA (50x from Eppendorf)
Gel Loading Dye 10x	0.14% Bromophenol Blue, 0.16% Orange G, 0.16% Xylene Cyanole FF, 15% Ficoll 400
Bacteria	
TSB medium	10% PEG, 5% DMSO, 10mM $MgCl_2$, 10mM $MgSO_4$; in LB medium
KCM Buffer 5x	0.5M KCl, 0.15M $CaCl_2$, 0.25M $MgCl_2$;
Yeast	
TE/LiAc 1.1x	1.1x TE, 0.11M LiAc (10x TE and 1M LiAc from BD Biosciences)
PEG/LiAc	1x TE, 0.1M LiAc, 40% PEG 3350 (10x TE, 1M LiAc, 50% PEG 3350 from BD Biosciences);
Buffer S	10mM KPO_4 pH7.2, 10mM EDTA pH 8.0, 50mM β-mercaptoethanol, 200µg/ml lyticase (SIGMA);
Lysis Buffer	0.25M Tris-HCl pH 7.5, 25mM EDTA pH 8.0, 2.5% SDS
Z-Buffer	60mM Na_2HPO_4, 40mM NaH_2PO_4, 10mM KCl, 1mM $MgSO_4$
Cracking Buffer Stock solution	8M Urea, 5% w/v SDS, 40mM Tris-HCl pH 6.8, 0.1mM EDTA, 0.4mg/ml Bromophenol Blue
X-α-Gal (Glycosynth)	20mg/ml in DMF
X-β-Gal (SIGMA)	20mg/ml in DMF
SDS-PAGE and Western Blot	
Acrylamide Stacking gel	3.75% acrylamide, 125mM Tris-HCl pH 6.8, 0.1% SDS, 0.06% APS, 0.2% TEMED
Acrylamide Resolving gel	7.5 -15% acrylamide, 375mM Tris-HCl pH 8.8, 0.1% SDS, 0.06% APS, 0.16% TEMED

(continued from previous page)	
Reducing Sample Buffer 2x	125mM Tris-HCl pH 6.8, 20% Glycerol, 4% SDS, 10% β-mercaptoethanol, 0.005% Bromophenol Blue
Running Buffer 1x	25mM Tris, 192mM Glycine, 0.1% SDS (10x Tris/Glycine/SDS: 0.25M Tris, 1.92M Glycine, 1% SDS; from National Diagnostics)
Transfer Buffer	25mM Tris, 192mM Glycine, 20% Methanol
TBS 10x	200mM Tris Base, 1370mM NaCl, pH 7.6
TBS-T	0.1% Tween-20 in TBS 1x
Amido Black staining	0.1% amido black 10B in destaining solution
Destaining solution	7.5% acetic acid, 20% methanol
Blocking Buffer I	5% non fat dry milk, 0.1% NaN_3 in TBS-T
Blocking Buffer II	5% BSA, 0.1% NaN_3 in TBS-T

Co-immunoprecipitation, Immunocytochemistry (ICC), Immunohystochemistry (IHC)	
PBS-TX100, 0.1/0.5%	0.1/0.5% Triton X-100 in PBS 1x (OXOID; tablets)
Protease and Phosphatase inhibitors cocktail (final concentrations)	1mM PMSF (from 100mM in isopropanol); 10µg/ml Aprotinin; 50µg/ml Benzamidin; 10µg/ml Pepstatin; 10µg/ml Leupeptin; 50mM NaF; 3mM Na-pyrophosphate; 1mM Na-orthovanadate (Na_3VO_4)
HEPES	0.5M pH 7.4
Blocking solution (for ICC)	0.5% BSA, 20mM Glycine, 0.1% NaN_3 in PBS 1x
Blocking solution I (for IHC)	5% Goat Serum, 0.1% NaN_3, in 0.1% PBS-TX100
Blocking solution II (for IHC)	2% Goat Serum, 0.1% NaN_3, in 0.1% PBS-TX100

All solutions, buffers and gels were made using distilled water unless otherwise stated.

Table 2.2. Summary of the different media used for bacteria and yeast

Bacteria Culturing Media	
LB agar	25g of LB powder in 1L, 20% agar
LB broth	25g of LB powder in 1L

Yeast Culturing Media	
YPAD agar	50% YPD, 0.04% Adenine Hemisulfate salt, 20% agar
YPAD broth	50% YPD, 0.04% Adenine Hemisulfate salt
SD Synthetic Minimal agar	6.7% Yeast Nitrogen Base, 18.2% D-Sorbitol, 2% Glucose, 20% agar
SD Synthetic Minimal broth	6.7% Yeast Nitrogen Base, 18.2% D-Sorbitol, 2% Glucose
Dropout Solutions	To render the medium selective, both SD agar and broth were added Dropout Solutions (BD Biosciences), lacking the appropriate aminoacids.

All media were made using distilled water and autoclaved.

2.3 Vectors

All the cloning work was performed using the Gateway Technology developed by Invitrogen.

2.3.1 Entry Vectors

pDONR201, pDONR/Zeo (Invitrogen): entry vectors with *att*P1 and *att*P2 sequences, kanamycin and zeocin resistance respectively; these vectors were used as shuttle vectors for subcloning into destination vectors (Yeast Two-Hybrid vectors and Mammalian Expression vectors, see following);

2.3.2 Yeast Two-Hybrid Vectors

1. pAD-GAL4-Amp-DEST*: GAL4 Activation Domain destination vector with *att*R1 and *att*R2 sequences; LEU2 marker, ampicillin resistance;
2. pBD-GAL4-Cam-DEST*: GAL4 DNA-Binding Domain destination vector with *att*R1 and *att*R2 sequences; TRP1 marker, chloramphenicol resistance;
3. pAD-WT (Stratagene): AD-GAL4 and aa 132–236 of wild type fragment C of lambda cI repressor, ampicillin resistance;
4. pBD-WT (Stratagene): BD-GAL4 and aa 132–236 of wild type fragment C of lambda cI repressor, chloramphenicol resistance;
5. pAD-MUT (Stratagene): AD-GAL4 and aa 132–236 of E233K mutant fragment of lambda cI repressor, ampicillin resistance;
6. pBD-MUT (Stratagene): BD-GAL4 and aa 132–236 of E233K mutant fragment of lambda cI repressor, chloramphenicol resistance;
7. pBD-pLC (Stratagene): BD-GAL4 and aa 67–230 of human lamin C, ampicillin resistance;

All the vectors listed in section 2.3.2 were a gift of Dr R. Roepman;

*original Stratagene vectors modified with *att* sequences to be used with the Gateway Technology.

2.3.3 Mammalian Expression Vectors

1. pDEST26 (Invitrogen): destination vector with *att*R1 and *att*R2 sequences, N-terminal 6xHis tag, ampicillin resistance;
2. pcDNA3.1V5DEST (Invitrogen): destination vector with *att*R1 and *att*R2 sequences, N-terminal V5 epitope tag, ampicillin resistance;

2.4 Cells

A number of different cell types were used throughout this project, details of the bacterial and yeast strains are as following.

2.4.1 *E.coli* Host Strains

1. AG1: *recA1, endA1, gyrA96, thi-1, hsdR17, (rk–, mk+), supE44, relA1* strain used for maintenance and propagation of plasmid DNA; maintaned in LB agar/broth without any selection; gift from Dr R. Roepman;
2. Library Efficiency DH5α Competent Cells (Invitrogen): F- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(rk-, mk+) *pho*A*sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ- strain used for maintenance and propagation of DNA; the manufacturer's transformation protocol was followed;
3. Library Efficiency DB3.1 Competent Cells (Invitrogen): F- gyrA462 endA1 Δ(sr1-recA) mcrB mrr hsdS20(rB-, mB-) supE44 ara-14 galK2 lacY1 proA2 rpsL20(SmR) xyl-5 λ- leu mtl1, ccdB-resistant strain used to propagate entry vectors and destination vectors; manufacturer's transformation protocol was followed;

The transformants were maintained on LB agar/broth, with antibiotic selection according to the genotype of the vector. The final antibiotic concentrations were as follows:

- ampicillin (Sigma) 100μg/ml;
- chloramphenicol (Sigma) 170μg/ml;
- kanamicin (Sigma) 50μg/ml;
- zeocin (Invitrogen, InvivoGen) 30μg/ml.

2.4.2 Yeast Host Strain

S.cerevisiae PJ69-4A strain (James *et al.*, 1996) was a gift from Dr R. Roepman, and used for all the yeast two-hybrid procedures.

- **PJ69-4A**: *MAT*α *ura3-52 his3-*Δ*200 leu2-3,112 lys2-801 trp1-901 gal4*Δ *gal80*Δ *GAL2- ADE2 LYS2::GAL1-HIS3 met2::GAL7-lacZ*
- **PJ69-4A**: *MAT*a *ura3-52 his3-*Δ*200 leu2-3,112 lys2-801 trp1-901 gal4*Δ *gal80*Δ *GAL2- ADE2 LYS2::GAL1-HIS3 met2::GAL7-lacZ*

2.4.3 Mammalian Cell Lines

(for culturing methods, see section 2.12.1)

1. HEK293 (personal communication): used for immunoprecipitation study;
2. MDCK (personal communication): used for protein localisation study.
3. SK-N-SH (personal communication): used for protein localisation study.

2.5 Antibodies

2.5.1 Primary Antibodies

1. Anti-GAL4 DNA-AD: mouse monoclonal antibody, against the activation domain of the yeast GAL4 protein (Clontech);
2. Anti-GAL4 DNA-BD: mouse monoclonal antibody, against the DNA-binding domain of the yeast GAL4 protein (Clontech);
3. Anti-6x His tag: rabbit polyclonal antibody, against 6x His tag (Abcam);
4. Anti-6x His tag: mouse monoclonal antibody, against 6x His tag (clone HIS-1, ascites fluid; Sigma);
5. Anti-V5: mouse monoclonal antibody, against the V5 epitope (Invitrogen).
6. Anti-PRPF31$_{484\text{-}497}$: rabbit polyclonal antibody raised against peptide 484-497 (AEFLKVKGEKSGLM) of human PRPF31 protein; kindly donated by Prof R. Lührmann (Makarova *et al.*, 2002);
7. Anti-PRPF31$_{140\text{-}154}$: goat polyclonal antibody raised against peptide KELGNSLDKCKNNEN, corresponding to internal sequence amino acids 140-154 of Human PRPF31 (Abcam);
8. Anti-PRPF6: goat polyclonal antibody, against a peptide mapping near the N-terminus of PRP6 of human origin (Santa Cruz Biotechnology);
9. Anti-GRIF1$_{8\text{-}633}$: rabbit polyclonal antibody, against peptide 8-633 of the rat GRIF1 protein, orthologue to human TRAK2 (Beck *et al.*, 2002); kindly donated by Prof A. Stephenson;
10. Anti-MAGI3: purified mouse anti-MAGI3, clone 46/MAGI3 (immunogen: recombinant protein human MAGI3 aa 194-298; BD Transduction Laboratories);
11. Anti-E-Cadherin: mouse monoclonal antibody against Human E-Cadherin, clone 36 (immunogen: C-terminal recombinant protein; BD Transduction Laboratories);

12. Anti-PNA: biotinylated anti-peanut agglutinin (Vector Laboratories, Burlingame, CA);
13. Anti-Rhodopsin: mouse monoclonal antibody 1D4 (National Cell Culture Center, Minneapolis, MN);
14. Anti-SC-35: mouse monoclonal antibody to SC-35, nuclear speckle marker (immunogen: partially purified mammalian spliceosomes; Abcam);
15. Anti-TOPORS: Mouse monoclonal antibody raised against a partial recombinant TOPORS (M01, clone 5G11; Abnova);
16. Anti-TRAK2: goat anti ALS2CR3/GRIF1 antibody, raised against the peptide CTSSPKMGVLKED (Everest Biotech);

2.5.2 Secondary Antibodies

Secondary antibodies were purchased from Jackson ImmunoResearch, except for Cy2-conjugated streptavidin (Molecular Probes). HRP-conjugated anti-rabbit and anti-mouse were raised in goat; HRP-conjugated anti-goat was raised in rabbit. Cy3-conjugated anti-goat, FITC-conjugated anti-rabbit and anti-mouse were raised in donkey. HRP-conjugated antibodies were diluted 1:10000 in TBS-T; FITC-, Cy3- and Cy2-conjugated antibodies were diluted 1:800 in blocking solution (for recipe of immunocytochemistry blocking solution and immunohistochemistry blocking solution II, see Table 2.1).

2.6 Retina cDNA Libraries

The bovine and human retinal cDNA libraries used for the yeast two-hybrid screening were a gift from Dr R. Roepman. Each cDNA library was custom-made (Roepman *et al.*, 2000b), constructed to be fused to the yeast GAL4 activation domain (AD, N-terminal) and transformed in the *S.cerevisiae* PJ69-4A strain, *MATα*. The human library (titre: 5.2×10^9 cells/ml) is oligo-dT primed, with 1.9×10^6 independent clones; the bovine library (titre: 4.6×10^9 cells/ml) is randomly primed and has 4.5×10^6 independent clones.

2.7 Full Length cDNA Clones

Full length cDNA of human wild type *PRPF6*, *TRAK2* and *ZNF143* were purchased from RZPD, sequence verified by the company (image IDs respectively: 2988165 in

pOTB7 vector; 4814594 in pBluescriptR vector; 4902668 in pOTB7 vector); human wild type *PRPF31* cDNA was obtained from previous work (Deery *et al.*, 2002); human wild type *MAGI3* cDNA was cloned in our lab (see section 2.9.7.2 of this chapter and section 3.8.1 of Chapter 3).

2.8 Patients

Mutation screening was carried out on a panel of 95 unrelated autosomal dominant retinitis pigmentosa (adRP) patients. Informed consent was obtained from each patient, in accordance with guidelines established by the Declaration of Helsinki and was approved by the Moorfields' Eye Hospital Ethics Committee. Genomic DNA was provided to by Naushin Waseem and Beverley Scott.

Allele frequencies were calculated as follows: if the population consists of x individuals, then the alleles (A and a) for a given locus are $2x$ (x individuals × 2 alleles at that locus per individual); AA homozygous individuals are multiplied by 2 to calculate the number of A alleles, so the total of A alleles is = (Aa heterozygotes + 2AA homozygotes), the A allele frequency is = (total of A alleles / total number of alleles in population); the frequency of the a allele is = (1- frequency of A allele).

Genotype frequencies were calculated as follows: the number of individuals with each genotype is determined; then to determine the frequencies, the number of individuals for each genotype is divided by the total number of individuals in the population (for example: frequency of AA genotype = total of AA individuals / population size).

Assuming absence of selection, the genotype frequencies can also be calculated according to the Hardy-Weinberg equilibrium:

$$(p + q)^2 = 1$$

where p = frequency of allele A and q = frequency of allele a. From the binomial expansion of the above equation,

$$p^2 + q^2 + 2pq = 1$$

one can derive the genotype frequencies: p^2 = frequency of genotype AA, q^2 = frequency of genotype aa, $2pq$ = frequency of genotype Aa. If the genotype frequencies calculated from a real population do not meet those predicted by the Hardy-Weinberg equilibrium, then evolution changes are in operation.

2.9 General Methods

2.9.1 Polymerase Chain Reaction (PCR)

Polymerase Chain Reaction was used in several steps along the project, using different polymerases, according to the purpose. Oligonucleotides for PCR reactions and sequencing were designed manually and purchased from Sigma. They were designed to be 20-30 bp long, with a 40-60 % GC content; the melting temperature was calculated according to the formula: $T_m = 2(A + T) + 4(G + C)$. Biotaq polymerase (Bioline), supplemented with *Pfu* polymerase (Promega) when needed, was used to amplify small fragments (up to 1kb usually); KOD Hot Start DNA Polymerase (Novagen) and Expand Long Template PCR System polymerase (for ELTPS PCRs, of which only buffer #2 was used, a buffer 10x concentration, with 22.5mM $MgCl_2$ –datasheet May 2000) were used to amplify longer fragments and for high fidelity. The Tables below show the PCR mix (Table 2.3) and the cycling conditions (Tables 2.4, 2.5 and 2.6) of each enzyme used; PCRs were normally carried out in a volume of 25µl, using ~1µg of DNA.

Table 2.3. PCR parameters for each of the polymerases used throughout this project.

	Biotaq	KOD	ELTPS
dNTPs	0.2mM	0.2mM	0.7mM
Mg	2-2.5mM	1mM	2.25mM
Buffer 10x	1x	1x	1x
Primer F	0.25µM	0.3µM	0.25µM
Primer R	0.25µM	0.3µM	0.25µM
Enzyme	0.016 U/µl (0.4U)	0.02 U/µl (0.5U)	0.08 U/µl (4U)
DNA	1µl	0.5µl	2µl
Water	up to final volume	up to final volume	up to final volume
Final volume	25µl	25µl	50µl

Table 2.4. Thermocycling conditions for BIOTAQ Polymerase.

BIOTAQ polymerase	Temperature	Time	Cycles
Initial Denaturation	95°C	5 minutes	-
Denaturation	95°C	15-30 seconds	30-40
Annealing	Ta °C	30-60 seconds	
Elongation	72°C	up to 4 minutes	
Final Elongation	72°C	5-10 minutes	-
Hold	4°C	∞	-

Table 2.5. Thermocycling conditions for KOD Polymerase.

KOD polymerase	Temperature	Time	Cycles
Initial Denaturation	95°C	2 minutes	-
Denaturation	95°C	15 seconds	35-40
Annealing	60°C	30 seconds	
Elongation	72°C	up to 2 minutes	
Final Elongation	72°C	10 minutes	-
Hold	4°C	∞	-

Table 2.6. Thermocycling conditions for ELTPS polymerase.

ELTPS polymerase	Temperature	Time	Cycles
Initial Denaturation	94°C	2 minutes	-
Denaturation	94°C	10 seconds	35
Annealing	59°C	30 seconds	
Elongation	68°C	up to 2 minutes	
Final Elongation	68°C	7 minutes	-
Hold	4°C	∞	-

2.9.2 Agarose Gel Electrophoresis

DNA molecules of various molecular weights were separated on agarose gels, for either visualisation, quantification or for subsequent gel extraction. Gels were prepared by dissolving 0.8-2% agarose (Sigma) in 1x TAE buffer (Eppendorf); DNA was stained with 0.5µg/ml ethidium bromide, added directly to the dissolved agarose before pouring it into a sealed casting plate containing a comb. Samples (0.5-5 µl of DNA) were loaded together with a 1x gel loading dye; a molecular weight marker (Smart Ladder,

Eurogentec) was also included to allow the sizing and quantification of samples. The gel was placed into an electrophoresis tank (Biorad), filled in with 1x TAE buffer, and electrophoresis was carried at 80-120V until the desired separation was achieved. DNA was visualised on a UV transilluminator (UVP).

2.9.3 Restriction Enzyme Digest

Plasmid DNAs were digested to confirm the expected digestion pattern after cloning. Digestion mixtures contained 2µl of 10X appropriate buffer, 1µg of DNA, 1.0-2.0 units of the appropriate enzyme and distilled water to a final volume of 20µl. The reactions were then incubated for 1 hour at the specified temperature. All enzymes and buffers were purchased from Promega and used according to the manufacturer's instruction.

2.9.4 DNA Purification

1. From PCR reactions: PCR products were purified from excess dNTPs and primers using ExoSAP-IT (USB); ExoSAP-IT consists of two enzymes, Exonuclease I and Shrimp Alkaline Phosphatase. Amplified PCR product (3µl) was mixed with 1µl of ExoSAP-IT and made up to a final volume of 25 µl with distilled water. The reaction was then incubated at 37°C for 15 minutes, followed by 80°C for 15 minutes to deactivate the enzyme. The product was then used for the downstream sequencing reactions;
2. From agarose gel: the desired DNA samples were excised from the gel, dissolved and purified using the Illustra™ GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare), following manufacturer's instructions;
3. From *E.coli* cultures: single colonies were grown overnight in 5ml selective LB; cells were collected and subjected to a series of steps to isolate the DNA, using GenElute Plasmid Miniprep Kit (Sigma)", following manufacturer's instructions. Alternatively, 200ml cultures were prepared and the DNA purified using "GenElute Plasmid Maxiprep Kit (Sigma).

2.9.5 DNA Sequencing

Samples underwent sequencing reactions, using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Each reaction was carried out in a final volume of 10µl, containing: 0.5-3µl of PCR product (100-300ng), 2.5µl of sequencing buffer, 0.5µM primers, 0.5µl of Big Dye 3.1; thermocycling conditions are shown in Table 2.7.

Table 2.7. Thermocycling conditions for sequencing reactions.

BigDye Terminator	Temperature	Time	Cycles
Initial Denaturation	95°C	3 minutes	-
Denaturation	96°C	10 seconds	25
Annealing	50°C	3 seconds	
Elongation	60°C	4 minutes	
Hold	4°C	∞	-

Unincorporated dye and salt was removed from the sequencing product using Sephadex G50 (Sigma). MultiScreen HTS Filter Plate (Millipore) was first filled with dry Sephadex G50; each well was then filled with 350µl water, and left to set for 30 minutes at room temperature. A collection plate was placed to the MultiScreen plate, and the plate was centrifuged for 5 minutes at 910g; 100µl of deionized water was again added to each well, and centrifuged again for 5 minutes at 910g. A fresh 96-well plate was placed underneath the MultiScreen plate before loading 20µl sequencing samples. The plate was centrifuged again at 910g for 5 min and the eluant was loaded on to ABI Prism 3730 DNA Analyzer.

2.9.6 Reverse Transcription (RT) PCR

Total retinal RNA (Clontech) was used as a template to prepare cDNA, using randomly primed oligonucleotides (Promega). As a precaution, prior to cDNA synthesis, 1 µg of total retina mRNA was incubated with DNase I Amplification Grade (Invitrogen) at room temperature, 15 minutes (10µl reaction); the DNase I was inactivated by adding 1µl of 25 mM EDTA and heated for 10 minutes at 65°C. The RNA sample was ready to be reverse transcribed, using Superscript II Reverse Transcriptase (Invitrogen),

following the manufacturer's instructions. The double stranded cDNA was then either used immediately as a template for PCR, or stored at –80^{o}C.

2.9.7 Cloning

2.9.7.1 Cloning via Recombination Reactions

Each clone was plated on selective LB agar and grown overnight at 37°C; single colonies were picked and cultured in 5ml selective LB, overnight. DNA was purified from the saturated culture. Oligonucleotides with *att*B overhangs were designed to amplify full length *PRPF31* (Table 2.8), *TRAK2* (Table 2.9), and *ZNF143* (Table 2.10) sequences, using the KOD or ELTPS polymerases. The three fragments of *PRPF31* were generated as follows: F1: 562 bp fragment encompassing the NOSIC domain; F2: 610 bp fragment encompassing the NOP domain; F3: 629 bp fragment encompassing the NLS domain (primers listed in Table 2.8). Purified PCR products were recombined in either pDONR201 or pDONR/Zeo to generate entry clones, and sequenced (primers listed in Table 2.11). The entry clones were then recombined into the desired destination vectors. Each new clone was first tested by restriction digestion followed by confirmation by sequencing (primers listed in Table 2.11) The *PRPF6* sequence was purchased already cloned into an entry vector, hence not requiring any preparatory amplification with *att*B overhung primers. The *PRPF6* vector was directly used in a recombination reaction, and sequenced (primers listed in Table 2.12). All the cloning steps described were done using the Gateway Technology from Invitrogen, using BP and LR recombinase mixes, following manufacturer's instruction.

The Gateway Technology is a rapid and efficient way to clone DNA sequences developed by Invitrogen, allowing use of multiple vector systems but maintaining orientation and reading frame, with no restriction/ligation of the DNA. This cloning method is based on site-specific recombination properties of bacteriophage lambda, a system which facilitates the integration of lambda into *E.coli* chromosome; the components of this system have been modified to improve specificity and efficiency. They include: the DNA recombination sequences (specific attachment sites, or *att* sites) and the proteins that mediate the recombination reaction. The recombination occurs

Table 2.8. Primers for *PRPF31* cloning.

PRPF31 att	Sequence (5' … 3')	Size bp
PRPF31-I_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAGTCTGGCAGATGAGCTC	562
PRPF31-I_R	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTAGACCATGATGGTGGCATTGG	
PRPF31-II_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAATGCCACCATCATGGTCGTC	610
PRPF31-II_R	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTACGGCTTCACCTGCTTCAC	
PRPF31-III_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGGGAAGGTGGGCTACGAAC	629
PRPF31-III_R	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATGTGGACATAAGGCCACTCTTC	
PRPF31-I_F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAGTCTGGCAGATGAGCTC	1568
PRPF31-III_R	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATGTGGACATAAGGCCACTCTTC	

For all primer pairs, the annealing temperature is 60°C, and Mg concentration is 1mM, using KOD polymerase.

Table 2.9. Primers for *TRAK2* cloning.

TRAK2 att	Sequence (5' … 3')	Size bp
TRAK2_F_full	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAGTCAATCCCAGAATGCA	2809
TRAK2_R_full	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATCAGTCCTCCTTCAGGACACC	

The annealing temperature is 59°C, and Mg concentration is 2.25mM, using ELTPS polymerase.

Table 2.10. Primers for *ZNF143* cloning.

ZNF143 att	Sequence (5' … 3')	Size bp
ZNF143_F_full	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGACAGAGTTTCCTGGAGGA	1945
ZNF143_R_full	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATTAATCATCCAACCCTGGCGT	

The annealing temperature is 59°C, and Mg concentration is 2.25mM, using ELTPS polymerase.

Table 2.11. Summary of the primers used to sequence the corresponding vectors.

Vector	Primer name	Sequence 5'... 3'
pDONR201	pDONR201_F	TCGCGTTAACGCTAGCATGGATCTC
	pDONR201_R	GTAACATCAGAGATTTTGAGACAC
pDONR/Zeo	M13F	GTAAACGACGGCCAG
	M13R	CAGGAAACAGCTATGAC
pAD-GAL4-Amp-DEST	pAD_BRP F	AAACCACTGTCACCTGGTTGG
	pAD_BRP R	GTGCACGATGCACAGTTGAAGT
pBD-GAL4-Cam-DEST	pBD-Dest_F	GGAGACTGATATGCCTCTAACAT
pDEST26	pDEST26_F	AGCTCGTTTGTGAACCGTCAG
pcDNA3.1V5DEST	T7	TAATACGACTCACTATAGGG

Table 2.12. Primers used to sequence *PRPF6*.

PRPF6	sequence (5' … 3')
PRP6 F1	CTCGTCATTGTTCTCGTTCCCT
PRP6 F2	AACCCACGCTATGAGAAGCTGA
PRP6 F3	AATGTGACAGGGCTGGGAGTGT
PRP6 R1	CAATGTCTGACCGCGTGGTGAAG
PRP6 R2	TAGATGGCTCGTGCACACTCCA

between *att* sites on the interacting DNA molecules, without gain/loss of nucleotides or synthesis of new DNA. The *att* sites are hybrid sequences made of sequences donated by each parental vector (i.e., *att*L sites consist of sequences from *att*B and *att*P sites). The proteins mediating this reaction are enzymes binding to the *att* sequences, which bring together the target sites, cleave and attach the DNA. BP reaction facilitates recombination of an *att*B substrate (*att*B-PCR product or a linearised *att*B expression clone) with an *att*P substrate (donor vector) to create an *att*L-containing entry clone; LR reaction facilitates recombination of an *att*L substrate (entry clone) with an *att*R substrate (destination vector) to create an *att*B-containing expression clone. Here, one example regarding the cloning of *PRPF31* sequence into the bait vector pBD-DEST is explained:

- the full length cDNA of *PRPF31* has been amplified, using the appropriate primers with matching *att*B site overhangs;
- through the BP reaction, a recombination between the *att*B-*PRPF31*_PCR_product and the donor vector pDONR201 (which has *att*P sites) occurs; the pDONR-*PRPF31* will now have *att*L sites and is named "entry" clone;
- through LR reaction, the new entry clone pDONR-*PRPF31* undergoes further recombination with the desired "destination" vector (for ex. pBD-DEST, the bait vector); the new expression clone pBD-*PRPF31* contains *att*B sites.

2.9.7.2 SPLICE Cloning

To clone *MAGI3*, we opted for a technique called SPLICE (Davies *et al.*, 2007), adapting it to suit our case. SPLICE is a rapid technique which allows intron removal from genomic DNA to generate a continuous coding sequence, and is based on an overlapping fusion-PCR strategy (swift PCR for ligating *in vitro* constructed exons; Davies *et al.*, 2007). This technique was adapted to amplify bits of coding region of *MAGI3*. Since MAGI3 presents 21 exons, retinal cDNA was chosen instead of genomic DNA, allowing for less rounds of amplification. According to the Ensembl database, *MAGI3* has four different splice variants and we have generated the variant MAGI3-004 (transcript ENST00000307546), hence primers were designed according to its sequence,

after having decided which DNA segments were to be amplified. External oligonucleotides were designed with *att*B overhangs (Table 2.13) to allow the downstream Gateway cloning (see previous section 2.9.7.1); and, internal primers were designed as follows: for forward primers, the first 15 nucleotides corresponded to the last 15 nucleotides of an upstream exon, and the last 15 nucleotides corresponded to the first 15 nucleotides of a downstream exon (Table 2.13); the reverse primers were exactly the reverse complement sequence of the forward primer. PCRs were performed using KOD polymerase; fragments were gel-extracted and purified. From the concentration of the amplified products, equimolar quantities were calculated* and added together for subsequent rounds of PCR to amplify longer fragments, including in the reactions only the outermost primers and increasing the quantity of polymerase. The fragment corresponding to the full length product underwent a "test" amplification with *att* primers, to make sure it was amplifiable by the flanking primers. Finally the product was sequenced (additional primers for sequencing are shown in Table 2.14), and used for downstream recombination.

* the mol/µl of each fragment were calculated dividing their mass (using the concentration data) by their molecular weight (the bp length); the results are normalised to the lowest and the final number represents the number of moles in 1 µl.

Table 2.13. Primers for *MAGI3* SPLICE cloning.

MAGI3	**sequence (5' … 3')**	**size**
MAGI3_F_full	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGTCGAAGACGCTGAAGAAGA	362 bp
MAGI3_ex1-2R	TATTAATGACTTTGCCTGGTTTCACAGTCT	
MAGI3_ex1-2F	AGACTGTGAAACCAGGCAAAGTCATTAATA	790 bp
MAGI3_ex7-8R	TTCTGGTTAAGGTGATCAACATAGTATGTC	
MAGI3_ex7-8F	GACATACTATGTTGATCACCTTAACCAGAA	1201 bp
MAGI3_ex13-14R	AGCCCCAATATATATAGACTGGTCAGGTCC	
MAGI3_ex13-14F	GGACCTGACCAGTCTATATATATTGGGGCT	1111 bp
MAGI3_ex20-21R	TAATATCCCAATCACCATGGTCAGGTATCA	
MAGI3_ex20-21F	TGATACCTGACCATGGTGATTGGGATATTA	1166 bp
MAGI3_R_full	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATTACTGCCGTTTCTCAGCCATA	
MAGI3_F_full	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGTCGAAGACGCTGAAGAAGA	4510 bp
MAGI3_R_full	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTATTACTGCCGTTTCTCAGCCATA	

For all primer pairs, the annealing temperature is 60°C, and Mg^{2+} concentration is 1mM, using KOD polymerase.

Table 2.14. Additional primers used to sequence the full length of *MAGI3*.

***MAGI3* (MAGI3-004)**	**sequence (5' … 3')**
ex2	GATCAGAGATAATCTCTACTT
ex5	ACTGGGAAATGGCCTACACTG
ex10	GATGACAGTGAAGATCCTGTT
ex12	AGGTCTACCTGAAATCTAAGA
ex16	GAATGGGCAGTCCATTGTTGA
ex20	CAATTGAGCTCATTCAGGCTG
ex21-1	AGAGGTAGATCGGTTAGTCCC
ex21-2	AATGGTTGAGAAATCTCTTCC

2.9.7.3 Preparation of AG1 Competent Cells and Transformation

A single colony of AG1 cells was inoculated into 2ml of LB medium. Cells were grown overnight at 37°C, with shaking; cells were diluted 1:100 dilution in fresh LB medium and grown until OD_{600} reached 0.3-0.6; cells were centrifuged and resuspended in TSB medium (10ml of TSB for 100ml of culture). For the transformation, 0.1-10ng of plasmid DNA was diluted in 20µl of 5x KCM Buffer and sterile water was added up to 100µl; 100µl of competent cells were added, and the mixture was incubated on ice for 20 min. This was followed by incubation at room temperature for 10 min; after which 1ml of LB was added and the mixture was incubated a 37°C for 1 hour, with shaking. Transformed cells were plated on selective LB agar.

2.10 Yeast Two-Hybrid Methods

(All yeast liquid cultures were grown with vigorous shaking, 230-250 rpm)

2.10.1 Yeast Transformation

PJ69-4A competent cells were prepared by inoculating liquid YPAD medium with one yeast colony, and grown (with shaking) at 30°C until OD_{600} reached 0.8-1.0. Cells were split in 50ml tubes and centrifuged at 1000g for 5min; for each tube, cells were washed with 30ml sterile distilled water, centrifuged at 1000g for 5min, and finally resuspended in 6ml 1.1x TE/LiAc solution. The suspension was split in several tubes and centrifuged at maximum speed (13000 rpm) on a tabletop centrifuge; the supernatant was removed and each pellet was resuspended in 600µl of 1.1x TE/LiAc.

Competent cells were transformed as follows: 500ng of plasmid DNA was combined with 5µg of herring testes carrier DNA (BD Biosciences) – previously denaturated for 5 minutes at 95°C - and with 50µl of competent cells, mixing gently. PEG/LiAc solution

(0.5ml) was added to the DNA/competent cell mixture, mixing gently, and the cells were then incubated for 30 minutes at 30°C. DMSO (20µl, Sigma) was added with gentle mixing, and the tube was placed in a water bath at 42°C for 15 minutes, gently vortexing every 5 minutes; cells were pelleted at maximum speed on a tabletop centrifuge; the supernatant was then removed and cells were resuspended in 0.5ml of 0.9% NaCl solution; 200µl of the suspension was spread on the appropriate SD agar medium and cells were incubated at 30°C until colonies appeared.

As the yeast two-hybrid requires a step of co-transformation, the procedure described above was followed, but two vectors instead of one were transformed into the cell: in this case, the total 500ng of plasmid DNA correspond to the combination of two desired vectors.

2.10.2 Bait Auto-activation and Toxicity Assays

To confirm the designed baits did not auto-activate the reporter genes, single yeast colonies of PJ69-4A *MATa* were transformed with each of the four constructs and empty pBD vector. The transformants were then plated on SD-LW and allowed to grow few days until colonies reached 1-1.5mm diameter. The toxicity was assessed by comparing the growth of the bait transformants with the empty vector pBD transformant. To confirm that the baits did not auto-activate the reporter gene, one colony for each transformant was tested for the assays listed in section 2.10.7.

2.10.3 Bait Protein Expression

To confirm that each bait construct expressed the recombinant protein, yeast colonies transformed with each bait construct were cultured and treated to prepare a protein extract to be assayed by western blot.

For each transformant, a single colony (not older than four days) was used to inoculate 5ml of selective SD-W liquid medium, and grown overnight; a positive control was included, consisting of the empty vector pBD transformant. Alongside, a negative control was also prepared by inoculating 10ml YPAD liquid medium with an untrasformed PJ69-4A *MATα* colony. The next day, for each transformant, 50 ml of YPAD was inoculated and the culture was grown at 30°C, with shaking, until the OD_{600}

reached 0.4-0.6. The culture was then centrifuged at 1000g for 5 minutes at 4°C in pre-chilled 50 ml tubes. The supernatant was discarded and the pellet resuspended in 25ml of ice-cold dH_2O, and then centrifuged at 1000g for 5 minutes at 4°C. The pellet was frozen by placing the tube in liquid nitrogen and stored at -80°C.

For the preparation of protein extracts, complete cracking buffer was prepared by adding to 1ml of cracking buffer stock solution: 10µl of β-mercaptoethanol, 50µl of 100mM PMSF (4.5mM final concentration), and a cocktail of protease and phosphatase inhibitors (see Table 2.1 of this chapter), to a final volume of 1.1ml. The buffer was prewarmed to 60°C and 100µl of buffer per 7.5 OD_{600} units of cells was added to each cell pellet (total number of OD_{600} units = $OD_{600\ [1\ ml\ sample]}$ multiplied by the culture volume). The suspension was transferred to a new tube, containing 80µl of glass beads (425-600µm, Sigma) per 7.5 OD_{600} units and heated to 70°C, for 10 minutes; the sample was vortexed vigorously for 1 minute and then centrifuged at maximum speed (13000 rpm) on a tabletop centrifuge for 5 minutes, at 4°C. The supernatant was transferred to a new tube and placed on ice. The remaining pellet was boiled in water for 3-5 minutes and vortexed vigorously for 1 minute, and then centrifuged at maximum speed on a tabletop centrifuge for 5 minutes, at 4°C. This second supernatant was combined with the corresponding first supernatant. Each sample was then boiled briefly and loaded on a gel (7.5% and 15% acrylamide), or stored at -80°C until ready to electrophorese it on a polyacrylamide gel.

2.10.4 Library Screening by Yeast Mating

This method is based on the mating of haploid yeast cell of the opposite mating type, to generate diploid cells. PJ69-4A *MATa* was transformed with the bait constructs and PJ69-4A *MATα* carried the libraries. The mating protocol of Letteboer and Roepman, (2008) was followed and briefly described below.

Three milliliter of SD –W medium was inoculated with one colony of PJ69-4A *MATa*, previously transformed with one bait construct, and grown overnight at 30°C with shaking. Hundred milliliter of SD –W (1:50 / 1:100 / 1:250 dilutions) was inoculated with the overnight culture and incubated overnight at 30°C. The OD_{600} was determined,

and the culture with OD_{600} = ~1 was used (=2.5 x 10^7 cells/ml). 4 x 10^8 library cells were thawed and incubated in 20ml of liquid YPAD, for 10 minutes, at 30°C. 8 x 10^8 bait cells were added to the library cells and mixed. The cell mixture was centrifuged for 5 minutes at 900g at room temperature; the pellet was resuspended in 2ml of YPAD and plated onto four 15-cm YPAD agar plates, incubating for 4.5 hours, at 30°C. Cells were scraped in 10 ml 0.9% NaCl using a cell scraper; plates were rinsed with 10 ml 0.9% NaCl; all the fractions were pooled and centrifuged for 5 minutes at 900g. Finally the pellet was resuspended in 10 ml 0.9% NaCl (and dilutions for mating efficiency were prepared).

For the screening, ~2.5 x 10^5 diploid clones were plated onto ten 15-cm SD–LWHA plates, ~1 x 10^6 diploid clones onto ten 15-cm SD–LWHA plates. The plates were incubated for 3-14 days, at 30°C and subsequently streaked onto fresh SD–LWHA plates and clones were assayed for interactions.

For the mating efficiency (percentage of diploids), different dilutions (10^1-10^6) were prepared in 0.9% NaCl, making serial dilutions: adding 100 µl of culture to 900 µl of 0.9% NaCl (10^1 dilution), adding 100 µl of 10^1 dilution to 900 µl of 0.9% NaCl (10^2 dilution), etc. Each dilution was plated onto a 9-cm SD –LW plate and incubated 2-4 days, at 30°C. The colonies were counted and the mating efficiency calculated as follows: number of colonies counted × dilution factor (the 10^1-10^6 dilutions) = number of cfu/100µl; × 10 = number of cfu/ml; × 10 (last resuspension volume in the mating procedure) = total number of cfu; / 4 x 10^8 (library cells used) × 100% = % of diploids, that is the mating efficiency. The expected mating efficiency is around 25% (total of 100 × 10^6 diploid cells).

2.10.5 Yeast Plasmid Preparation

A single fresh yeast colony was inoculated into 0.5ml of selective SD medium and grown overnight, at 30°C. Cells were centrifuged for 5 minutes at full speed (13000 rpm) on a tabletop centrifuge and then resuspended in 100µl of Buffer S (Table 2.1) and incubated at 37°C for 30 minutes. The mixture was then added to 25µl of Lysis Buffer

(Table 2.1) and incubated at 65°C for 30 minutes; 50µl of 3M KAc was added to each sample, incubated on ice for 10 minutes and afterwards, centrifuged at maximum speed on a tabletop centrifuge at 4°C; 150µl of the suspension was combined with 150µl of ethanol and incubated overnight at -20°C (or 1 hour at -80°C) and the mixture was centrifuged at maximum speed on a tabletop centrifuge at 4°C. The supernatant was discarded and the pellet was washed with 70% ethanol, centrifuged at maximum speed on a tabletop centrifuge at 4°C, the supernatant was then removed and the pellet was air-dried and finally dissolved in 50µl of deionised water.

2.10.6 Long-term Storage of Strains

Liquid cultures of *S.cerevisiae* PJ69-4A strain and the range of transformants were prepared and grown overnight at 30°C. Cells were collected and resuspended in the appropriate medium (YPAD or selective SD) with 20% glycerol (filter sterilised), and finally placed at -80°C for long term storage.

2.10.7 Yeast Two-Hybrid Assays for Identification of Positive Interactors

Colonies obtained from the mating step and growing on SD-LWHA were re-streaked onto fresh SD-LWHA plates, one with X-α-Gal and two without, to assess the following reporter genes

1. Selective Medium Assay for *HIS3* and *ADE2* reporter genes: yeast colonies were plated on SD-LWHA and allowed to grow at 30°C for 3-4 days. The colonies which were growing healthily were considered positive for this assay, as they were able to synthesise hystidine and adenine to support their growth.
2. α-Galactosidase Assay for *MEL1* reporter gene: yeast colonies were plated onto SD-LWHA supplemented with 20 mg/l X-α-Gal (20mg/ml in DMF; Glycosynth), followed by incubation at 30°C for 3-4 days. The colonies that developed a blue-green colour (directly on the solid medium) were considered positive for this assay, as they were able to synthesise α-galactosidase;
3. β-Galactosidase Assay (or filter lift assay) for *LacZ* reporter gene: yeast colonies were plated on SD-LWHA, and, after growing at 30°C for 3 days, they were covered with Whatman 3M paper and grown for a further 2 days at 30°C. The filter was then

dipped into liquid nitrogen for 1 minute, thawed to room temperature (colonies facing up) and finally placed onto filter paper soaked with Z-Buffer (containing 0.27% β-mercaptoethanol and 0.334mg/ml X-β-gal). The filter was placed on a petri dish and sealed with parafilm and incubated at 37°C, for up to 4 hours – the time it takes for a blue colour to appear; produced by the yeast colonies due to β-galactosidase activity.

After these three assays performed on co-transformed clones, the positive clones were sequenced with pAD-GAL4-Amp-DEST primer (see Table 2.11), and with the following primers (Table 2.15), designed according to the clone sequence:

Table 2.15. Primers used to sequence the library clones positive for the yeast two-hybrid assays.

library clones	sequence (5' … 3')
3.9 I	
1F	AGAGGCATTCTGAGCCGTCTG
2F	ACTCGAGATCCATCCCAGCCTTA
4F	CTAGTGACTATCCCTCTGATC
3.10 I	
1F	AGAGGATGGGCAAGCAGTGCA
2F	TTACCAGTGTGAGCATGCAGG
4F	GATCACATATGTTACAGGTGT
4.17 I	
1F	TCGATCCAAGGCTTGTCACAT
2F	GGTATGAACTACTGACTCATGC
4F	GATGCTGATCATCTAGTCTAC

2.10.8 Selection of clones

After a yeast two-hybrid screening, a number of putative interactors will be isolated. Some will be false positives, and they will be eliminated with the co-transformation experiment and with the assay described in the previous section 2.10.7. After this initial selection, the clones will be analysed according to the goodness of the sequence: the vector sequence GAA TTC GGC ACG AGN will be looked for, with the reading frame defined by the GAATTC sequence, corresponding to the *Eco*RI restriction site used to clone the library. If the clones match this requisite, they will be included in the putative interactants list. As this list may be composed of several candidates, a number of criteria will be applied to prioritise certain clones among others:

1) investigation in database and literature about the expression in eye of the putative candidate;
2) investigation in database and literature about domains of the protein which can relate to splicing activity;
3) nuclear localisation;
4) that there is expression of a protein sequence (no early stop codons) from the prey vector and not just a small peptide.
As many clones may be positive for some criteria and negative for others, this is meant to be more a guideline, and hence these criteria will be applied with flexibility, according to the type of putative interacting partner.

2.11 SDS-PAGE and Western Blot

SDS-Polyacrylamide mini-gels (7.5-15%) were cast using a Biorad kit, following the recipe stated in the Table 2.1. Samples were prepared previously with the appropriate amount of 2x Reducing Sample Buffer and denatured at 70°C for 5 minutes. The samples were then loaded on the gel along with a protein marker (Precision Plus Protein Standard, Biorad) and electrophoresed in a Biorad tank (150-200 V) until the desired separation was achieved. The proteins were transferred onto a nitrocellulose membrane (0.45µm, Biorad) using a Biorad tank (100 V, 350 mA max, 1 hour). To ensure that the proteins were transferred onto the membrane, the latter was stained for 2 minutes with Amido Black Staining solution and bands corresponding to the proteins were visualised after 2 washes of 10 minutes each in destaining solution. The membrane was then briefly rinsed with TBS or TBS-T and blocked with Blocking Buffer I for 30 minutes at room temperature. If the membrane was incubated with an antibody diluted in Blocking Buffer II, then the membrane was washed briefly with TBS-T to eliminate any trace of milk prior to incubation with the primary antibody. After the incubation with the primary antibody (overnight at 4°C, for dilutions see Table 2.16), the membrane was washed in TBS-T 3x10 minutes, and incubated 45 minutes at room temperature with HRP-conjugated secondary antibody, diluted 1:10000 in TBS-T. The membrane was briefly rinsed with TBS-T and washed twice for 15 minutes each in TBS-T, and ultimately washed in TBS only, twice for 10 minutes each. For the development, ECL

solution (GE Healthcare) was put on top of the membrane for 1 minute and the membrane was exposed to detect any signal; the film (Fuji Super RX) was developed in Konica SRX-101A medical film processor.

Table 2.16. Summary of the antibody concentrations and blocking buffers used for western blot.

Primary Antibody	Blocking buffer	Dilution
Anti-GAL4 DNA-AD	II	1:2000
Anti-GAL4 DNA-BD	I	1:2000
Anti-6xHis (Sigma)	II	1:2000
Anti-V5	II	1:2000
Anti-$PRPF31_{484-497}$	II	1:2000
Anti-$PRPF31_{140-154}$	II	1:1000
Anti-PRPF6	II	1:1000
Anti-MAGI3	II	1:500
Anti-TRAK2	II	1:1000

2.12 Cell Culture

2.12.1 Maintenance of Cell Lines

HEK293, MDCK and SK-N-SH cells were grown either in 6-well, 24-well, or 85mm plates according to the experimental requirements. Cells used for immunocytochemistry (MDCK, SK-N-SH), were seeded on sterile glass coverslips, placed in each well of a 24-well plate. In each case, cells were fed with D-MEM medium (4.5 g/l D-glucose) supplemented with 100 units of penicillin and 100µg of streptomycin and 10% fetal calf serum; the cells were grown in a humidified incubator at 37°C in 5% CO_2. For transfection, cells were cultured in the same medium but lacking the antibiotics. Cells were passaged when a high confluence was reached, and hence were seeded at a lower cell concentration: cells were washed briefly with sterile PBS to eliminate any trace of medium, and a solution of Trypsin-EDTA (0.025%-0.01%) was added, ensuring the entire surface was covered by it; cells were incubated 5-10 minutes at 37°C to allow the detachament from the surface of the well/plate, and afterwards they were resuspended in medium and centrifuged for 3 minutes at 700 rpm; the supernatant was discarded and the cells were resuspended in the desired amount of medium and dispensed into new wells/plates.

2.12.2 Cell Transfection

Mammalian cell transfection was performed using Lipofectamine 2000 reagent (Invitrogen). The protocol involves the formation of complexes between the DNA and lipofectamine reagent. Prior to this step, the DNA and lipofectamine were both incubated with OptiMEM medium separately, the latter for 5 minutes at room temperature. Afterwards, the lipofectamine-OptiMEM complex was mixed with the DNA-OptiMEM complex and the suspension was incubated for 20 minutes at room temperature before adding it to the cells.

One day before transfection, cells were plated in medium without antibiotics (for final volume, see Table 2.17); once the lipofectamine-DNA suspension was ready to use, it was added to the cells, rocking the plate back and forth to allow even dispersion. Cells were incubated at 37°C, and after 4h the medium was replaced with fresh medium. Cells were incubated for a total of 24 hours prior to testing for transgene expression.

HEK293 cells used for co-immunoprecipitation experiments were transfected on 85mm plates (~75% confluency), whereas MDCK cells used for immunocytochemistry were transfected on 24-well plates (~50-75% confluency); the quantities of reagents used per well are summarised in Table 2.17.

Table 2.17. DNA/Lipofectamine ratio for the cell lines used for transfection.

Cell line	DNA	Lipofectamine	OptiMEM	plating medium
HEK293	24 µg	60µl	2 x 1.5ml	15ml
MDCK	0.6µg	2µl	2 x 50µl	0.5ml

2.12.3 Preparation of Cell Extracts

Wild type and transfected HEK293 cells, wild type and transfected MDCK cells and wild type SK-N-SH cells were grown in 85mm plates until ~75% confluent; the plates were washed twice with cold 1x PBS, and then scraped in 1ml of cold PBS-0.5% TX100, supplemented with a cocktail of protease and phosphatase inhibitors (see Table 2.1). The suspension was collected in 1.5ml tubes and kept on ice throughout the procedures to minimise protein degradation. The extracts were syringed at least 10 times (or more, until the suspension did not clump the needle), to homogenise the sample using a 25G needle. The extracts were either placed in -80°C for long term storage or

used in the downstream application (immunoprecipitation), or treated with reducing sample buffer (Table 2.1) for SDS-PAGE.

2.12.4 Methanol Fixation of Cells

After 24 hours post-transfection (if implied by the experiment), cells to be used for immunocytochemistry were fixed with methanol. Cells were initially incubated in ice-cold methanol for 5 minutes, at -20°C; cells were then rehydrated with PBS for 5 minutes, at room temperature, followed by incubation in blocking solution for at least 15 minutes before proceeding to staining. At this stage, cells were either stained or stored at 4°C in blocking solution.

2.12.5 CDK Inhibitors Treatment

MDCK cells were cultured on sterile glass coverslips, in 24 well plates, until 50-60% confluent (growing in islands). Medium was then replaced with fresh medium supplemented with either one of CDK inhibitors, Olomoucine (CDK2 inhibitor, diluted 1:1000; Calbiochem) and CDK4 inhibitor (Calbiochem, diluted 1:1000), and the cells left at 37°C, for 24 hours before fixing them with methanol.

2.12.6 Temperature Stress Treatment

MDCK cells were cultured on sterile glass coverslips, in 24 well plates, until ~50% confluent (growing in islands). Cells were washed and medium replaced with normal medium supplemented with 10mM HEPES pH 7.4, and cells were placed at 43°C for two hours, after which the cells were immediately fixed in methanol. HEPES buffer was added, as the incubator used for the higher temperature was not buffered with CO_2.

2.12.7 Osmolarity Stress Treatment

MDCK cells were cultured on sterile glass coverslips, in 24 well plates, until ~50% confluent (growing in islands). The physiological osmolarity of the media used for cell culture is 300 mOsm and with these treatments, we achieved hyper- and hypo-osmolarity of the medium by the addition of salt or water respectively. For hyper-osmolarity stress, 150mM NaCl was added to the medium, for a final osmolarity

of 600 mOsm; hypo-osmotic stress was achieved by adding water to the medium, for a final 100 mOsm (10 ml of medium + 20 ml of water). Cells were cultured with these media accordingly for 2 hours, before fixing them in methanol.

2.13 Co-immunoprecipitation

40µl of a 30% suspension of Protein-G Sepharose beads (Pharmacia) were mixed with 2mg of antibody and 500µl of PBS-0.5% TX100; the mixture was gently shaken for 2 hours at room temperature to allow the antibody to bind to the beads; afterwards, beads were briefly spun at maximum speed (13000 rpm) on a tabletop centrifuge and washed once with PBS-0.5% TX100. HEK293 cell extracts were thawed on ice and supplemented with the protease/phosphatase inhibitors cocktail, and finally incubated with the equivalent of 50µl of Protein-G Sepharose beads for 30 minutes on ice, inverting the tubes every now and then. Extracts were then centrifuged for 10 minutes at maximum speed on a tabletop centrifuge, at 4°C, and the supernatant was distributed to each immunobead reaction previously prepared and each tube was rocked for 4 hours, at 4°C. Beads were finally centrifuged at maximum speed on a tabletop centrifuge at 4°C, for 2 minutes, and washed twice with PBS-0.5% TX100 and once with PBS. Finally, 40µl of sample buffer were added and the beads were heated at 75°C, for 10 minutes. The samples were then ready for SDS-PAGE and immunoblot analysis, otherwise were stored at -20°C until required.

2.14 Immunofluorescence

The immunofluorescence images were processed using the Adobe Photoshop 7.0 software. The colour was removed to facilitate the examination on paper after printing; therefore the panel for a single signal is shown in black and white throughout the thesis. The type of visualised channel is indicated by the colour of the legend on the top-left on each panel (for example: PRPF31 detected with Cy3 and identified by a red signal is written in RED). The colours composing the final picture (red, green and blue) were then restored on one panel, termed “merged”.

2.14.1 Immunocytochemistry

Coverslips of MDCK cells, fixed in methanol and blocked with blocking solution (for ICC, see Table 2.1), were stained as follows: 20µl of primary antibody dilutions (Table 2.18) were prepared in blocking solution and were placed on a piece of parafilm and the coverslip was placed on top of it, cells facing down. The cells were incubated for 1h at room temperature, placing the parafilm on blotting paper soaked in water to prevent evaporation. Afterwards, the coverslips were placed back into a 24-well plate, cells facing up, and were washed three times in blocking solution; ~300µl of secondary antibody dilution together with Hoechst 33258 (0.25ug/ml) (Table 2.17) were added to each coverslip directly on the well, and incubated for 45 minutes at room temperature in the dark. Subsequently, coverslips were washed twice with blocking solution, and once with PBS only. For each coverslip, a drop of ProLong Gold anti-fade reagent was placed on a clean glass slide, and the coverslip (after removal of any excess liquid) was lowered carefully onto the drop, avoiding air bubbles; the samples were left to set overnight at room temperature before examining them on fluorescence (Leica DM-IRB) or confocal (Zeiss LSM510UV; Leica TCS-SP2) microscopes.

Table 2.18. Summary of the concentration of antibody used for immunocytochemistry (for antibody reference, see section 2.5).

Antibody	Dilution in Blocking Solution (ICC)
Anti-PRPF31$_{484-497}$	1:200
Anti-V5	1:500
Anti-PRPF31$_{140-154}$	1:100
Anti-TRAK2	1:50
Anti-MAGI3	1:50
Anti-E-cadherin	1:500
Anti-SC35	1:200
Donkey FITC anti-mouse	1:800
Donkey Cy3 anti-rabbit	1:800
Donkey Cy3 anti-goat	1:800
Hoechst 33258	1:2000 (from 0.5mg/ml stock solution)

2.14.2 Immunohistochemistry

Adult mice (C57Bl/6J strain, six-month old) were maintained in 12-hour light/dark cycle (lights on at 7 am). At 10 am, the animals were euthanised by asphyxiation (carbon dioxide) followed by cervical dislocation; dark adapted-mice were sacrificed at 10 am maintaining them in darkness from 7 am. Eyes were enucleated with 2-3 mm of the optic nerve; the cornea and the lens were removed. The eyes were then incubated in 4% PFA at 4°C overnight. The second day, the eyes were washed three times with PBS, for 15 minutes each, at room temperature with agitation. Subsequently, the eyes were incubated in sucrose-PBS four times for 30 minutes each, increasing the sucrose concentration each time: 5%, 10%, 12.5% and 15% sucrose respectively (at room temperature with agitation). The eyes were then left in 20% sucrose-PBS overnight at 4°C without shaking. On the third day, the eyes are shaken slowly for 30 minutes at room temperature in 13% sucrose-PBS and 1/3 of OCT freezing medium (Thermos Biolabs UK). The final wash of 30 minutes at room temperature was in OCT only. Each eye was then placed in a cryo-tube, filled with OCT and snap-frozen by immersion in a freezing medium: 2-methylbutane was frozen in liquid nitrogen and the cylinder was dipped into the 2-methylbutane to freeze. The eyes were stored at -80°C, until ready for cryosectioning; sections were prepared at 10 µm using a cryotome. The mouse retina sections were collected on glass slides and were stored at -80°C until ready for use. The above procedure was performed by Dr C. Chakarova.

Once a section was at room temperature, the blocking procedure was done rapidly: a hydrophobic pen was used to define a well around the section, and a drop (~30µl) of blocking solution I was added gently onto the section, and incubated for 1 hour at room temperature. Each slide was placed on a case fitted with blotting paper soaked in water to prevent evaporation. After the blocking step, each slide was washed 10 minutes three times in PBS-0.1% TX100; the sections were then incubated with the primary antibody, diluted in blocking solution II (see Table 2.19), for 1 hour at room temperature. Next, the slides were washed for 10 minutes three times in PBS-0.1% TX100 and sections were incubated for 45 minutes at room temperature with FITC/Cy3 conjugated secondary antibodies (diluted in blocking solution II, 1:400), in the dark. The slides were then washed twice in PBS-0.1% TX100 for 10 minutes each, and once in PBS only for

10 minutes, adding Hoechst 33258. Finally, slides were wiped from excess liquid and a drop of ProLong Gold anti-fade reagent was placed on top of each section and a glass coverslip was lowered carefully onto the drop, avoiding air bubbles. The sections were left to dry overnight at 4°C before examining them on an epifluorescent microscope (Leica DM-IRB).

Table 2.19. Summary of the antibody dilutions used for immunohistochemistry (for antibody reference, see section 2.5).

Antibody	Dilution in Blocking Solution II (IHC)
Anti-PRPF31$_{484-497}$	1:200
Anti-GRIF1$_{8-633}$	1:50
Anti-MAGI3	1:10
Anti-PRPF3	1:200
Anti-PRPF8	1:200
Anti-PRPF16	1:200
Anti-Rhodopsin	1:200
Anti-TOPORS	1:50
Anti-PNA	1:200
Donkey FITC anti-mouse	1:800
Donkey Cy3 anti-rabbit	1:800
Cy2 streptavidin	1:800
Hoechst 33258	1:2000 (from 0.5mg/ml stock solution)

3. Identification of Interactants of PRPF31

A powerful approach to study protein interactions *in vivo* is the yeast two-hybrid system. This technique was developed by Fields and Song (1989), and since then has become a popular tool in the examination of protein associations. Since its development, the technique has undergone many improvements hence it is highly standardised. Moreover the yeast two-hybrid technique has the advantage of getting physiologically relevant information by manipulating DNA only: when it comes to the identification of the protein interactor, its corresponding cDNA sequence can be identified through sequencing. Since the aim of this project was to test the hypothesis of a retina-specific partner for PRPF31 (see section 1.6), this method was therefore chosen as it allows the detection of potential interactants of PRPF31. Furthermore, the selection of candidates was restricted to the variety of proteins present in retina by choosing two retina libraries. Hence, the yeast two-hybrid and the retina libraries will allow for the discovery of PRPF31 associations in retina, if retina-specific proteins are isolated.

3.1 The Yeast Two-Hybrid System

The yeast two-hybrid system takes advantage of eukaryotic gene transcription, relying on unique properties of the yeast GAL4 transcription factor (TF). A TF is composed of two domains: a DNA-Binding Domain (BD), which binds to the upstream activation sequence (UAS) of a promoter, and an Activation Domain (AD) which elicits the transcription of the gene, recruiting the transcription apparatus. When the domains are separated, they are unable to activate the transcription, but the function can be restored if they are brought close enough. So if a test protein is attached to one domain, e.g. BD, its binding with an interactant will bring AD closer, resulting in the increase of transcription of particular genes. This is the essence of the technique, as it depends on the production of two hybrid proteins: one is the fusion between BD and a “bait” protein (PRPF31 in this study), the second is the fusion between AD and a “prey” sequence (a cDNA sequence from retina libraries in this study).

The fusion proteins BD-bait and AD-prey carry a nuclear localisation sequence to be transported into the nucleus, and their expression is regulated by the *ADH1* promoter. BD-bait initially binds to the UAS upstream a reporter gene; only when there is a direct interaction between bait and prey, and the AD is close to the BD will it initiate the transcription of the reporter gene (Figure 3.1 D).

Along with the two hybrid proteins, reporter genes are the third component of the system. Reporter genes are used to detect any occurring interaction, as they are under the control of a promoter which has BD binding sites (the UAS). Reporter genes can be either integrated in the yeast genome (as in this case) or carried by a plasmid, and are of two types: nutritional selection markers (coding for a gene involved in the biosynthesis of a particular organic compound), and genes coding for α-galactosidase and β-galactosidase enzymes.

The approach consists of an initial mating of yeast cells, where BD-bait vector is carried by a mating type yeast cell and the AD-prey vector by the opposite mating type. Following mating, the new diploid yeast cells harbour the two vectors together and the interaction is assessed through the transcription of reporter genes. Finally the vectors are isolated and sequenced to determine each interacting cDNA.

The steps characterising the yeast two-hybrid screening for the isolation of PRPF31 interactants described in this chapter are as follows:

1. Cloning and characterisation of a suitable bait;
2. Transformation of yeast mating type A with a bait construct;
3. Mating of strains A and α bringing bait and library respectively;
4. Plating and identification of positives;
5. DNA extraction of putative positive prey constructs from yeast and transformation in bacteria;
6. Co-transformation into yeast cells of prey construct and bait construct;
7. Plating and identification of positive clones (and elimination of false positives);
8. Sequencing of positive clones' constructs;
9. Exchanging of bait sequence into prey vector and human ORF cDNA corresponding to prey sequence into bait vector;

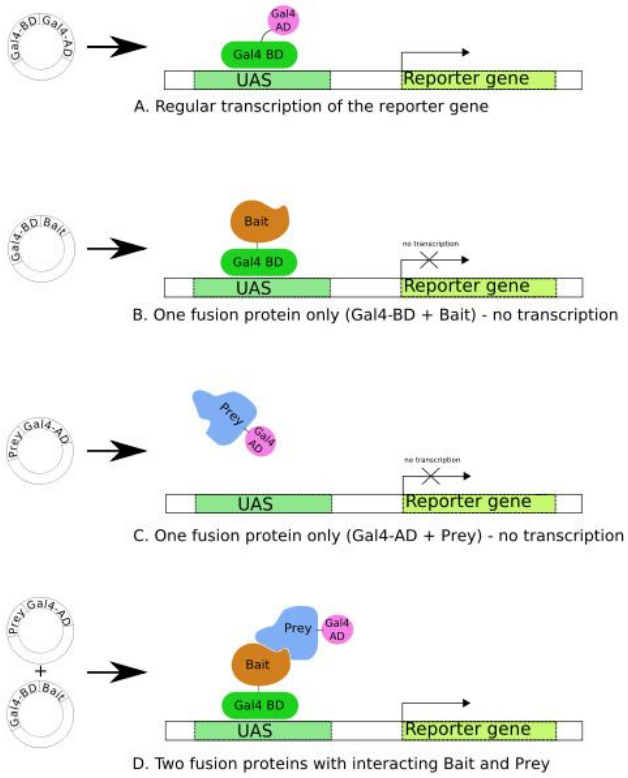


Figure 3.1. Overview of the yeast two-hybrid system.
In A, the transcription factor GAL4 induces the transcription of a reporter gene, as it would normally. In B, the fusion protein formed by the binding domain of GAL4 and the bait is able to bind a DNA sequence, but it is not sufficient to start the transcription of the downstream reporter gene. In C, the fusion protein formed by the activation domain of GAL4 and the prey does not bind the UAS sequence, nor does it start the transcription. In D, the presence of BD-bait bound to the UAS sequence, and the AD-prey interacting with the bait is sufficient to restore the GAL4 function, which in turn sets off the transcription of the reporter gene. Image adapted from http://en.wikipedia.org/wiki/Two-hybrid_screening.

10. Co-transformation into yeast cells;
11. Plating and identification of positives.

Before I joined the project, Dr C. Maubaret had already cloned the bait constructs and was about to perform the screening of the human and bovine libraries, in collaboration with Dr R. Roepman. After I took over the project, I characterised the bait constructs used for the screening (point one) and continued the project from point five onwards, collaborating again with Dr R. Roepman. Due to time constraints, the yeast two-hybrid screening and the characterisation of the bait constructs were carried out in parallel by Dr C. Maubaret and myself, respectively.

As a matter of completeness and clarity, I have included also Dr C. Maubaret's initial results.

3.2 Construction of PRPF31 Baits

The pBD-GAL4-Cam-DEST vector was used to generate BD-PRPF31 recombinant proteins, under the control of the *ADH1* promoter. Each bait protein was a hybrid of the DNA-Binding Domain (BD) of the transcription factor GAL4 (at the N-terminus) and PRPF31. We decided to create not just full length PRPF31 as the bait, but also specific fragments encompassing the single domains, in order to increases the chances of identifying potential interactants.

From the human wild type sequence of *PRPF31*, four different bait constructs were generated (Figure 3.2), following the protocol described in the section 2.9.7.1 of Materials and Methods (Chapter 2):

1. F1: 166 amino acid fragment encompassing the NOSIC domain, corresponding to exon 2-5 and part of exon 6; the predicted molecular weight of the fusion protein is ~37 kDa;
2. F2: 182 amino acid fragment encompassing the NOP domain, corresponding to part of exon 6, exons 7-9 and part of exon 10; the expected molecular weight of the fusion protein is ~39 kDa;
3. F3: 186 amino acid fragment encompassing the NLS domain, and corresponding to exons 10-14; the expected molecular weight of the fusion protein is ~40 kDa;

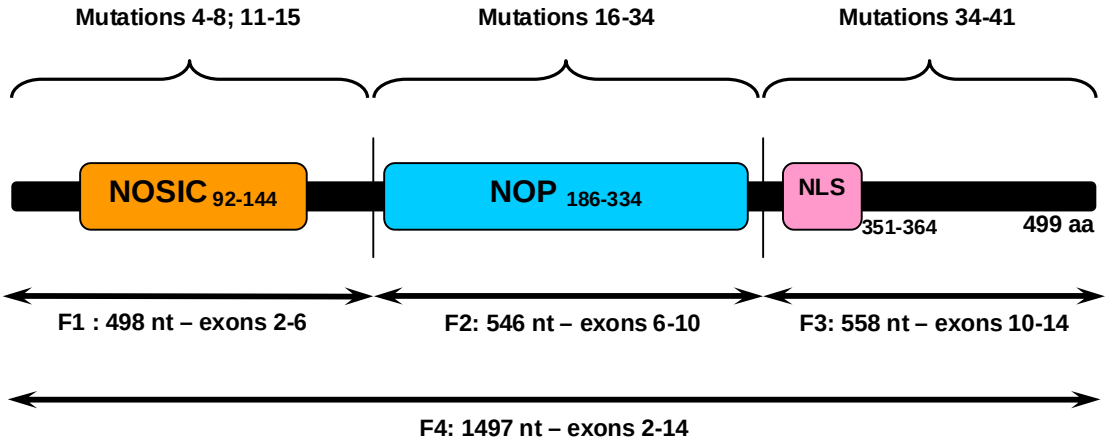


Figure 3.2. Illustration of the various fragments of PRPF31 used as baits for the yeast two-hybrid system.
The coloured boxes highlight the protein domains and their amino acid position along the protein sequence. The arrows illustrate the fragments (named F1-F4), reporting the length in bp and the corresponding exons. The upper brackets group the mutations according to the bait fragment; mutation numbers correspond to those in table 1.2 (section 1.6.1 of the Introduction).

4. F4: 499 amino acid fragment corresponding to the full length sequence of human PRPF31; the predicted molecular weight of the fusion protein is ~76 kDa.

Each clone was restricted and sequence verified (primer pBD-Dest_F, Table 2.11, section 2.9.7.1).

3.3 Characterisation of the Baits

Ideally, a bait to be used in a yeast two-hybrid screening should:

1. Not activate the transcription of the reporter genes (auto-activation test);
2. Not be toxic to the yeast cell;
3. Be expressed in the yeast cell;

The characterisation of the bait constructs consists in verifying these three characteristics.

3.3.1 Bait Auto-activation Assay

To confirm each bait did not activate the transcription of any reporter gene, PJ69-4A *MATa* yeast strain was transformed with each bait construct, and the activation of the reporter genes was assessed through the selective medium assay (yeast two-hybrid assay #1, paragraph 2.10.7). A suitable bait should not activate any reporter gene without the presence of an interacting prey, therefore one should not expect growth of the yeast transformant in this type of assay. As shown in Figure 3.3, clones were not able to grow on SD-WHA medium, as *HIS3* and/or *ADE2* genes were not expressed, therefore not permitting any growth due to the lack of histidine and adenine.

In parallel, an analogous test was carried out: PJ69-4A *MAT*α yeast cells were co-transformed (as described in paragraph 2.10.1) with the bait constructs and with the (empty) pAD-GAL4-Amp-DEST vector; and also with empty vectors pBD-GAL4-Cam-DEST and pAD-GAL4-Amp-DEST vectors. Co-transformed yeast cells were plated on SD-LW, and then plated onto SD-LWHA medium carrying out the three tests, observing no activation of the reporter genes (Figure 3.4, panels A-E; note that the α-galactosidase assay was such that any picture taken was not representative of the result, hence only β-galactosidase is presented herein).

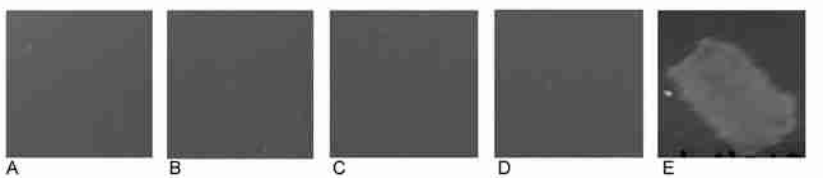


Figure 3.3. Selective medium assay for PRPF31 baits.
Panels A-D represent snapshots of SD-WHA plates: A: BD-F1; B: BD-F2; C: BD-F3; D: BD-F4. None of the baits is able to grow on the selective medium lacking Histidine and Adenine. Panel E is an example of a patch of healthily growing yeast.

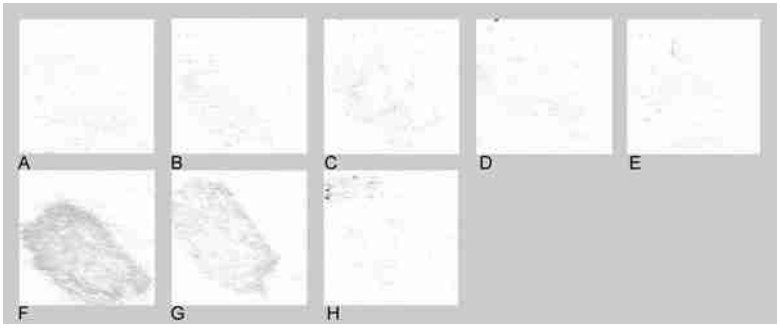


Figure 3.4. β-galactosidase assay.
Assay for bait auto-activation: pAD-GAL4-Amp-DEST + (A) pBD-GAL4-Cam-DEST, (B) BD-F1, (C) BD-F2, (D) BD-F3, (E) BD-F4.
Assay of the control plasmids: (F) pAD-WT and pBD-WT, (G) pAD-MUT + pBD-MUT, (H) pAD-WT + pBD-pLC.

In conclusion, all the four baits were suitable for use in a library screening as no activation of the reporter genes was observed: as shown by both Figures 3.3 and 3.4 (A-E panels), no *HIS3*, *ADE2* or *LacZ* activity was observed.

For comparison, a set of control vectors were separately co-transformed: the pBD-WT control plasmid expresses the BD of GAL4 and amino acids (aa) 132–236 of wild type lambda cI, fragment C. The pAD-WT control plasmid expresses the AD of GAL4 and aa 132–236 of wild type lambda cI, fragment C. (The lambda cI gene product (cI-wt) naturally forms homodimers). The pAD-MUT control plasmid expresses the AD of GAL4 and aa 132–236 of E233K mutant lambda cI, fragment C. The pBD-MUT control plasmid expresses the BD of GAL4 and aa 132–236 of E233K mutant lambda cI, fragment C. The cI-E233K mutation encodes a substitution in the gene product that interferes with the interaction between the homodimers, resulting in a weaker protein–protein interaction (from HybriZAP-2.1 Two-Hybrid Libraries, manual #838401-13).

The combination pAD-WT + pBD-WT gives a bright blue colour on assays #2 and #3, a sign for strong interaction; the combination pAD-MUT + pBD-MUT gives a pale blue colour on assays #2 and #3, for a weak interaction; the combination pAD-WT + pBD-pLC does not show any colour, a sign of no interaction (Figure 3.4, panels F-H respectively). Hence, it is possible to compare the panels A-E for the auto-activation, to panel H (corresponding to pAD-WT + pBD-pLC), indicating no activation of *LacZ* reporter gene.

3.3.2 Bait Toxicity Assay

The toxicity of pBD-F1, pBD-F2, pBD-F3 and pBD-F4 was analysed by transforming PJ69-4A *MATa* yeast strain with the four bait constructs. The bait growth was compared to the empty vector pBD-GAL4-Cam-DEST as described in section 2.10.2 (Figure 3.5). Colonies per quarter section of plate were counted, and estimated for a full 15cm plate (Table 3.1); they were, approximately:

Table 3.1. Number of colonies estimated for a 15cm plate.

	PBD-GAL4-CAM-DEST	PBD-F1	PBD-F2	PBD-F3	PBD-F4
COLONIES	3036	2972	3264	1136	1648

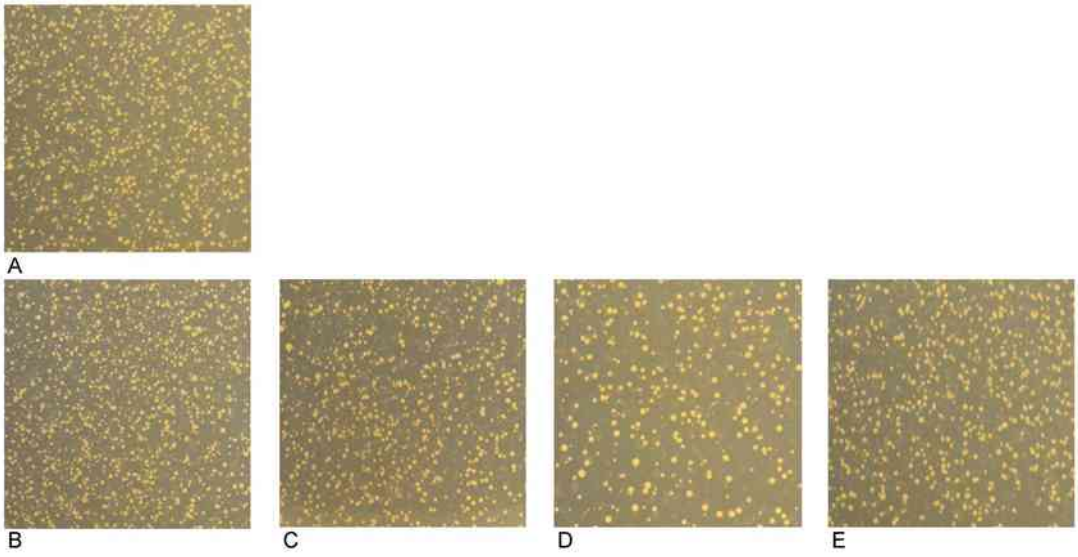


Figure 3.5. Bait toxicity assay.
A-E panels represent snapshots of plates of yeast colonies growing on SD-W medium. A: pBD-GAL4-Cam-DEST. B: BD-F1. C: BD-F2. D: BD-F3. E: BD-F4.

From the number of colonies, it was observed that pBD-GAL4-Cam-DEST and F1 and F2 transformants are similar (panels A, B and C respectively); on the other hand, F3 and F4 show less than half the colonies of the pBD-GAL4-Cam-DEST transformant, which could be due to a slight toxicity of the recombinant protein to the yeast cell. To the eye, the growth and the dimensions of the colonies were comparable for all of them, except for F3, which showed slightly larger colonies. Smaller number of colonies in the F3 and F4 baits could be due to less number of cells plated after the transformation than that of the pBD-GAL4-Cam-DEST, hence explaining the lower number of colonies.
Nevertheless, as all the baits looked generally healthy compared to the control, each was considered suitable to be used in a yeast two-hybrid screening.

3.3.3 Bait Expression

The expression of the recombinant baits was assessed by western blot analysis, as described in section 2.10.3. Figure 3.6 shows the bands corresponding to the fusion proteins, expressed in PJ69-4A *MATa*. The anti-GAL4 DNA-BD antibody (used to detect the bait fusion proteins) revealed two bands corresponding to bait BD-F1 and the BD alone (in panel A, lanes pBD-F1 and pBD-GAL4-Cam-DEST respectively). On closer inspection of the blot, the remaining three baits also presented a weak band each (indicated by a red arrow in panel A). The presence of bait BD-F4 was confirmed also with the anti-PRPF31$_{484\text{-}497}$ antibody (detecting the C-terminal of the prpf31 protein): its corresponding band is indicated by the red arrow, lying below the 75kDa as expected.
Considering the band corresponding to BD-F4 (full PRPF31 protein) in panel B is very strong, and the one in panel A is not, it may be that the anti-GAL4 DNA-BD antibody does not bind as efficiently as anti-PRPF31$_{484\text{-}497}$ antibody. The epitope of the anti-GAL4 antibody might be masked in the F2, F3 and F4 fragments.
Although some baits did not show a high level of expression, they were nonetheless used to screen the libraries, as this characterisation could not take place at the same time of the library screening (due to facilities and knowledge for the yeast two-hybrid method not in-house). Based on the western blot data, BD-F1 was considered a high quality bait, as it was expressed and easily detected on western blot; baits F2, F3 and F4 were considered less suitable/desirable than F1, since their level of expression was lower.

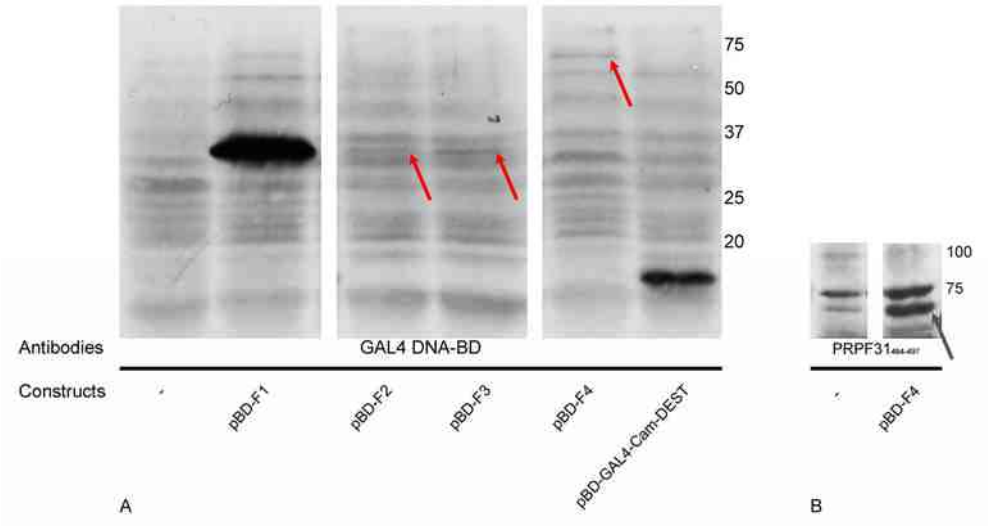


Figure 3.6. Expression analysis of PRPF31 baits.
In A: yeast protein extracts of PJ69-4A strain (lane -), of yeast expressing pBD-F1, pBD-F2, pBD-F3, pBD-F4, pBD-GAL4-Cam-DEST (as labelled on bottom of each lane). Extracts were probed using anti-GAL4 DNA-BD antibody, which revealed bands at ~35-37kDa for BD-F1, and at ~17kDa for BD domain (pBD-GAL4-Cam-DEST lane). The comparison between lanes for pBD-F2, pBD-F3, pBD-F4 with the negative control (lane -) revealed weak bands at a molecular weight corresponding to BD-F2, BD-F3, BD-F4 (red arrows).
In B, yeast protein extracts of PJ69-4A strain (lane -) and of yeast expressing pBD-F4 probed with anti-$PRPF31_{484-497}$ antibody: the antibody revealed a band lower than 75kDa, evidencing for expression of PRPF31.

The low level of expression may be explained either by low abundance of the constructs within the yeast; or the hybrid protein is rapidly degraded although not toxic (see section 3.2.2). However, since a band was detected for each bait, they were considered adequate for the screening.

3.4 Yeast Library Mating and Co-transformation

A bovine retina cDNA library and a human retina cDNA library were screened with the four bait constructs. Both libraries were pre-transformed into the PJ69-4A *MATα* yeast strain. The mating consisted of combining yeast cells of opposite mating types (*a* and *α*), one type carrying one of the bait vectors, the other one carrying the library; this step was carried out as described in section 2.10.4. The mating efficiency was calculated and summarised in the following Table 3.2:

Table 3.2. Summary of mating efficiencies per bait and per library.

	BD-F1	BD-F2	BD-F3	BD-F4
HUMAN LIBRARY	**2.8%**	**2.7%**	**0.5%**	**1.3%**
BOVINE LIBRARY	**4.2%**	**4.1%**	**2.6%**	**2.8%**

It appeared that the mating efficiency for the human library was lower compared to the bovine library; also, in both screenings, the efficiencies of baits F1 and F2 were higher, suggesting that perhaps the yeast cells harbouring these baits were more suitable than F3 and F4, which instead gave a low mating efficiency in both libraries. However, as the initial number of library cells used for the mating was 4×10^8, even a 0.5% mating efficiency allowed to screen more than a million clones ($\sim 2 \times 10^6$ clones precisely).

After the mating procedure, the total number of colonies positive for the three reporter genes was 68 for the bovine library and 48 for the human library (assays described in section 2.10.7). The above steps were carried out by Dr. Maubaret, and from here onwards I took over the project.

As different unrelated multicopy plasmids can be present in the same yeast cell, the prey plasmids were isolated from yeast (as described in section 2.10.5) and used to transform *E.coli* (section 2.9.7.3). Yeast prey vector was isolated from each bacterial colony which was then used for co-transformation with the appropriate bait (described in the last paragraph of section 2.10.1). This was done to determine which of the several plasmids, if more than one per yeast cell, triggered the reporter genes.

Makarova and co-workers (2002) have previously shown interaction between PRPF6 and PRPF31 by yeast two-hybrid, so we cloned the full length sequence of human *PRPF6* into pAD-GAL4-Amp-DEST and included it in the co-transformation experiment as a positive control. Control vectors were co-transformed alongside to check for reporter genes activation. Figure 3.7A shows the expression of the AD-PRPF6 protein: the anti-GAL4-AD antibody reacted with a band between 100 and 150 kDa believed to be the recombinant protein (indicated by a green arrow), as the same band was not visible on the yeast strain lysate nor on the yeast expressing only the AD domain (indicated by a red arrow). Figure 3.7B shows the β-galactosidase assay of PRPF6 with the four baits, to be compared with the controls in Figure 3.4. As can be seen in the figure, the interaction of PRPF6 with fragment F1, F2 and F3 of PRPF31 did not give any blue colour (panels A-C), however full length PRPF31 gave a weak interaction (D). The presence of mutations in the commercially bought clone of PRPF6 was excluded, since the entire cDNA was sequenced, using the primers listed in Table 2.12.

3.5 Sequence Analysis of Positive Clones

After the co-transformation step, the number of putative interactants positive on the three assays was reduced to 11 for the bovine library. All the remaining clones from the bovine library and all the clones isolated from the human library appeared to be false positives: they did not show either growth on selective medium or colour development on the assays, therefore they were not included for further analysis. At this stage, the verification of the 11 bovine clones was done by sequencing (as described in section 2.9.5; primers pAD_BRP F and pAD_BRP R in Table 2.11, section 2.9.7.1). Sequencing

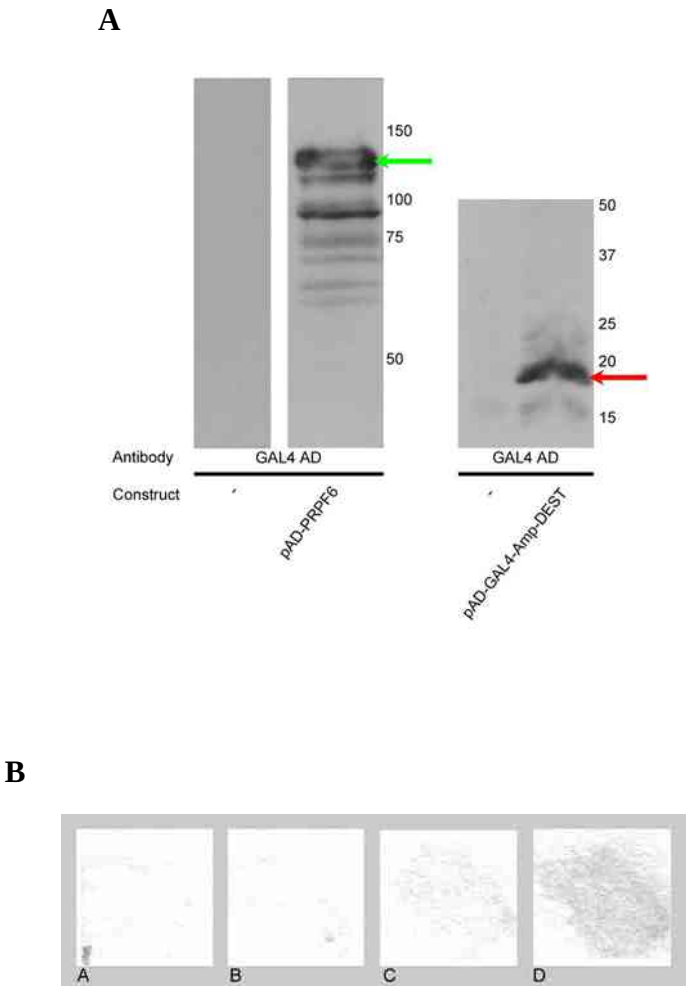


Figure 3.7. Expression of AD-PRPF6 (A) and β-galactosidase assay of AD-PRPF6 (B).
Figure A shows the expression of AD-PRPF6 prey (indicated by green arrow) upon incubation with the anti-GAL4-AD antibody; the figure shows also the blot of yeast protein extracts of PJ69-4A strain (lane -) and of pAD-GAL4-Amp-DEST (AD domain is indicated by red arrow). Figure B shows the β-galactosidase assay between pAD-PRPF6 and pBD-F1, pBD-F2, pBD-F3, pBD-F4 (panels A, B, C, D respectively).

was required to check if the prey cDNA, cloned in the prey vector (pAD-GAL4-Amp-DEST), was in frame with the AD domain, such that the expression of a recombinant protein corresponding to the clone was possible.

The identification of the cDNA sequence was possible by searching the vector sequence for GAA TTC GGC ACG AGN, in which the *Eco*RI restriction site (GAATTC) defined the reading of the frame. Six clones out of 11 from the bovine library appeared to be in frame and thus were then considered for further analyses; instead the remaining five were considered artifactual as not leading to protein expression and thus rejected. The six clones were fully sequenced (primers listed in Table 2.15 of section 2.10.7), of which three showed early stop codons leading to the production of truncated proteins (20-40 aa long). Nonetheless they were also considered as it is known that yeast tolerates translational frameshift, skipping stop codons and continuing the translation (Matchmaker Pretransformed Libraries User Manual PT3183-1, published 5 February 2007). The clones were analysed performing BLASTn searches: one corresponded to the MAGI3 protein, four clones corresponded to TRAK2 (three of which were those leading to a truncated protein), and one to ZNF143. The remaining five clones were not considered for further analysis, as the cloned sequence was not in frame with the AD domain sequence, therefore not leading to the translation of a known protein. Considering that the number of isolated clones was little (six only) and that four out of six corresponded to the same protein species, the criteria for the selection of clones listed in section 2.10.8 was not applied. We instead continued with the investigation of the putative candidates, even though not retina-specific, as they may shed new light into PRPF31 protein network, which may be basis for further speculation of RP in the future.

3.6 Putative Positive Interactants

The yeast two-hybrid method resulted in the identification of three potential interactants of PRPF31 and they will be described in the following sections. The corresponding prey constructs were tested for auto-activation in PJ69-4A *MATα* (following the same procedure used for PRPF31 baits, described in section 3.3.1): each hybrid protein did not activate the transcription of the reporter genes.

3.6.1 Membrane-associated Guanylate Kinase, WW and PDZ Domain-containing Protein 3, MAGI3 (clone 3.9)

A 1646bp cDNA fragment (clone 3.9) encoding a protein sequence similar to human MAGI3, variant Q5TCQ9-4 (Swiss-Prot) was found to interact with bait F3. The translated sequence of clone 3.9 corresponds to residues 127-678 of Q5TCQ9-4 (which is 1481 amino acids long), and this fragment encompasses the guanylate kinase domain, WW domains, and PDZ-1 and PDZ-2 domains. Figure 3.8A shows the β-galactosidase assay, with development of pale blue colour. It is possible that this interaction was weak, however, the positive control with PRPF6 also showed for weak interaction. Figure 3.8B illustrates the alignment between clone 3.9, the bovine protein (ENSBTAP00000026128) and the human protein Q5TCQ9-4 (Clustal W method).

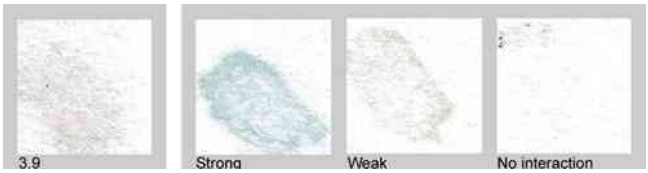
MAGI3 belongs to a family of scaffolding proteins named MAGUKs, for membrane-associated guanylate kinases.

Cell polarisation is essential for selective and asymmetrical transport of substrates: thus, in epithelial cells, apical and basolateral membrane compartments are separated by tight junctions, which form beneath the most apical part of the lateral membrane. Tight junctions provide a physical barrier yet allowing permeability, and contribute to cell-cell adhesion.

Scaffolding proteins are involved in cellular adhesion organising and anchoring protein complexes at cell-cell contact sites, contributing to tissue development and differentiation. They organise macromolecular protein complexes, tether receptors, intracellular signalling enzymes and adhesion molecules.

Previous studies by various groups have identified MAGUKs as ubiquitous scaffolding proteins concentrated at cell-cell contacts. MAGUKs are characterised by a guanylate kinase (GK) domain of ~300 amino acids, catalytically inactive. Other features include an SH3 domain, usually upstream of the GK domain, or a WW domain following the GK domain, and also PDZ domains. MAGIs are MAGUK proteins with an inverted arrangement of protein-protein interaction domains (GK domain is N-terminal, and the PDZ domains are C-terminal). Three MAGIs have been identified to date, (MAGI1, MAGI2, MAGI3), and all contain an N-terminal GK domain followed by two WW

A



B

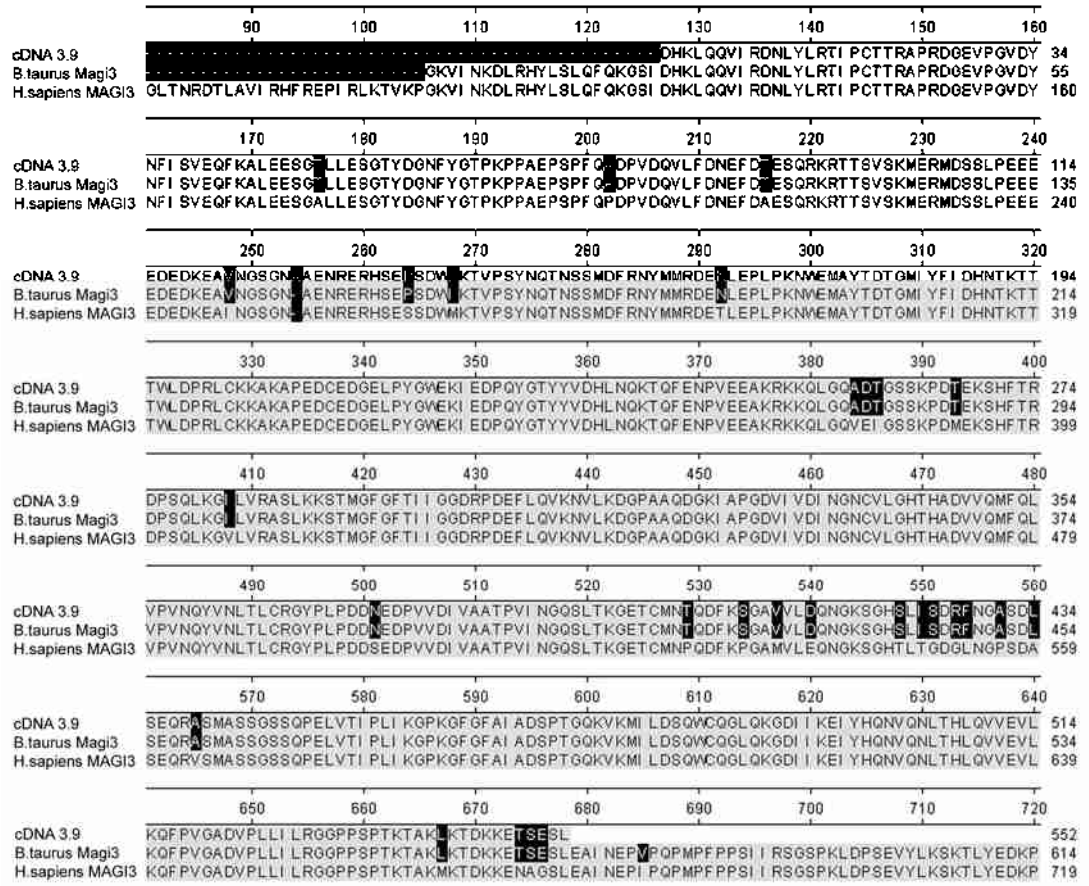


Figure 3.8. Clone 3.9 interacts with BD-F3.
A: β-galactosidase assay after co-transformation of PJ69-4A *MATα* with clone 3.9 and bait BD-F3; the three panels on the right represent the controls.
B: alignment of the amino acid sequences of clone 3.9 cDNA (detected as BD-F3 interactant), the predicted bovine and the human MAGI3 sequences. In light grey residues that match Q5TCQ9-4 exactly, in solid black residues that differ from Q5TCQ9-4.

domains and six PDZ domains (Figure 3.9). WW domains are functionally similar to SH3 domains (they both bind to proline-rich regions), perhaps explaining the substitution (as reviewed by Funke *et al.*, 2005).

MAGI3 was first identified (Wu *et al.*, 2000) through a yeast two-hybrid screening in search for interacting partners of PTEN/MAC, a phosphatase which quenches PKB activity, the latter implicated in the inhibition of apoptosis; MAGI3 interacts with PTEN/MAC through the PDZ-2 domain. Wu and colleagues documented that MAGI3 is expressed in various foetal and adult tissues, such as brain, liver and lung.

Later it was discovered that another phosphatase, RPTPβ (receptor protein tyrosine phosphatase β) interacts with rat MAGI3, through its PDZ-3 or PDZ-4 (Adamsky *et al.*, 2003): RPTPβ is implicated in cell-cell and cell-matrix interactions, and in cell adhesion and migration; however, MAGI3 was not found to be tyrosine-phosphorylated.

Adamsky and colleagues performed western blot analysis on rat cell and tissue samples, revealing a major band at 140 kDa, and two minor bands at 170 kDa and 130 kDa, suggesting that MAGI3 has multiple isoforms. In epithelial cells, MAGI3 co-localises with ZO-1 (Wu *et al.*, 2000) and cingulin to tight junctions, and it partially overlaps with β-catenin and desmoplakin (proteins found in adherens junctions and desmosoms respectively; Adamsky *et al.*, 2003). Adamsky and colleagues also showed that in transfected MDCK cells, MAGI3 was found at cell-cell contacts, with a strong punctate nuclear staining, and a weakly labelled cytoplasm. In rat primary astrocytes, MAGI3 was localised to the nucleus and at adherens junctions and cellular protrusions with β-catenin. The authors concluded that MAGI3 is found in the cell nucleus (MAGI3 contains two putative nuclear localisation signals) and in specific structures along the plasma membrane, and, according to cell type and culture conditions, it could be present in tight and adherens junctions, cellular protrusions and focal adhesion sites (Adamsky *et al.*, 2003).

Other interacting partners of MAGI3 include: the oncogenic HPV E6 protein (with the PDZ-1 domain of MAGI3), which targets MAGI3 (and MAGIs 1 and 2 as well) to degradation (Thomas *et al.*, 2002); the terminus of the ATEV motif (with PDZ-4) found in Delta proteins (interacting also with MAGI1 and MAGI2; Wright *et al.*, 2004); the cytoplasmic region of proTGFα (transforming growth factor-α precursor; with PDZ-1;

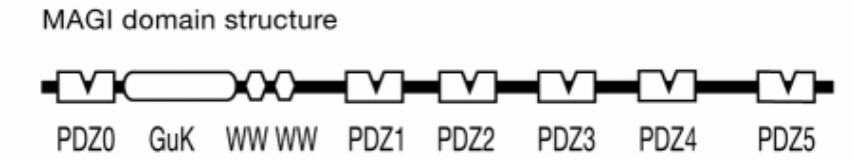


Figure 3.9. Domain structure of the MAGI family of proteins. (Image taken from Wright *et al.*, 2004).

Franklin *et al.*, 2005); the C-terminal region of β_1-Adrenergic Receptor (with PDZ-1), a G protein-coupled receptor (β_1AR; He *et al.*, 2006); and the G protein-coupled receptor LPA_2 (which mediates signaling by lysophospahtidic acids; with PDZ-5), and when MAGI3 was over-expressed, a clear increase in RhoA activation was observed (Zhang *et al.*, 2007).

3.6.2 Zinc Finger Protein 143, ZNF143 (clone 3.10)

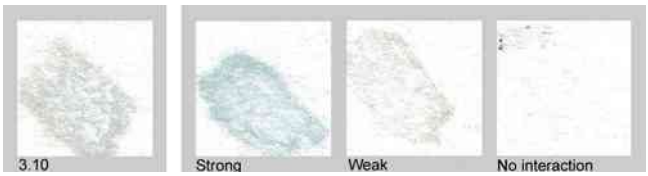
A 1960bp cDNA fragment (clone 3.10) encoding a protein sequence similar to human ZNF143, P52747 (Swiss-Prot), was detected with bait F3. The translated sequence of clone 3.10 matches the complete sequence of P52747 (626 amino acids). According to Pfam, the protein contains seven central Zn finger (zf_C2H2) domains.

Figure 3.10A shows the β-galactosidase assay: also in this case, as the colour is a pale blue, the interaction between BD-F3 and clone 3.10 was a weak one, in accordance with the α-galactosidase assay, which was comparable to the control for weak interaction. Figure 3.10B shows the alignment between clone 3.10, the bovine protein (ENSBTAP0000001394) and the human protein P52747 (Clustal W method).

ZNF143 encodes a zinc finger protein that is expressed in all normal adult tissues. ZNF143, or hStaf, is the human ortholog of *Xenopus laevis* Staf (sharing 84% of identity) and it is a seven-zinc finger transcription factor of 638 amino acids, and the zinc finger domains are repeated in tandem. Staf was originally identified in *X.laevis* as a transcription factor, involved in the transcriptional activation of snRNA and snRNA-type genes transcribed by RNA polymerase II and III. It was reported to target several human snRNA genes (U2, U4, U6, U11 and others) for transcription. As Myslinski and co-workers (1998) identified the human ZNF143 to be functionally equivalent to *Xenopus* Staf, it appears that the transcriptional activation of these genes relies on ZNF143.

Recently, it was reported a region on the human U6 snRNA promoter is able to recruit the ZNF143 (Yuan *et al.*, 2007): the authors showed that ZNF143 activates transcription from a U6 gene template assembled into chromatin and identified CHD8 as a ZNF143-associated protein required for efficient U6 transcription *in vivo*.

A



B

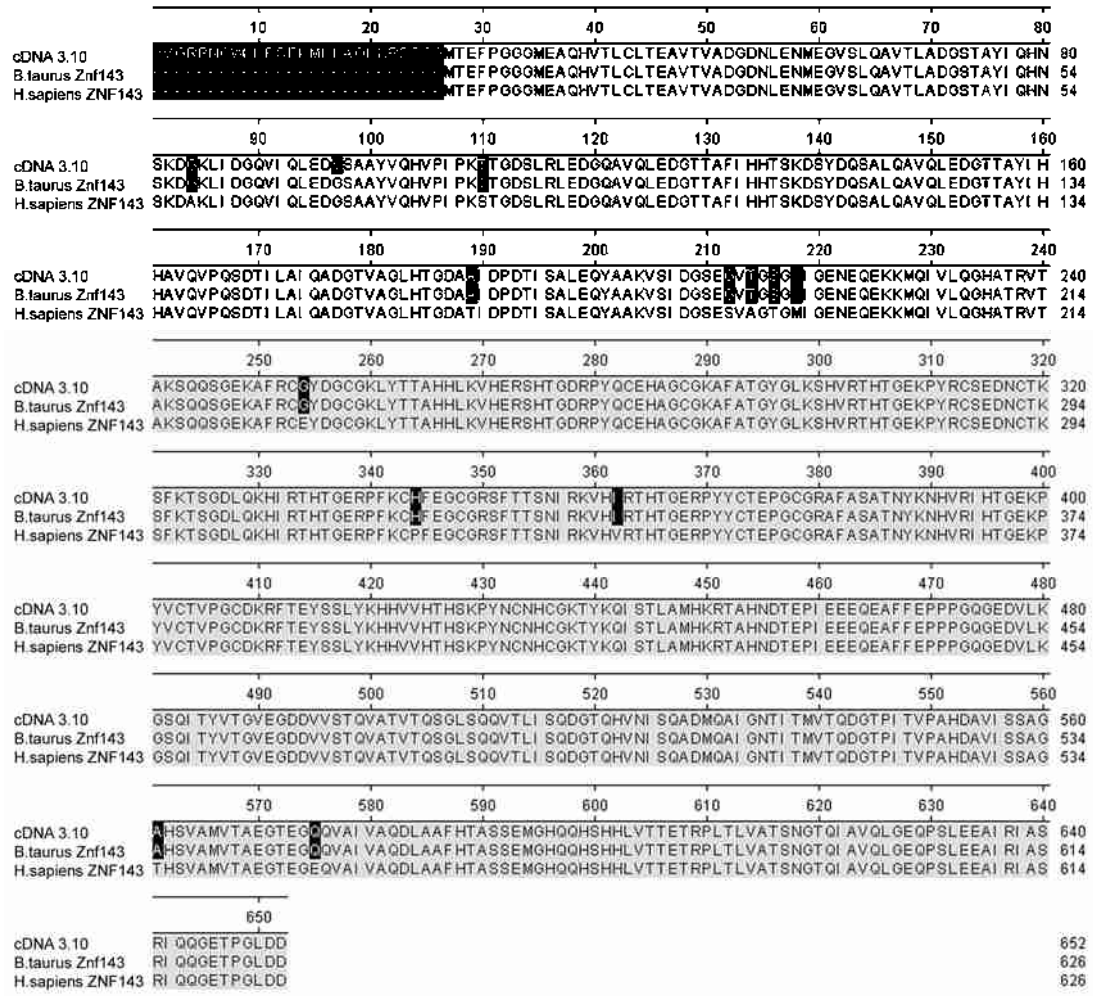


Figure 3.10. Clone 3.10 interacts with BD-F3.
A: β-galactosidase assay after co-transformation of PJ69-4A *MATα* with clone 3.10 and bait BD-F3; the three panels on the right represent the controls.
B: alignment of the amino acid sequences of 3.10 cDNA detected as BD-F3 interactant, the predicted bovine and the human ZNF143 sequences. In solid light grey residues that match P52747 exactly, in solid black residues that differ from P52747.

Also, ZNF143 regulates the transcription of genes involved in DNA repair (through association with tumor suppressor p73; Wakasugi *et al.*, 2007), cell survival (*TAL*; Grossman *et al.*, 2004), cell cycle (*BUB1B*; Myslinski *et al.*, 2007), and mitochondrial transcription (*TFAM*; Gerard *et al.*, 2007).

3.6.3 Trafficking Kinesin-binding Protein 2, TRAK2 (clone 4.17)

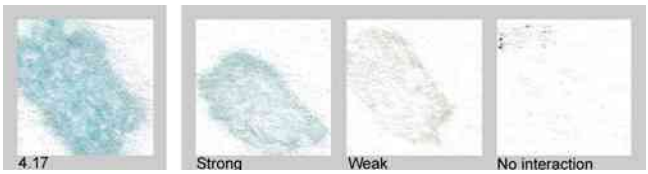
A 580bp cDNA fragment (clone 4.17) encoding a protein sequence similar to human TRAK2, O60296 (Swiss-Prot), interacted with bait F4. The translated sequence of the clone corresponds to residues 107-299 of O60296 (914 amino acids). According to Pfam, this sequence shows a HAP1_N domain.

Figure 3.11A shows the β-galactosidase assay: considering the vivid blue colour developed on the filter-lift assay, the interaction between clone 4.17 and BD-F4 was strong – confirmed by the α-galactosidase assay, comparable to the control for strong interaction (not shown). Figure 3.11B shows the alignment between clone 4.17, the bovine protein (ENSBTAP00000019878) and the human protein O60296 (Clustal W method).

The rat orthologue GRIF1 was firstly isolated from rat brain through a yeast two-hybrid screening, and was among the partners interacting with the intracellular loop of the $GABA_A$ β2 subunit. The screen took place in order to identify the proteins which influence the receptor localisation, which shows different localisations according to its subunit composition. GRIF1 is highly expressed in excitable tissues, mostly brain and heart. The protein contains a coiled-coil region (HAPN) homologous to HAP1A, HAP1B; GRIF1 also contains a weak leucine zipper motif. The expression of GRIF1 in HEK293 cells revealed a punctate, plasma membrane localisation, along with cytoplasmic localisation (Beck *et al.*, 2002).

The human GRIF1, also called TRAK2 or OIP98, is a 98 kDa protein, and was found both an interactor and a substrate of OGT, which is responsible for the *O*-GlcNAc (acetylglucosamine) post-translational modification (OGT transfers the saccharide to the proteins; Iyer *et al.*, 2003): in fact TRAK2 is modified by GlcNAc. Brickley and co-workers (2005) proposed a role for TRAK2 in membrane trafficking of vesicles and/or organelles, via motor proteins, demonstrating that TRAK2 interacts with kinesin.

A



B

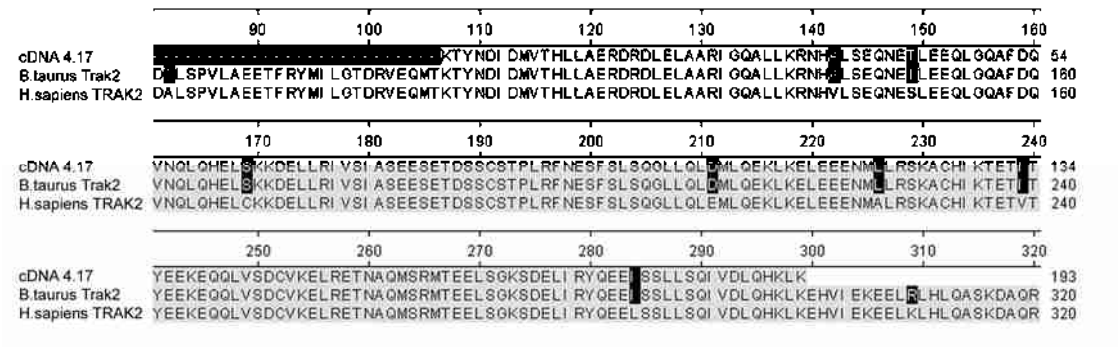


Figure 3.11. Clone 4.17 interacts with BD-F4.
A: β-galactosidase assay after co-transformation of PJ69-4A *MATα* with clone 4.17 and BD-F4; the three panels on the right represent the controls.
B: alignment of the amino acid sequences of 4.17 cDNA detected as BD-F4 interactant, the predicted bovine and the human TRAK2 sequences. In solid light grey residues that match O60296 exactly, in solid black residues that differ from O60296.

Further experiments reinforce the hypothesis of TRAK2 being involved in trafficking: TRAK2 interacts with K^+ channel Kir2.1, facilitating its trafficking to the cells surface (Grishin *et al.*, 2006). GRIF1 also interacts with Hrs, and may regulate endosomal trafficking since Hrs is a component of the endosome-to-lysosome trafficking (Kirk *et al.*, 2006).

Given these experimental data, TRAK2 was proposed as an adaptor protein involved in motor-dependent trafficking of proteins to synapses (Brickley *et al.*, 2005).

As TRAK2 displays a HAPN domain, it is of interest describing the HAP1 protein. HAP1 (huntingtin-associated protein 1) is expressed primarily in neurons, it is a cytoplasmic protein with no transmembrane domains or nuclear localisation signals; it contains several coiled-coil domains, which may regulate or mediate its interaction with a number of proteins. Scientific evidence supports the hypothesis of HAP1 being involved in intracellular trafficking and as a scaffold protein in the assembly of protein complexes: in fact, it interacts with a subunit ($p150^{Glued}$) of dynactin (a component of the dynein motor complex), and also interacts with kinesin light chain 2 (KLC2 – a subunit of kinesin motor), suggesting that HAP1 is involved in both the retrograde and anterograde transport respectively. Moreover, HAP1 interacts with Hrs, involved in endosome-to-lysosome trafficking, and it also interacts with $GABA_A$ receptor, modulating the inhibitory synaptic transmission. Still, it is not clear how HAP1 participates in those diverse processes: it may act as a scaffold to stabilise protein complexes for trafficking, or it could also be an adaptor to link microtubules with different cargoes (as reviewed by Rong *et al.*, 2007).

The *Drosophila* orthologue of GRIF1 protein is Milton, which is responsible for the transport of mitochondria in the drosophila retina (Stowers *et al.*, 2002). There is also evidence that *Drosophila* photoreceptors lacking Milton, fail to target mitochondria to their terminals, remaining abundant and present throughout the soma (Gorska-Andrzejak *et al.*, 2003).

TRAK2, together with its homologous TRAK1 (or OIP106), define a new family of kinesin-associated coiled-coil proteins, proposed to function as adaptors in the anterograde trafficking of organelles to synapses. It is also interesting to note that

TRAK1 localises to distinct punctate regions in the nucleus, and co-localises with RNA polymerase II *in vivo* (Iyer *et al.*, 2003).

Clones 4.6, 4.12, 4.31 also correspond to TRAK2, but showed early stop codons after the *Eco*RI site, producing 20-40 amino acid long peptides. It has been reported that yeast is capable of skipping stop codons (Matchmaker Pretransformed Libraries User Manual PT3183-1, published 5 February 2007), and continuing the translation process. These clones were not characterised further but were taken as a further confirmation of TRAK2 as an interactant of PRPF31.

3.7 Expression Profiles on Interactants

The *MAGI3*, *TRAK2*, and *ZNF143* genes are ubiquitously expressed, according to their UniGene expression profiles. However, the proteins were identified as PRPF31 interactants from a bovine library. In order to make sure they are all present in human retina, the three transcripts were amplified by RT-PCR from human retina cDNA (total RNA from Clontech). Along with retina, other human tissues were tested: brain, heart, liver, lung, pancreas, placenta and skeletal muscle. Oligonucleotides (Table A.1 in appendix) were designed to amplify a region spanning two adjacent exons (*MAGI3* exons 17-18, *TRAK2* exons 11-12, *ZNF143* exons 4-5), using Biotaq polymerase. Results are shown in Figure 3.12, indicating that all three genes are expressed in the human retina, since bands of the expected sizes were amplified (264 bp for *MAGI3*, 247 bp for *TRAK2*, 157 bp for *ZNF143*). Genomic DNA was also included in the reaction, and its amplification showed a band for each gene at a higher molecular weight due to inclusion of one intron separating the amplified exons (expected sizes for genomic DNA: 844 bp, 544 bp, 682 bp, for *MAGI3*, *TRAK2*, *ZNF143* respectively).

The RT-PCR amplification also showed that each gene is expressed in other tissues, and with the abundance in retina comparable to the rest of the tested tissues. I would point out that this result is only indicative, as no quantification was performed.

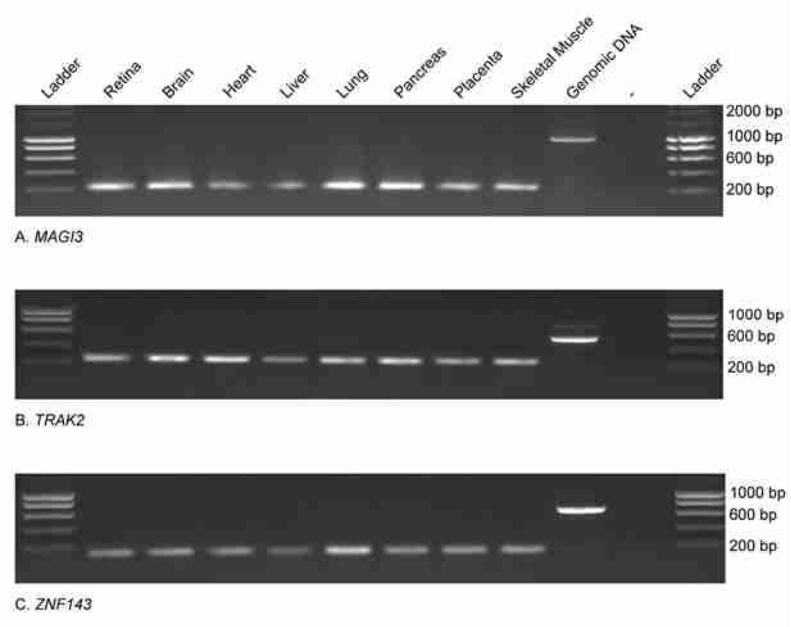


Figure 3.12. Expression profiles of interacting partners.
cDNA types are specified on top of panel A and refer to panels B and C. "-" lanes for negative control. For each gene, the cDNA lanes show a band of the size of the predicted transcript fragment.

3.8 Co-transformation of Human Wild Type Sequences

To further validate the positive interactions, the human wild type cDNA corresponding to each of the isolated clones was used in a co-transformation step. Full length *MAGI3*, *TRAK2*, and *ZNF143* cDNAs were cloned into pBD-GAL4-Cam-DEST as described in section 2.9.7.1, to generate BD-fused proteins. For the generation of the wild type *MAGI3* sequence (corresponding to isoform MAGI3-004, Ensembl database), refer to section 2.9.7.2 and the following section 3.8.1. Full length PRPF31 was cloned into pAD-GAL4-Amp-DEST, using the gateway cloning system.

The bait constructs were transformed in PJ69-4A *MATα* to verify they did not activate the reporter genes: the selective medium assay is shown in Figure 3.13A, and as no colonies were observed, the baits were considered suitable for the further experiments. Figure 3.13 B shows the expression of baits and prey in yeast: pBD-MAGI3 and pBD-TRAK2 (green and yellow arrows respectively) were not as expressed as pBD-ZNF143 (blue arrow); incubation with the anti-GAL4-AD antibody showed a band at ~75 kDa compatible with AD-PRPF31 (red arrow). As the expression was confirmed by the western blot, we decided to proceed with the experiments.

Each new bait construct was co-transformed in PJ69-4A *MATα* with the pAD vector expressing PRPF31. Selective medium assay and α-galactosidase assay were positive for each combination (not shown). However, as shown in Figure 3.14, the β-galactosidase assay confirmed the interactions previously observed between PRPF31 and MAGI3, and PRPF31 and TRAK2 (panels A and B respectively, showing a vivid blue colouration), whereas PRPF31 and ZNF143 interaction was weak.

As the experiment described above partially supported the outcome of the original screening (confirming the interaction with PRPF31 of two out of three partners), it was decided to continue the characterisation of those showing a strong interaction (MAGI3 and TRAK2). However, the fact that in this experiment ZNF143 failed to interact strongly with PRPF31 does not exclude it as a partner: in fact, it is possible that the DNA-BD tag affects ZNF143, for example by changing its folding, thus preventing a strong contact with PRPF31.

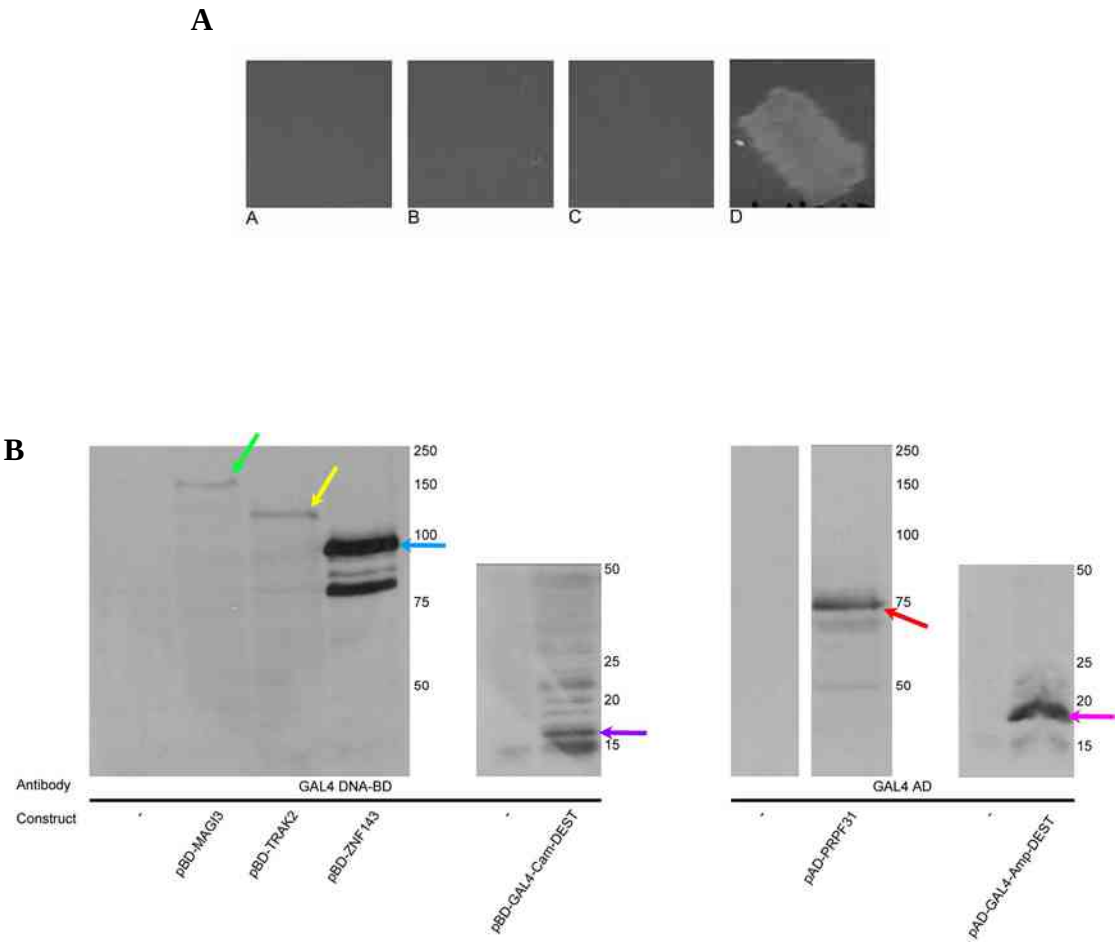


Figure 3.13. Selective medium assay for human baits (A) and expression of new baits and prey (B). In A, selective medium assay for pBD-MAGI3 (panel A), pBD-TRAK2 (panel B), pBD-ZNF143 (panel C): no growth was observed after streaking a colony of yeast bearing the respective construct. Panel D illustrates a healthily growing patch of yeast.
In B, expression of BD-MAGI3, BD-TRAK2 and BD-ZNF143 (green, yellow and blue arrows respectively; left-hand sided blots) and AD-PRPF31 (red arrow, right-hand side blot). BD domain is indicated by the purple arrow; AD domain is indicated by a pink arrow.

Figure 3.14. Co-transformation β-galactosidase assay.
pAD-F4 + (A) pBD-MAGI3, (B) pBD-TRAK2, (C) pBD-ZNF143. The comparison between these panels and the three panels on the right for control plasmids suggests a strong interaction for both MAGI3 and TRAK2 with PRPF31, whereas ZNF143 represents a weaker interaction.

3.8.1 Generation of Human Wild Type *MAGI3*, MAGI3-004 Isoform

As described in section 2.9.7.2 of the cloning procedures, SPLICE approach was used to generate the full coding sequence of *MAGI3*. According to the alignment of the bovine sequences of MAGI3 (prey and database sequences), it was decided to amplify MAGI3-004 (transcript ENST00000307546 in Ensembl; Q5TCQ9-4 protein in Swiss-Prot), from human retinal cDNA.

Figure 3.15A and 3.15B show the bands corresponding to exons 1, 2-7 and exons 8-13, 14-20, 21 respectively, after the first round of PCR; Figure 3.15C shows the bands corresponding to exons 1-13 and 14-21, after the second round of PCR, and finally, Figure 3.15D shows the band corresponding to the full length sequence. The last was verified by sequencing (primers listed in Tables 2.13 and 2.14, Chapter 2). The final sequence was then used to generate BD-MAGI3 and V5-MAGI3 constructs as described in section 2.9.7.1.

Oligonucleotides (Table 2.12) were designed for the *MAGI3* transcript sequence (~4.5 kb) which generates protein Q5TCQ9-4; each oligonucleotide had 15 nucleotides overhangs at the 5' and 3' ends, pairing the upstream and downstream exons, and designed to amplify portions of cDNA of ~1 kb each. KOD Hot Start polymerase was used for the purposes of this experiment.

Exons 1, 2-7, 8-13, 14-20 and 21 were amplified using the following thermocycling settings: 40 cycles, 22 sec of elongation (Figure 3.15 A and B). The amplicons were gel-purified, mixed in an equal molar ratio and two greater fragments were amplified (corresponding to exons 1-13 and exons 14-21; Figure 3.15 C), using the following conditions: 8 initial cycles, 50 seconds of elongation, without addition of oligonucleotides (the products have base-pairing sequences); secondly, 32 cycles, 50 seconds of elongation with the addition of the flanking primers. As the PCR resulted in several amplification products, the fragments for exons 1-13 and 14-21 were identified in the gel according to their size (expected to be 2353 bp and 2277 respectively); they were purified and mixed in equal molar ratio, and then underwent a third round of PCR to amplify the whole 4.5 kb transcript, using conditions as stated before (8 + 32 cycles), but with longer elongation time (92 seconds).

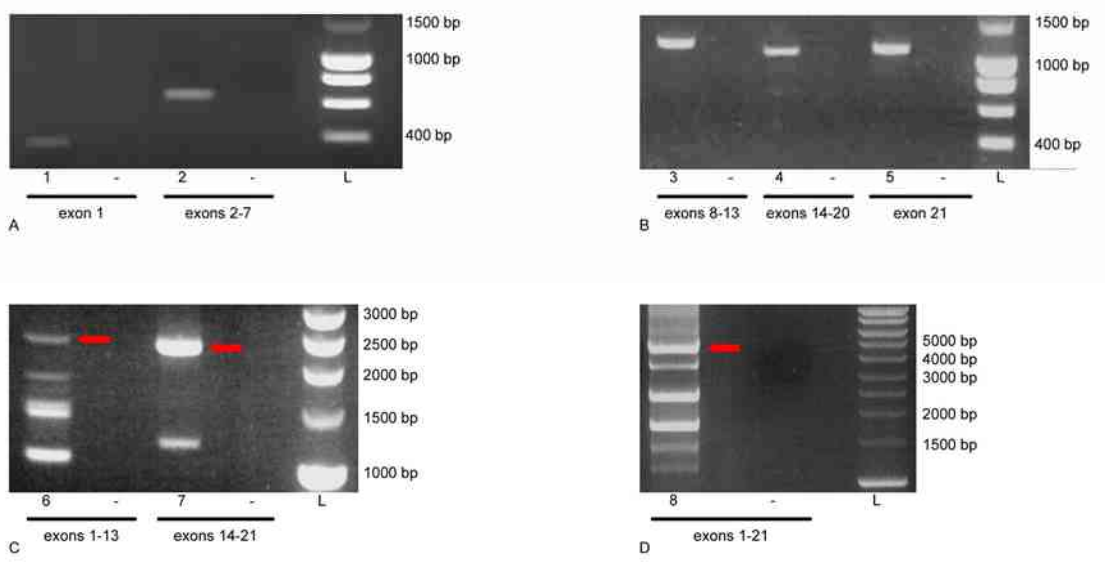


Figure 3.15. Generation of human wild type *MAGI3* using the SPLICE method.
In A, amplification of exon 1 (lane 1), exons 2-7 (lane 2). In B, amplification of exons 8-13 (lane 3), exons 14-20 (lane 4), exon 21 (lane 5). In C, amplification of exons 1-13 (lane 6, red arrow), exons 14-21 (lane 7, red arrow). In D, amplification of exons 1-21 (lane 8, red arrow). L is for ladder, - is for negative control.

The amplicon (Figure 3.15 D) corresponding to the expected size was gel-extracted and purified, recombined in pDONR/Zeo and fully sequenced to ensure the accuracy (oligonucleotides for sequencing listed in Table 2.13, Chapter 2). The full length sequence was then recombined in pBD-GAL4-Cam-DEST vector.

3.9 Amplification of *PRPF6* from Bovine and Human Libraries

As PRPF6 is known to interact with PRPF31 (see section 1.6 of the Introduction), it was expected to be among the positive interacting partners from the library screens. As described in the previous paragraph 3.6, PRPF6 does not appear among the positive interactants. Hence, we decided to amplify it from the retina libraries used for the yeast two-hybrid. Plasmid DNA aliquots of the libraries were kindly donated to us from Dr R. Roepman. The concentration of the human library aliquot was 150ng/µl, whereas the bovine was 4.5ng/µl. One microliter of the human library was sufficient to transform AG1 cells for amplification, whereas 10µl of the bovine library was used for the transformation in the same *E.coli* strain. Transformed cells were plated onto several plates and the following day all the colonies were collected into 5 ml, of which 2.5 ml was used to inoculate a 250 ml culture, for subsequent DNA extraction. The amplified library was then used for a PCR, using three different pairs of PRPF6 primers (to amplify two adjacent exons, listed in Table A.2 in Appendix), using an equal amount of template DNA from each library, and applying identical thermocycling conditions. Figure 3.16 shows the amplification of PRPF6 in both libraries: panel A shows the results for the human library where lanes 1-3 show the expected fragments (sizes 415 bp, 582 bp, 622 bp respectively); panel B represents the bovine library, and only lanes 1 and 3 show a weak band each, of the expected size (415 bp and 622 respectively, red arrows).

Several reasons may explain the different levels of amplification of *PRPF6* in the two libraries: firstly, the process of amplification is likely to have affected the presence of the various transcripts constituting the libraries; secondly, the human library might be more abundant in *PRPF6* transcript than the bovine one. The reason why PRPF6 was not picked up among the partners in both screenings may be: 1) the human library screening failed for some reasons, not allowing PRPF6 to be detected together with other

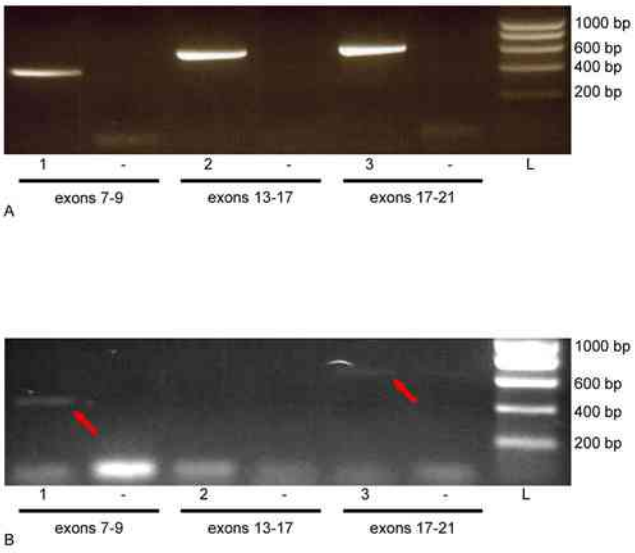


Figure 3.16. PRPF6 amplification on retina libraries.
In A, the human library: lanes 1, 2 and 3 show amplicons at the expected molecular weight.
In B, the bovine library: only lanes 1 and 3 showed the expected bands, whereas no fragment was detectable on lane 2. L is for ladder, - is for negative control.

interactants, and 2) regarding the bovine library screening which was successful, the interaction with PRPF31 could be weak and hence not detectable by the system used. On the other hand, it is to note that the primers used to amplify were designed on a human sequence, hence the paucity of product in the bovine library may be due to the oligonucleotides failing to anneal to the template.

3.10 Conclusion

From the analysis of the results presented in this chapter, only the bovine retina library screen was successful in identifying interacting partners for PRPF31. Despite the use of four different baits, the screen of the human library was not as successful and did not yield any positive interactants. Three novel partners of PRPF31 were isolated from the bovine library, corresponding to the following human proteins: TRAK2 (cytoplasmic protein involved in trafficking), MAGI3 (both nuclear and cytoplasmic/junctional protein), and ZNF143 (a zinc finger, nuclear protein); none of which being retina-specific. As the yeast two-hybrid system allows the detection of direct interactions, these partners are to be considered potential direct interactants. Results from co-transformation experiments using human MAGI3, TRAK2 and ZNF143 sequences alongside PRPF31 (section 3.8), indicated the former were both strong interactors, whereas ZNF143 appeared to be quite weak. Therefore we selected MAGI3 and TRAK2 for further experiments to confirm these findings. However, notwithstanding the co-transformation experiments (sections 3.4 and 3.8), performed to exclude any false positives, still care must be taken when considering MAGI3 and TRAK2 as good candidates. In fact, both proteins MAGI3 and TRAK2 present domains which are protein-protein interaction platforms, hence it may be that they are “sticky” proteins that can be isolated in another, unrelated screen with a different bait other than PRPF31. The results need a cautious interpretation also because the clones were isolated either only once (MAGI3), or with apparent artifactual clones (TRAK2, clones 4.6, 4.12 and 4.31, see section 3.6.3). Hence, the results presented in this chapter need further validations to test their genuinity as PRPF31 interactors, especially because, if they are genuine, they may be very interesting and suggestive.

4. Confirmation of Novel PRPF31 Interactants, MAGI3 and TRAK2

Having identified new interacting partners of PRPF31 through the yeast two-hybrid system, a validation of these interactions with a different method is necessary. Though the isolated partners came from a retinal library but none appeared to be retina-specific, we still pursued the validation experiments as they may enrich the knowledge regarding network of interactions of PRPF31. We chose to co-immunoprecipitate the proteins to verify the interactions illustrated in Chapter 3, using transfected mammalian cells expressing the proteins of interest. We followed up MAGI3 and TRAK2 as they appeared to be the strongest interacting partners of PRPF31 in the direct yeast two-hybrid experiments, as demonstrated in section 3.8 of Chapter 3.

Immunoprecipitation (IP) is a technique which allows the precipitation of an antigen (protein) using a specific antibody directed against the antigen. Thereby, the specific protein can be isolated from a mixture of many other proteins, simply by incubating the protein solution with the specific antibody. However, the antibody must be coupled to a solid substrate, commonly beads coated with immunoglobulin-binding protein, which binds the Fc region of an immunoglobulin (*i.e*. the tail region of an antibody).

The term co-immunoprecipitation (co-IP) describes the application of the IP procedure to precipitate protein complexes. In this case, the antibody of choice will target the desired protein, and the entire complex will be extracted from the solution, making it possible to identify the various components of the complex (Figure 4.1). For successful pull-down, the known protein must have the epitope exposed, and not masked by any other member of the complex. Thus any partner binding on the epitope will not be identified.

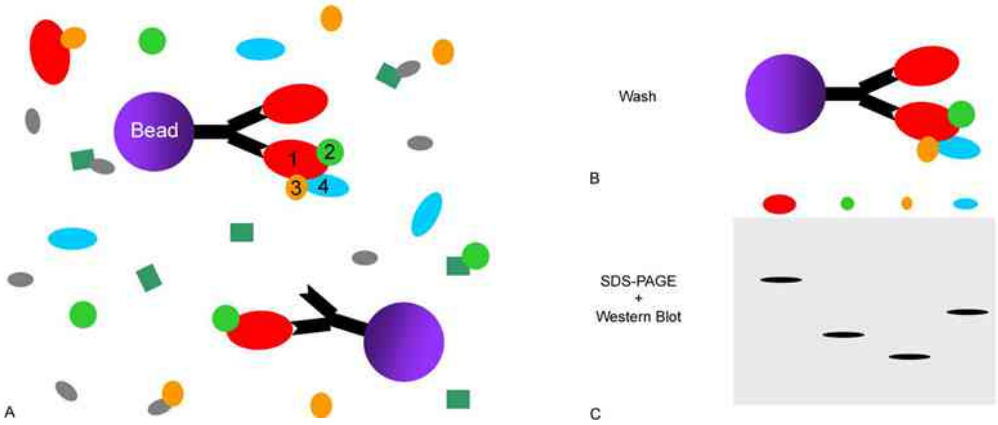


Figure 4.1. Illustration of a traditional co-immunoprecipitation.
Panel A is a depiction of a protein mixture, combined together with beads coupled to antibodies, in a tube. Panel B represents the situation after a washing step: the final product bead+antibody+complex is cleared from the unbound proteins, and ready for SDS-PAGE and western blot as pictured in C.

For this study, the co-IP approach involved the over-expression of the desired proteins in the HEK293 cell line. We chose to over-express tagged proteins as antibodies directed against the tags are commercially available. These will also facilitate the pull down of the complex and its detection on western blot. However this method presents a drawback in that it raises the likelihood of formation of inclusion bodies, in which case the over-expressed protein aggregates and is less accessible. Also, the over-expressed protein may interfere with the cellular pathway in which the equivalent endogenous protein is involved.

For the purpose, the proteins were expressed as N-terminally 6xHis tagged and V5 tagged proteins. This procedure was also used to detect association between wild type PRPF31 and V5-MAGI3, and V5-TRAK2. In addition, the reverse approach was also performed.

4.1 Generation of N-terminally Tagged Proteins

The pDEST26 vector was used to generate construct 6xHis-PRPF31; pcDNA3.1V5DEST vector was used to generate constructs V5-MAGI3, V5-PRPF6, V5-TRAK2. All hybrid proteins were expressed in HEK293 cells, under the control of the CMV promoter. Full-length open reading frames, listed in section 2.7, were used for this purpose. All the clones were checked first by restriction digestion followed by direct DNA sequencing (primers listed in Table 2.11).

4.2 Expression of Tagged Proteins in HEK293 Cell Line

HEK293 cells were cultured in 85 mm plates, and when ~75% confluent, they were transfected with Lipofectamine 2000 reagent (as described in section 2.12.2) with the following combinations of DNA constructs (ratio 1:1):

1. pDEST26-PRPF31 + pcDNA3.1V5-MAGI3;
2. pDEST26-PRPF31 + pcDNA3.1V5-PRPF6;
3. pDEST26-PRPF31 + pcDNA3.1V5-TRAK2;

After 24 hours, cells were scraped and lysed into PBS-0.5% TX100 as described in section 2.12.3. In order to see whether these proteins are expressed in the transfected

cells, each cell extract was treated with sample buffer and loaded on an acrylamide gel for immunoblot analysis. Figure 4.2 shows the bands corresponding to the various fusion proteins: 6xHis-PRPF31 is shown on blot A, as a band migrating at ~65kDa as expected; untransfected cells were also blotted with anti-6xHis antibody (lane "-" in blot A), and the antibody revealed a weak band at the same molecular weight of 6xHis-PRPF31. V5-tagged proteins are shown in blots B and C: V5-MAGI3 migrates at a ~150 kDa (lane "V5-MAGI3"), and V5-TRAK2 is detected at ~120 kDa; V5-PRPF6 is shown in blot C, and as expected, migrates at ~100 kDa; untransfected cells (lane "-", blot B) were also probed, revealing no significant background. As MAGI3 is 1481 amino acids long, its corresponding band was expected to be at ~165 kDa; however, the actual band lies at a lower molecular weight. This may be caused by incomplete denaturation of the protein; formation of secondary structures; presence of detergents. However, it is also important to consider that protein size in Daltons is estimated as an average for all the amino acids whereas each actually contributes a different amount according to the type of amino acid. Therefore proteins of similar length in amino acids may migrate differently depending on the combination amino acids they contain, and as such cannot be predicted accurately.

The backgrounds from the two antibodies were different: the α-V5 antibody gave a clean blot even after long exposure time, whereas the background from α-6xHis antibody (Sigma) showed extra bands, among which a non-specific band at ~65kDa.

4.3 Co-immunoprecipitation of 6xHis-PRPF31 and Interactants, Using anti-6xHis Antibody (Abcam)

The method for validating the yeast two-hybrid data was the co-immunoprecipitation of PRPF31 with the said proteins. Having shown that the fusion proteins were successfully expressed (according to section 4.2), the first step was to precipitate PRPF31 and the splicing factor PRPF6. The same procedure was then applied to validate MAGI3 and TRAK2 with PRPF31.

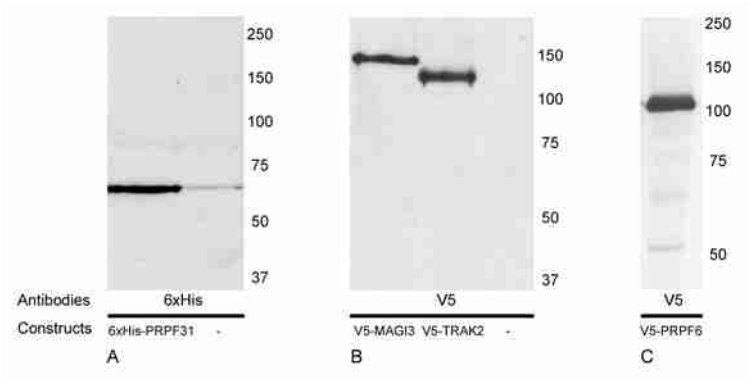


Figure 4.2. Expression analysis of tagged proteins in HEK293 cells.
A, expression of 6xHis-PRPF31: anti-6xHis antibody (Sigma) revealed a band at ~65 kDa, corresponding to the tagged protein; however, a faint band was observed in wt cells (lane -). B and C: expression of V5-MAGI3, V5-TRAK2, V5-PRPF6: anti-V5 antibody revealed bands at the corresponding sizes of the fusion proteins (~150 kDa, ~120 kDa, ~110 kDa respectively); the incubation of the wt cell lysate (lane -) did not reveal any significant band. The positions of molecular mass standards (kDa) are shown on the right.

4.3.1 Co-immunoprecipitation of 6xHis-PRPF31 and V5-PRPF6

For this assay, HEK293 cells were co-transfected with the following combinations of DNA constructs (ratio 1:1):

1. pDEST26-PRPF31 + pcDNA3.1V5-PRPF6;
2. pDEST26-PRPF31 + pcDNA3.1V5DEST; (control: PRPF31 and empty V5 vector);
3. pDEST26 + pcDNA3.1V5-PRPF6; (control: empty 6xHis vector and PRPF6);
4. pDEST26 + pcDNA3.1V5DEST; (control: empty 6xHis and V5 vectors).

Protein-G beads were coupled with anti-6xHis (Abcam) and crude cell lysates were incubated with the immunobeads, according to the protocol described in section 2.13. The samples were then electrophoresed on SDS-polyacrylamide gel, and probed with the antibodies directed against the tags.

Figure 4.3 (panel A) shows the co-IP of PRPF31 and PRPF6 expressed in HEK293 cells, pulled down with the anti-6xHis antibody and probed with the anti-6xHis and the anti-V5 antibodies. The complex 6xHis-PRPF31 + V5-PRPF6 shows strong positive interaction. In fact, panel A confirms the interaction between the two protein species, when construct combination #1 was used for immunoprecipitation: the 6xHis blot shows a band for 6xHis-PRPF31 in both IP (immune pellet) and input (transfected cell lysate) lanes (6xHis-PRPF31 is marked by a green asterisk, and represents the lower band at ~65 kDa); the V5 blot shows V5-PRPF6 in both IP and input lanes (immunoreactive band observed at ~100 kDa, marked by a red asterisk).

Panels B, C and D show all the possible controls run alongside to exclude the possibility of nonspecific interactions (construct combinations #2, #3, and #4 respectively). Panel B shows anti-6xHis antibody immunoreactive bands at ~65 kDa (6xHis-PRPF31) in both the IP and input lanes. Panel C demonstrates that V5-PRPF6 does not precipitate in the absence of 6xHis-PRPF31, as an anti-V5 antibody immunoreactive band is detected at ~110 kDa in the input lane (indicated by the red asterisk) but not in the IP lane. A comparison between the 6xHis blots of panel B and panels C and D, illustrates that 6xHis-PRPF31 is represented by the lower band (panel B, green asterisks), as only the upper band is present in C and D (yellow asterisks).

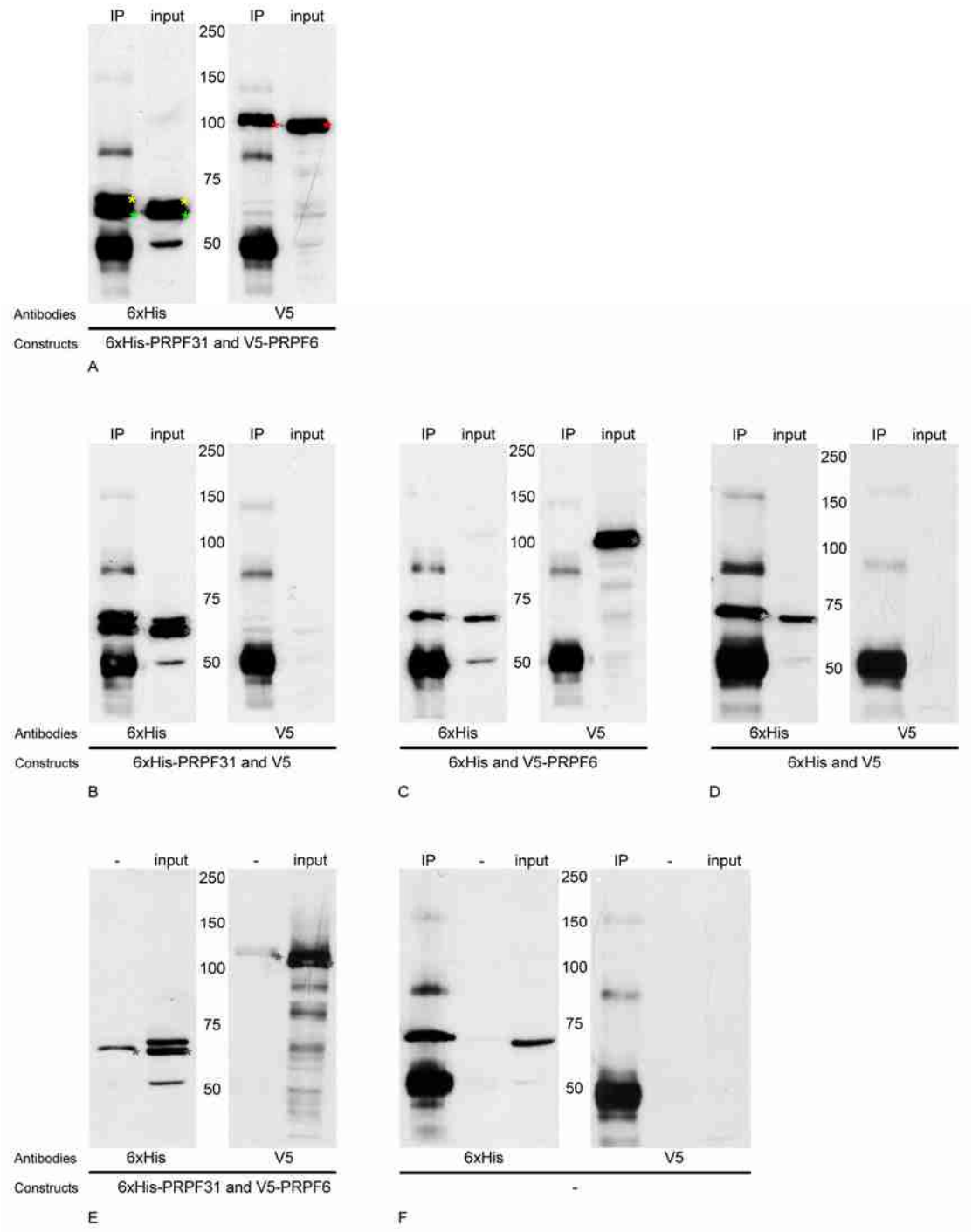


Figure 4.3. PRPF31 associates with PRPF6, as shown by immunoprecipitation.
Gel lanes are: IP, anti-6xHis immune pellet; input: transfected cell homogenate; -: non-immune pellet/negative control. Pellets and cell homogenates were analysed by immunoblotting with the antibody specificities as indicated under each blot, "antibodies" (anti-6xHis antibody is from Sigma); the type of homogenate analysed is stated under the "constructs". Coloured asterisks highlight relevant bands (green for 6xHis-PRPF31, red for V5-PRPF6, yellow for the un-specific band, as explained in text). The positions of molecular mass standards are indicated in kDa.

A mock immunoprecipitation for construct combination #1 was also performed, as illustrated in panel E: the anti-6xHis immunoreactive band at ~65 kDa is still observed on the non-immune pellet lane ("-", negative control). However the upper band observed in the previous panels (A, B, marked by the yellow asterisks) is not detected. The anti-V5 immunoreactive band at ~100 kDa is detected only in the input lane, although a faint band is present also in the negative control (red asterisks). These results may indicate a modest binding of 6xHis-PRPF31 to the beads, which did not strongly affect the result shown in panel A (care was taken to ensure all the gels -blots A and E- were equally loaded).
Panel F is dedicated to lysates of mock transfected cells, which were used to produce immune-pellets and non-immune pellets: anti-6His immunoreactivity is observed at ~65 kDa in both the IP and input lanes, but not in the negative control; this allowed us to conclude that the upper band observed in panels A and B is due to a protein species other than 6xHis-PRPF31, but which binds to the anti-6xHis antibody. No significant V5-immunoreactivity was observed.

As the membranes incubated with the anti-6xHis antibody revealed an extra band close the molecular weight of 6xHis-PRPF31 (see Figure 4.3, yellow asterisks), the samples were re-electrophoresed on a gel and probed with an anti-PRPF31 antibody (Abcam) and anti-PRPF6 antibodies, for which the blots are shown in Figure 4.4.
Using the anti-PRPF31$_{140\text{-}154}$ antibody (Abcam), an immunoreactive band at ~65kDa (6xHis-PRPF31) was detected in the immune pellets corresponding to construct combinations #1 and #2 (panels A and B respectively, green asterisks). The anti-PRPF6 immunoreactivity with ~110 kDa (V5-PRPF6) was detected in the immune pellets corresponding to construct combinations #1 (panel A, red asterisks), confirming the interaction observed in the previous experiment, but not in the IP lane of panel B (where only the empty V5 vector was transfected), or panel C (construct combination #3) where 6xHis-PRPF31 is not expressed. Panel D shows only the endogenous PRPF31 and PRPF6, as only empty vectors were transfected.
In panel E, a very faint band at ~65 kDa (green asterisk) is visible in the IP lane of PRPF31 blot, validating what was previously observed (panel E, Figure 4.3), that:

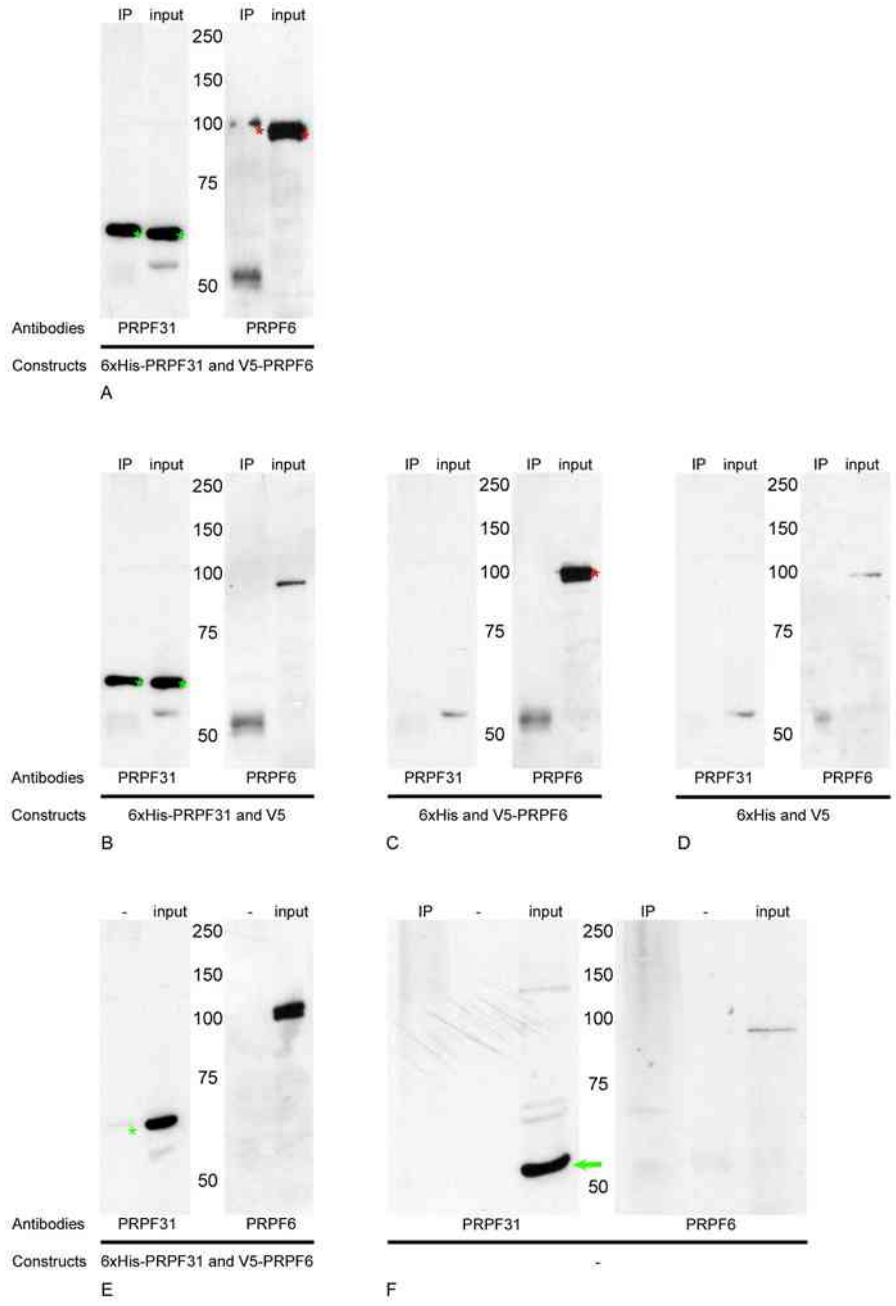


Figure 4.4. PRPF31 associates with PRPF6, as shown by immunoprecipitation.
Gel lanes are: IP, anti-6xHis immune pellet; input: transfected cell homogenate; -: non-immune pellet/negative control. Pellets and cell homogenates were analysed by immunoblotting with the antibody specificities as indicated under each blot, "antibodies"; and the type of homogenate analysed is stated under "constructs". Coloured asterisks highlight relevant bands (green for 6xHis-PRPF31, red for V5-PRPF6, as explained in text). The positions of molecular mass standards are indicated in kDa.

6xHis-PRPF31 is able to bind the beads.

No immunoreactivity for the anti-PRPF31$_{140-154}$ antibody (Abcam) is detected on either the immune-beads or the non-immune beads, as shown on panel F (PRPF31 blot), therefore concluding the un-specific nature of the band observed at ~65kDa in panel F (6xHis blot) of Figure 4.3. In panel F of Figure 4.4, the only band visible on the input (green arrow, PRPF31 blot) is the band corresponding to the PRPF31 protein, confirmed by its position at ~61 kDa (as indicated by the green arrow).

The experiments shown in this section proved that the anti-6xHis antibody (Abcam) was unsuitable for immunoprecipitation, and a different, perhaps monoclonal antibody could have been more suitable for the purpose, avoiding un-specificity.

4.3.2 Co-immunoprecipitation of 6xHis-PRPF31 and V5-MAGI3, 6xHis-PRPF31 and V5-TRAK2

Once the conditions were optimised for successful co-IP in detecting the 6xHis-PRPF31 + V5-PRPF6 interaction, they were applied to investigate any association between 6xHis-PRPF31 + V5-MAGI3, and 6xHis-PRPF31 + V5-TRAK2. As in section 4.3.1, HEK293 cells were co-transfected with the following combinations of DNA constructs (ratio 1:1):

1. pDEST26-PRPF31 + pcDNA3.1V5-MAGI3;
2. pDEST26-PRPF31 + pcDNA3.1V5-TRAK2;
3. pDEST26-PRPF31 + pcDNA3.1V5DEST; (control: PRPF31 and empty V5 vector);
4. pDEST26 + pcDNA3.1V5-MAGI3; (control: empty 6xHis vector and MAGI3);
5. pDEST26 + pcDNA3.1V5-TRAK2; (control: empty 6xHis vector and TRAK2);
6. pDEST26 + pcDNA3.1V5DEST; (control: empty 6xHis and V5 vectors).

Figure 4.5 shows association between PRPF31 and MAGI3, and between PRPF31 and TRAK2, upon probing with anti-tag antibodies. An anti-6xHis immunoreactive band at ~65 kDa (6xHis-PRPF31) was detected in the immune pellets corresponding to constructs PRPF31 and MAG13 co-transfection (combination 1), PRPF31 and TRAK2 (combination 2) and PRPF31 and empty V5 vector (combination 3), as shown in Figure

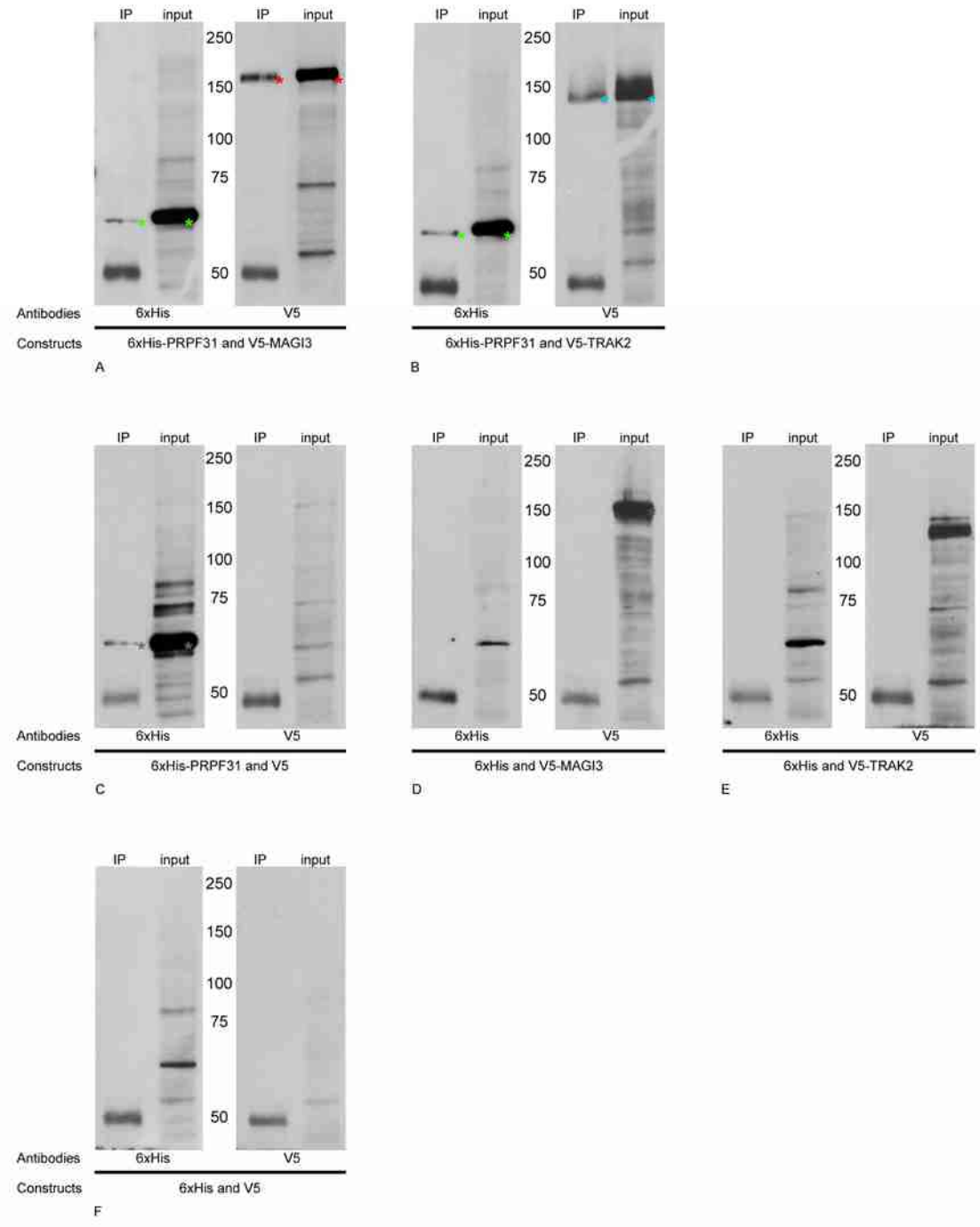


Figure 4.5. PRPF31 associates with MAGI3 and with TRAK2.
Gel lanes are: IP, anti-6xHis immune pellet; input: transfected cell homogenate. Pellets and cell homogenates were analysed by immunoblotting with the antibody specificities as indicated under each blot, "antibodies"; the type of homogenate analysed is stated under the "constructs". Red asterisks point out V5-MAGI3; blue asterisks point out V5-TRAK2; green asterisks point out 6xHis-PRPF31. The positions of molecular mass standards (kDa) are shown on the centre.

4.5, panels A, B and C respectively (green asterisks). An anti-V5 immunoreactive band at ~150 kDa (V5-MAGI3) was detected in the immune pellets corresponding to constructs PRPF31 and MAG13, but not constructs empty-6xHis-vector and MAGI3 (combinations 1 and 4, panels A and D respectively, marked by red asterisks). The anti-V5 immunoreactive band at ~120 kDa (V5-TRAK2) was detected in the immune pellets corresponding to construct PRPF31 and TRAK2, but not empty-6xHis-vector and TRAK2 (combinations #2 and #5, panels B and E respectively, indicated by blue asterisks).

As previously observed, also this experiment confirms the presence of an extra band around the molecular weight of 6xHis-PRPF31 (see Figure 4.5, 6xHis blot in panel C as an example), but it does not affect the outcome of the immunoprecipitation between 6xHis-PRPF31 and the interactants.

Table 4.1 summarises the protein combinations and the outcome (interaction or non-interaction), when the anti-6xHis antibody (Abcam) was coupled to the beads.

Table 4.1. Summary of co-immunoprecipitations coupling anti-6xHis antibody (Abcam) to the beads.

	V5-PRPF6	V5-MAGI3	V5-TRAK2	V5
	Section 4.3.1	Section 4.3.2		
6xHis-PRPF31	+	+	+	-
6xHis	-	-	-	-

The table provides an overview of the various combinations between 6xHis-PRPF31/6xHis and V5-proteins, stating in which section each interaction is described. It also provides a first glance of the immunoprecipitation outcomes, where "+" stands for interaction and "-" for no interaction observed.

4.4 Reverse Co-immunoprecipitation

To further validate these interactions, reverse co-immunoprecipitation was performed. In the previous section the immunoprecipitation reactions were performed with the anti-6xHis antibody conjugated to the beads. On the reverse co-IP, it is the anti-V5 antibody coupled to the beads, and therefore the V5-tagged protein that is linked to the solid substrate.

4.4.1 Co-immunoprecipitation of V5-MAGI3 and 6xHis-PRPF31, V5-TRAK2 and 6xHis-PRPF31

This experiment was performed on HEK293 cells transfected with the same combinations of DNA constructs stated in section 4.3.2. Figure 4.6 shows the immunoblots of the co-immunoprecipitation reactions as well as inputs, although on separate blots: panels A and B show the immune pellets of the different cell homogenate type stated on "constructs", whereas panels C and D represent the input blots. The membranes were probed twice: first with the anti-V5 antibody, and subsequently with the anti-6xHis antibody (from Sigma), and due to which a background is observed on blots B and D.

Upon incubation with the anti-V5 antibody, protein bands corresponding to V5-MAGI3 (red asterisks) and V5-TRAK2 (blue asterisks) were detected both in the IP blot (panel A) and in the input blot (panel C), confirming that the IP reaction took place in the first instance. However, the anti-6xHis reactive band at ~65 kDa (for 6xHis-PRPF31) was observed not only where expected (when 6xHis-PRPF31 was expressed together with a V5-protein, shown by green asterisks in IP panel B), but also when expressed together with the V5 tag alone (indicated by a green arrow in panel B). Hence a cross-reaction between the anti-V5 antibody and 6xHis-PRPF31 must have occurred.

Panel A confirms what is observed in panel B (green arrow): an anti-V5 immunoreactive band at ~65 kDa was detected in the immune pellets corresponding to construct combinations #1, #2 and #3, that is when 6xHis-PRPF31 is expressed (green arrows in panel A), but not #4, #5 and #6, that is when only 6xHis tag is produced (in this case, it is not possible to see a band corresponding to the 6xHis tag only as it is too small). This may be due to a cross reaction between the anti-V5 antibody and the 6xHis tag.

4.4.2 Co-immunoprecipitation of V5-MAGI3 and Wild Type PRPF31, V5-TRAK2 and Wild Type PRPF31

To evaluate whether or not the anti-V5 antibody cross-reacted with the 6xHis tag and not with PRPF31 protein, and furthermore to prove that the wild type protein also co-precipitated with each interactant, a set of immunoreactions prepared corresponding to construct combinations #4, #5 and #6 (that is the cells expressing the V5-proteins and

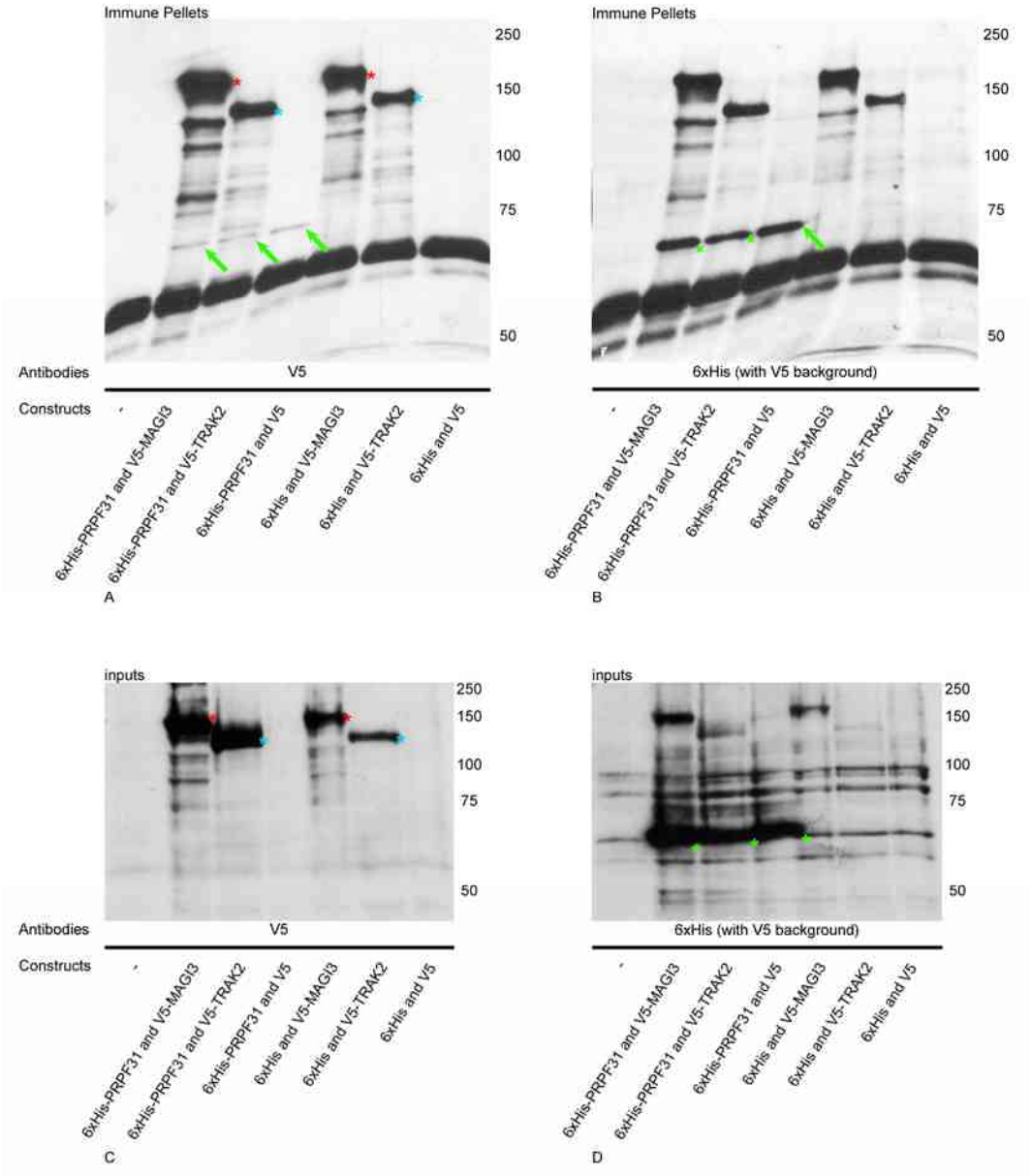


Figure 4.6. Cross reaction between the anti-V5 antibody and 6xHis-PRPF31, as shown by immunoprecipitation.
Immune pellets (A and B) and inputs (cell homogenates; C and D) were analysed by immunoblotting with the antibody specificities as shown (the membranes were probed first with an anti-V5 antibody and secondly with an anti-6xHis antibody (Sigma), therefore the 6xHis immunoblots have a background from the previous incubation with the anti-V5 antibody); the type of homogenate analysed is stated at "constructs". Red asterisks indicate V5-MAGI3; blue asterisks indicate V5-TRAK2; green asterisks and green arrows indicate 6xHis-PRPF31. The positions of molecular mass standards (kDa) are shown on the right

the empty 6xHis tag vector), were probed with the anti-PRPF31$_{484-497}$ antibody.

The anti-PRPF31$_{484-497}$ immunoreactive band at ~61 kDa (wild type PRPF31) was detected in the immune pellets corresponding to construct combinations #4 and #5 (expressing V5-MAGI3 and V5-TRAK2 respectively, shown in panels A and C of Figure 4.7, green asterisks), but no detectable immunoreactivity in the immune pellet for combination #6 (cells expressing both tags only; shown in panels B and D), confirming there was no cross-reaction between the anti-V5 antibody and wild type PRPF31.

Therefore, not only was the wild type splicing factor found complexed with the V5-interactant, corroborating the previous results, but also it was not implicated in any cross-reaction with the anti-V5 antibody, the cross-reaction can then be attributed to only the 6His tag.

Table 4.2 summarises the results of the co-IPs showed in the sections 4.4.1 and 4.4.2:

Table 4.2. Summary of co-immunoprecipitations coupling anti-V5 antibody to the beads.

		V5-MAGI3	V5-TRAK2	V5
6xHis-PRPF31	Section 4.4.1	+	+	+
6xHis		-	-	-
wt PRPF31	Section 4.4.2	+	+	-

The table offers a synopsis of the various combinations between the V5-proteins and the 6xHis-PRPF31/6xHis/wtPRPF31, referring to the section each interaction is described. It also provides a first glance of the immunoprecipitation outcomes, where "+" stands for (potential) interaction and "-" for no interaction observed.

4.5 Conclusion

Based on the results presented so far, MAGI3 and TRAK2 may be considered as partners of PRPF31 *in vitro*; however, to undoubtedly consider them as part of the PRPF31 protein complex, the interaction must be further proven *in vivo*.

Before attempting any further experiment to confirm or rule out MAGI3 and TRAK2 as true interacting partners of PRPF31, conditions were optimised against a known PRPF31 interactant, V5-PRPF6 (Makarova et al., 2002). Results showed that PRPF6 was successfully detected in the immune pellets. Similarly, the co-immunoprecipitation

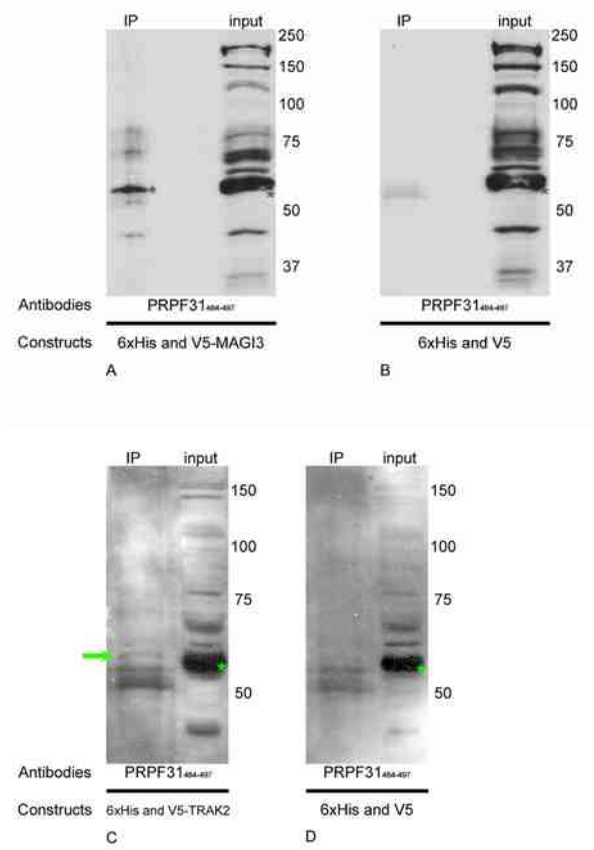


Figure 4.7. Wild type PRPF31 associates with MAGI3 and with TRAK2, as shown by immunoprecipitation.
Gel lanes are: IP, anti-V5 immune pellet; input, transfected cell homogenate. Pellets and cell homogenates were analysed by immunoblotting with the antibody specificities as shown in the abscissae; the type of homogenate analysed is stated at the "constructs" abscissae. The positions of molecular mass standards (kDa) are shown on the right.

reactions between 6xHis-PRPF31 and V5-MAGI3 and between 6xHis-PRPF31 and V5-TRAK2 were also successful, when pulled down using the 6xHis species. The interaction was also proven by the reverse reaction: when the immune pellet was pulled down using the V5-species, wild type PRPF31 was detected in the complex.

These experiments also revealed a cross-reaction between the anti-V5 antibody and the 6xHis tag, which hindered detection of the interactants in reverse co-immunoprecipitation conditions.

5. Co-localisation of PRPF31 with MAGI3 and TRAK2

In order to validate *in vivo* the interaction of PRPF31 with MAGI3 and TRAK2, co-localisation in cells was performed, using both tagged and endogenous proteins.

For this purpose, MDCK cells were transfected with the same constructs as for the co-immunoprecipitation experiments (encoding N-terminally tagged proteins), to observe potential associations through the immunofluorescence technique. As the HEK293 cell line shows a low adhesion to glass coverslips, the MDCK epithelial cell line was preferred. MDCK cells grow in a monolayer and strongly adhere to glass coverslips, rendering this cell line more suitable for microscopic analysis than the HEK293 cell line. As mentioned at the beginning of Chapter 4, experiments were carried out even if the interacting candidates were not retina-specific: still, new information can be gathered on PRPF31 role. That is also why we selected the MDCK cell line as a model: non-neuronal, to identify whether the interactions are happening not just in particular cell types, since all three proteins (PRPF31, MAGI3, TRAK2) are expressed ubiquitously; differentiated, to be able to infer the significance of any hypothetical co-localisation of PRPF31 with the interactants.

To ensure results were robust, we performed a test staining in order to ascertain whether the cross-reaction between the 6xHis-tag and the anti-V5 antibody occurred also in the immunofluorescence method, since it was observed in western blot and co-IP procedures. The experiment is described in the “Preliminary Work” section (section 5.2.1).

The chapter also addresses the localisation of the endogenous proteins: PRPF31 and MAGI3 were investigated in MDCK cells, whereas a potential association between PRPF31 and TRAK2 was examined in the SK-N-SH cell line; the choice of the SK-N-SH cell line was dictated by the fact that TRAK2 is enriched in neuronal tissue, and it was not detected at satisfying levels in MDCK cells.

5.1 Expression of Tagged Proteins in MDCK Cells

To verify the expression of the tagged proteins, MDCK cells were cultured in 24-well plates, transfected with the combinations of constructs as used for the co-IP experiments (see section 4.3.2 of Chapter 4), using the Lipofectamine 2000 reagent (section 2.12.2). After 24 hours, the cells were harvested in PBS-0.5% TX100 (according to protocol in section 2.12.3). The cell lysates were then treated with sample buffer for SDS-PAGE (section 2.11).

Figure 5.1 shows the expression of the tagged proteins in MDCK cells. The anti-6xHis antibody (Sigma) revealed bands at ~65 kDa likely corresponding to 6xHis-PRPF31 (indicated by red asterisks on blot A); lanes corresponding to cells transfected with only the 6xHis tag, mock transfected cells or wild type cells show an overlapping band, at the molecular weight of 6xHis-PRPF31 (~65 kDa), suggestive of non-specificity. Anti-V5 antibody incubation revealed bands corresponding to V5-MAGI3 (~150kDa), V5-TRAK2 (~120kDa) and V5-PRPF31 (~70kDa).

5.2 Co-localisation of 6xHis-PRPF31 and V5-MAGI3, 6xHis-PRPF31 and V5-TRAK2

MDCK cells were cultured on sterile glass coverslips placed in 24-well plates, and transfected with Lipofectamine 2000 reagent according to protocol in section 2.12.2. After 24 hours, cells were fixed with methanol and stained (as described in sections 2.12.4 and 2.14.1 respectively). Over-expressed proteins were detected upon incubation with the anti-V5 and the anti-PRPF31$_{484\text{-}497}$ antibodies. We chose not to use any anti-6xHis antibody (Sigma and Abcam) to detect the 6xHis-tagged protein as the outcoming signals were indistinct (not shown).

The localisation experiments involving over-expressed protein (paragraphs 5.2 and 5.3) were performed with the following secondary antibodies: a donkey Cy3-conjugated anti-rabbit secondary antibody to detect the anti-PRPF31$_{484\text{-}497}$ antibody, and a donkey FITC-conjugated anti-mouse secondary antibody to detect the anti-V5 antibody.

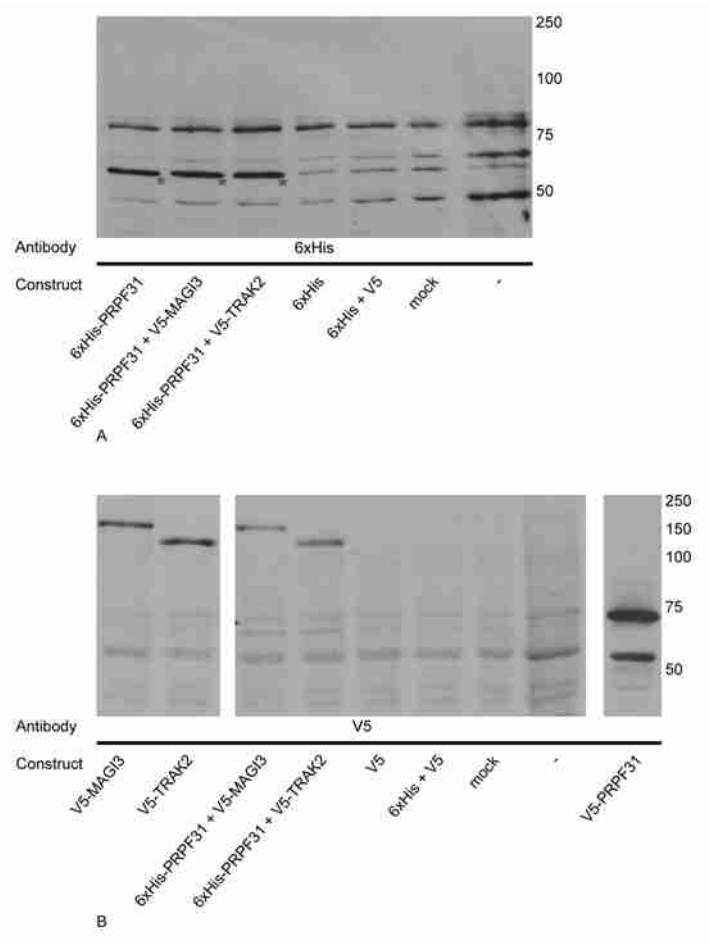


Figure 5.1. Expression of tagged proteins in MDCK cells.
In A, detection of tagged proteins upon incubation with the anti-6xHis antibody (Sigma): the blot shows three strong bands at ~65 kDa corresponding to 6xHis-PRPF31 (red asterisks). In B, the detection of tagged proteins upon incubation with the anti-V5 antibody revealed bands at ~150 kDa for V5-MAGI3, at ~120 kDa corresponding to V5-TRAK2 and at ~70 kDa for V5-PRPF31.

5.2.1 Preliminary Work

To rule out a cross reaction between the anti-V5 antibody and the 6xHis tag, the pcDNA3.1V5DEST vector was used to generate the V5-PRPF31 construct. A staining of two sets of cells, expressing 6xHis-PRPF31 or V5-PRPF31 each, was performed to ascertain whether the cross-reaction still persisted in immunofluorescence techniques.

In Figure 5.2, panels A1 and B1 provide two examples of cells expressing V5-PRPF31, detected with the anti-V5 antibody: the distribution of the protein was observed to be both nuclear, in speckles, and in round aggregates resembling the nucleoli - and cytoplasmic too; in each case, the level of expression was high and the protein was easily detected.

Cells expressing 6xHis-PRPF31 were also stained with the anti-V5 antibody, to examine the possible occurrence of a cross-reaction in these conditions: the staining revealed no cross-reaction, as no fluorescence was detected (Figure 5.2, panel C); if any signal was to be seen, it would have resembled the one shown in panels A and B. The fact that no fluorescence was detected indicates that the anti-V5 antibody recognised its specific epitope, even though 6xHis-PRPF31 and V5-tagged protein were simultaneously present.

Furthermore, the V5 tag was expressed in cells, in order to assess whether an N-terminal tag would affect PRPF31 localisation. Although PRPF31 was expressed with the 6xHis tag, its detection was not clear using an anti-6xHis antibody in the conditions adopted for immunofluorescent staining. We therefore expressed the V5 tag instead, since its detection was more feasible. The anti-V5 staining did not show any cytoplasmic or nuclear staining pattern (panel D of Figure 5.2), thus confirming that the tag does not alter the protein localisation. However, a small peptide such as the V5 tag may be quickly degraded by the cell, making its detection difficult.

Panel E illustrates cells which underwent treatment with the transfecting agent alone; after methanol fixation, cells were incubated with anti-V5 antibody followed by donkey FITC-conjugated anti-mouse secondary antibody. No fluorescence was detected.

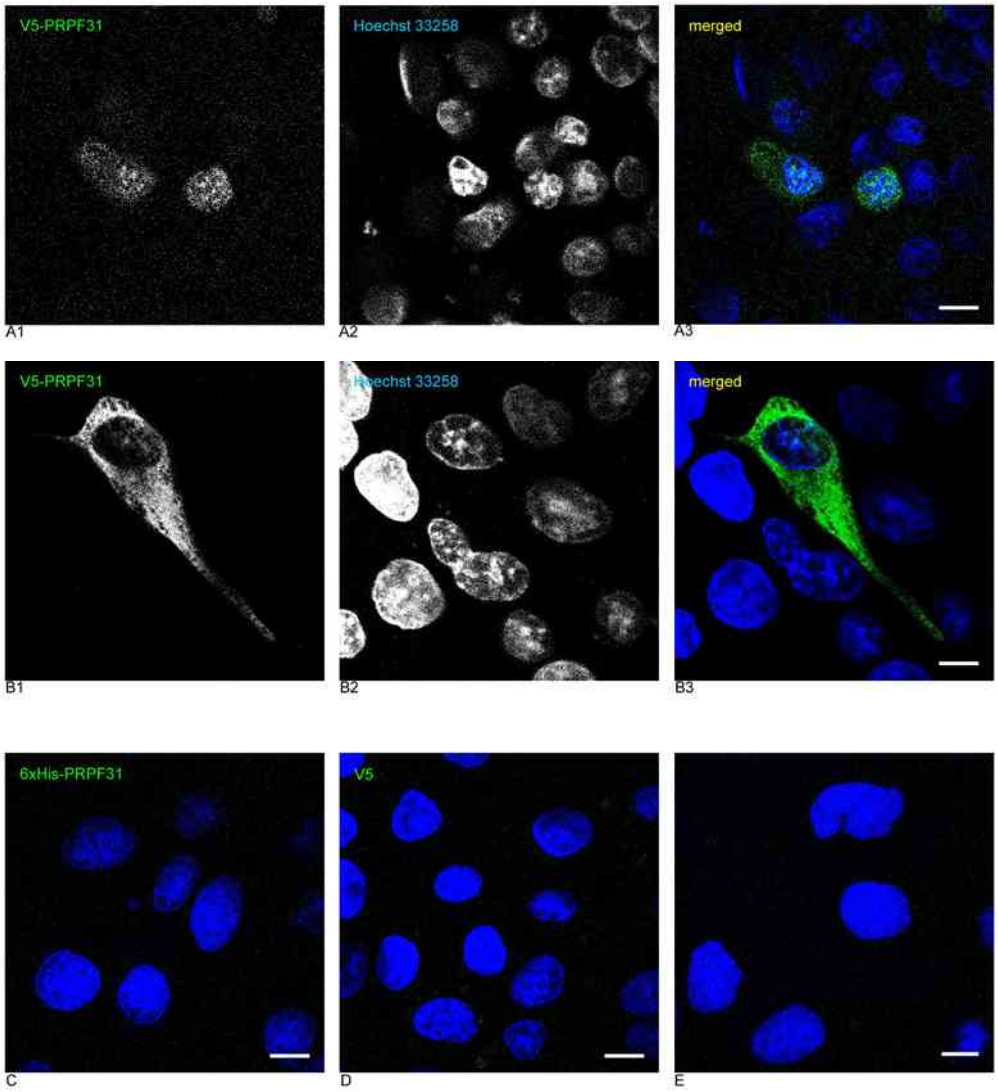


Figure 5.2. V5-PRPF31 and 6xHis-PRPF31 localisation in MDCK cells, using anti-V5 antibody. V5-PRPF31 (FITC signal) was localised in nuclear speckles (A1) and in the cytoplasm (A1 and B1). In C, the presence of 6xHis-PRPF31 is not detected by the anti-V5 antibody. In D, cells expressing only the V5 peptide do not show any significant staining upon incubation with the anti-V5 antibody. In E, mock transfected cells incubated with the anti-V5 antibody. Nuclei were stained with Hoechst 33258 (blue signal). Scale bars, 10μm. Colour is indicated by legend on each black and white panel.

5.2.2 Localisation of 6xHis-PRPF31 in MDCK Cells

As no cross reaction between 6xHis-tag and the anti-V5 antibody was observed (as described in the previous section 5.2.1), we proceeded with the localisation of 6xHis-PRPF31, using the anti-PRPF31$_{484-497}$ antibody. Figure 5.3 shows that the staining of 6xHis-PRPF31 is similar to that observed for V5-PRPF31 (compare with Figure 5.2). The transfected cells showed both a nuclear localisation (in small speckles and in round aggregates resembling the nucleoli) and a distinct cytoplasmic presence.

The expression of the tagged protein, and hence the signal intensity of the PRPF31, was very high and the exposure time of the image was adjusted in accordance with this; for this reason no signal was observed from the endogenous PRPF31 in the mock transfected cells (panel D).

5.2.3 Localisation of V5-MAGI3 in MDCK Cells

In epithelial cells, the distribution of the V5-MAGI3 protein was predominantly cytoplasmic (Figure 5.4): in fact the expressed protein localised strongly at intercellular junctions - likely tight junctions as previously reported by Wu *et al.* (2000) and Adamsky *et al.* (2003) - , and around the nucleus, its presence becoming progressively weaker towards the distal part of the cytoplasm. A less intense signal was detectable in the nucleus too. No formation of aggregates was observed. Figure 5.4 shows two examples of the localisation of the V5-MAGI3 protein.

5.2.4 Localisation of V5-TRAK2 in MDCK Cells

V5-TRAK2 was localised primarily in the cytoplasm (Figure 5.5): the signal was present throughout the entire cell. The signal was also observed in the nucleoplasm, although weakly. In cells expressing less V5-TRAK2 (judged by lower signal intensity), the protein was mainly detected in a perinuclear fashion; the signal also had a streaked pattern. Also for V5-TRAK2, no aggregates were observed. Approximately 50% of the cells expressing the protein, exhibited a very intense fluorescence at the end of cellular extension, detected as several speckles in a limited area (Figure 5.5, panel A1 upper portion of the cell as indicated by green arrows; panel B1, stronger signal in upper portion of the cell, weaker in the lower, as indicated by green arrows).

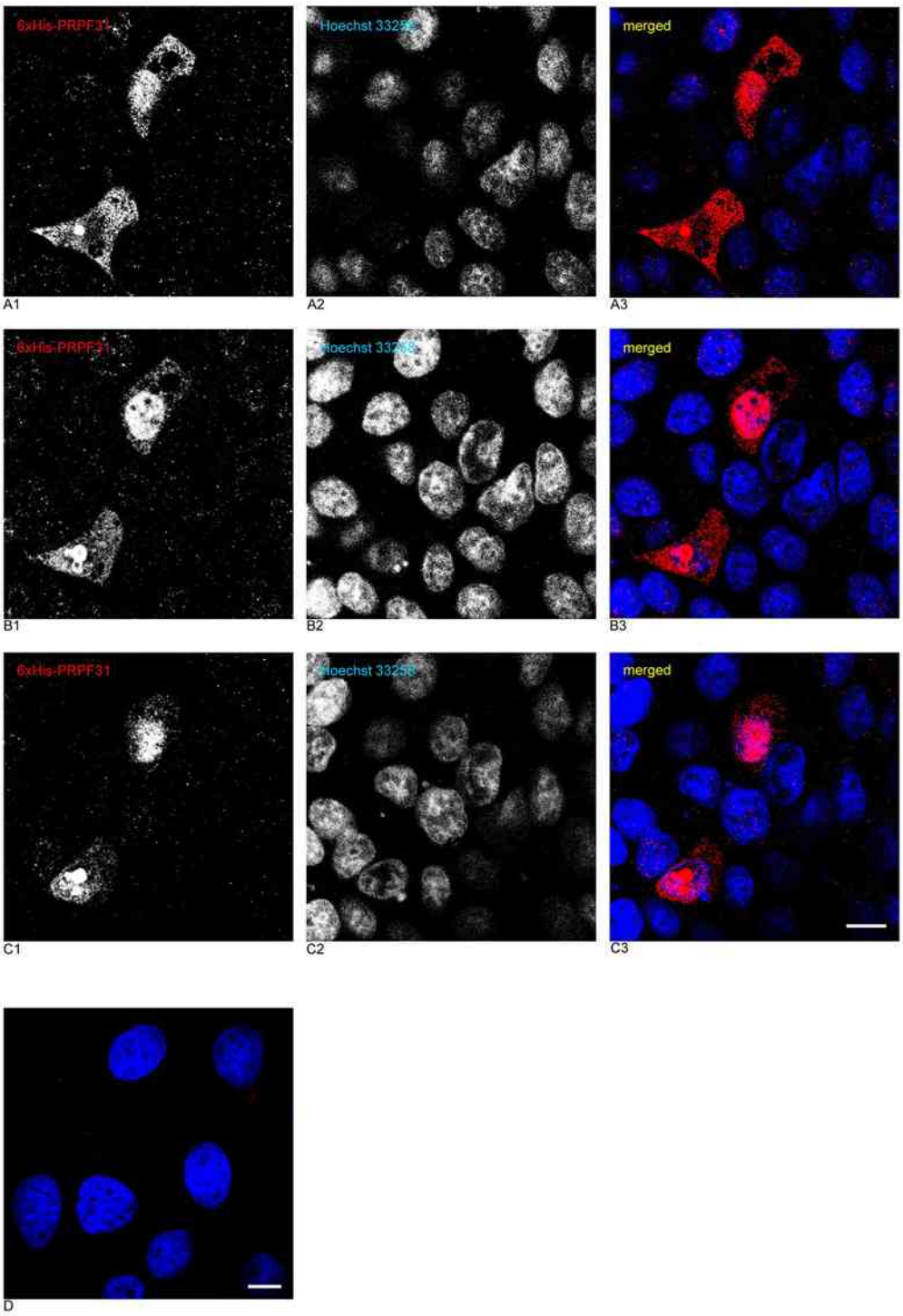


Figure 5.3. 6xHis-PRPF31 localisation in MDCK cells, using anti-PRPF31$_{484-497}$ antibody.
The series of panels represent the same area of MDCK cells, expressing 6xHis-PRPF31, detected with the anti-PRPF31$_{484-497}$ antibody (Cy3 signal). Panels A, B and C are three different *z* sections, of 0.70 µm each. Panel D represents mock transfected cells. Nuclei were stained with Hoechst 33258 (in blue). Scale bar, 10µm. Colour is indicated by legend on each black and white panel.

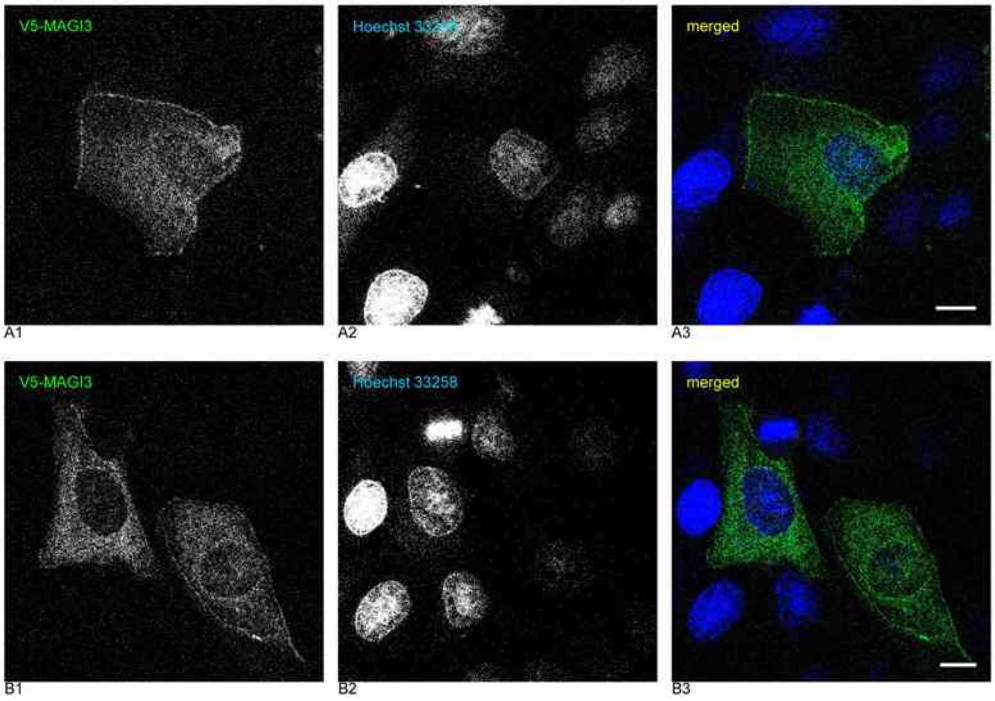


Figure 5.4. Localisation of V5-MAGI3 in MDCK cells, using the anti-V5 antibody.
A, B: cells over-expressing V5-MAGI3 (A1, B1; FITC) presented a staining pattern which included intercellular junctions (fluorescence at the edge of the cell), cytoplasm (stronger around the nucleus, progressively fading towards the cytoplasm), and a little the nucleus (visible in A1). Nuclei were stained with Hoechst 33258 (blue signal). Scale bars, 10µm. Colour is indicated by legend on each black and white panel.

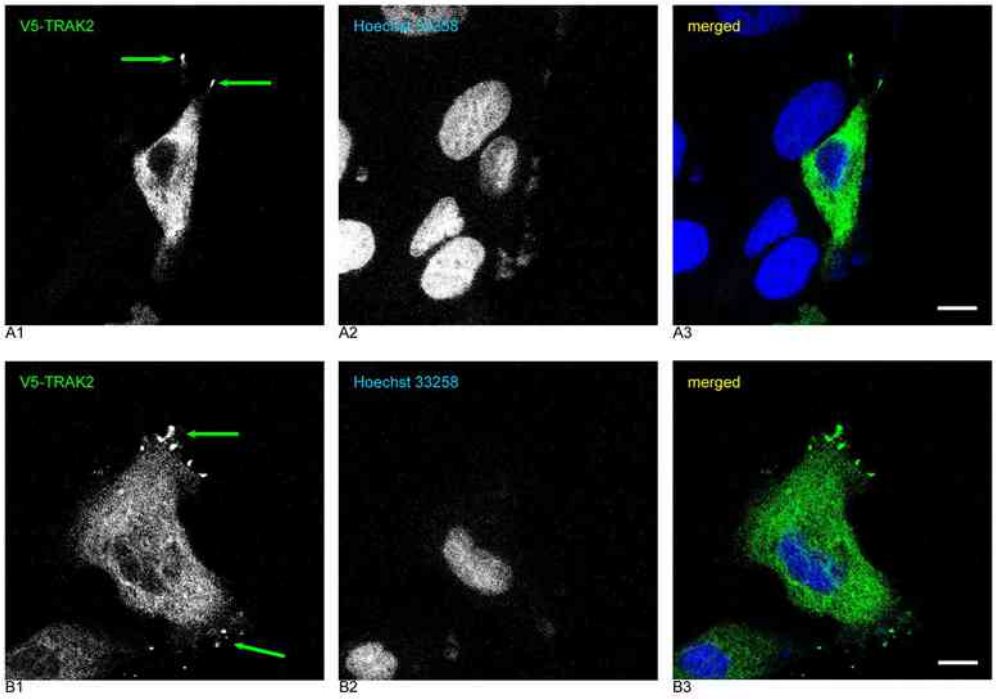


Figure 5.5. Localisation of V5-TRAK2 in MDCK cells, using the anti-V5 antibody.
A, B: cells over-expressing V5-TRAK2 (A1, B1; FITC) presented a staining pattern which included mainly the cytoplasm, and a little the nucleus (right-hand border of the nucleus in both A and B). Nuclei were stained with Hoechst 33258 (blue signal). Scale bars, 10µm. Colour is indicated by legend on each black and white panel.

5.2.5 Co-localisation of 6xHis-PRPF31 and V5-MAGI3

Figures 5.6, 5.7 and 5.8 show three examples of the localisation observed for both 6xHis-PRPF31 and V5-MAGI3, co-expressed in MDCK cells. It was observed that in the presence of V5-MAGI3, 6xHis-PRPF31 was expressed mainly in nuclear aggregates, with little or no cytoplasmic staining; less frequently, it was also observed dispersed throughout the nucleus, again with little cytoplasmic staining. On the other hand, V5-MAGI3 was detected largely at cell-cell junctions, and faintly in the cytoplasm, but also in nuclear aggregates together with 6xHis-PRPF31. According to the observations, the two recombinant proteins localised mainly to the nucleus, in large round aggregates, in which 6xHis-PRPF31 stained mostly the outer area and V5-MAGI3 was detected on the inner side: as the overlay shows, there is a region of overlap (denoted by a yellow colour), which suggests that they co-localise.

To summarise, two staining patterns were observed (100 co-transfected cells counted for each of three independent transfections):

- cytoplasmic, junctional and nuclear V5-MAGI3 and nuclear 6xHis-PRPF31 (~57% of cells co-expressing the proteins);
- cytoplasmic and junctional V5-MAGI3, and nuclear 6xHis-PRPF31 (~43% of cells co-expressing the recombinant proteins).

Because the anti-PRPF31$_{484\text{-}497}$ antibody was used for the detection of 6xHis-PRPF31, the observed fluorescence is the contribution of two signals, one originating from the 6xHis-tagged protein, and a second from the endogenous PRPF31. However, the abundance of the over-expressed protein was such that the exposure had to be adjusted, therefore the visibility of the endogenous signal is very reduced, explaining why it cannot be observed in non-transfected cells.

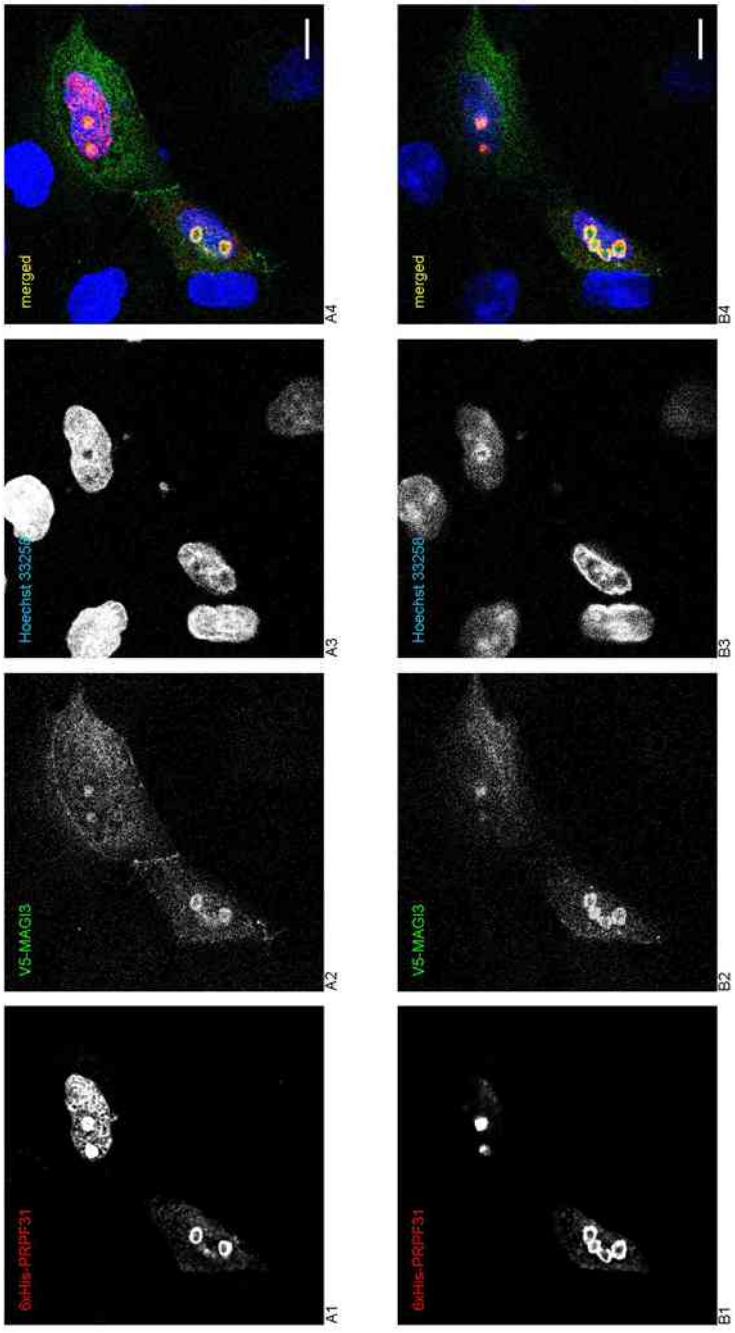


Figure 5.6. Co-localisation of 6xHis-PRPF31 and V5-MAGI3 in MDCK cells. Co-expression of 6xHis-PRPF31 (in A1, B1; Cy3) and V5-MAGI3 (in A2, B2; FITC), both proteins co-localised to the nucleus, particularly in round aggregates within the nucleus (yellow signal in A4, B4). Scale bars, 10µm. The two sets of panels represent two *z* sections (0.45µm) of the same area. Colour is indicated by legend on each black and white panel.

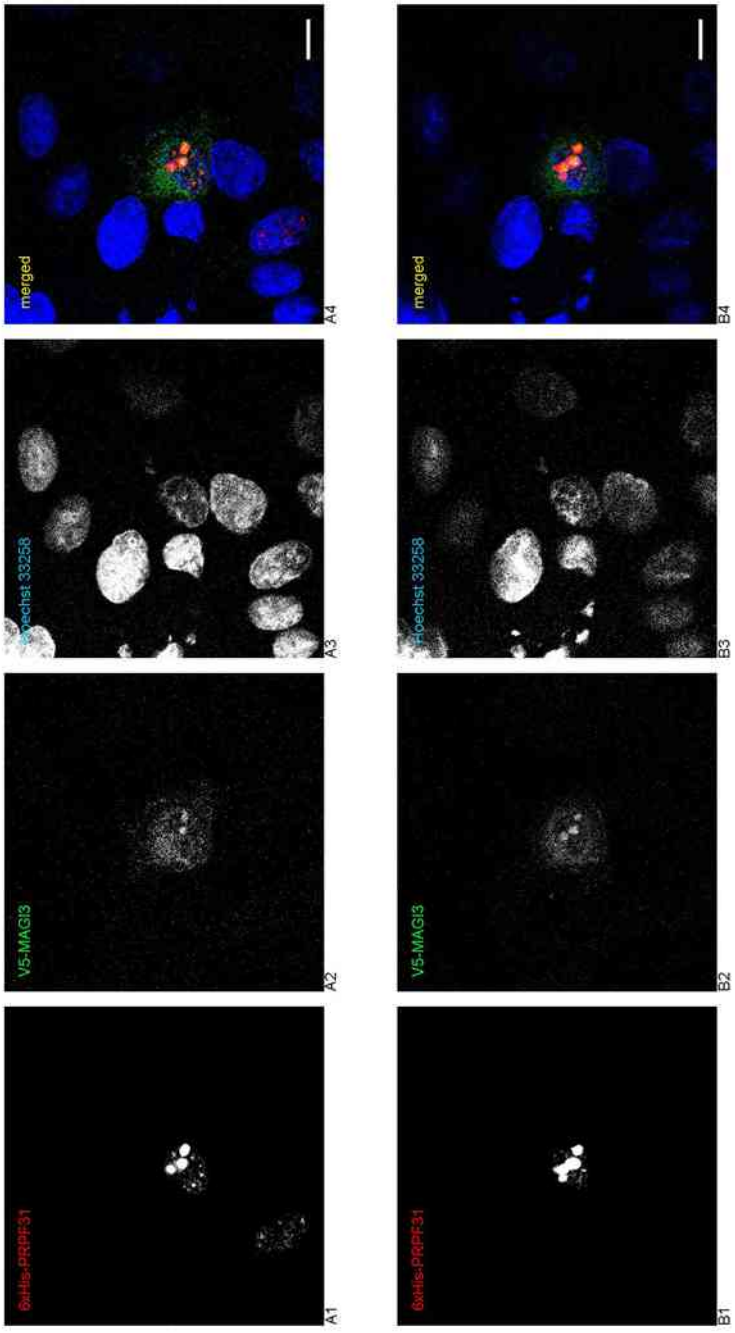


Figure 5.7. Co-localisation of 6xHis-PRPF31 and V5-MAGI3 in MDCK cells. Co-expression of 6xHis-PRPF31 (in A1, B1; Cy3) and V5-MAGI3 (in A2, B2; FITC), both proteins co-localised, particularly in round nuclear aggregates (yellow signal in A4, B4). Scale bars, 10µm. The two sets of panels represent two *z* sections (0.45µm) of an identical area. Colour is indicated by legend on each black and white panel.

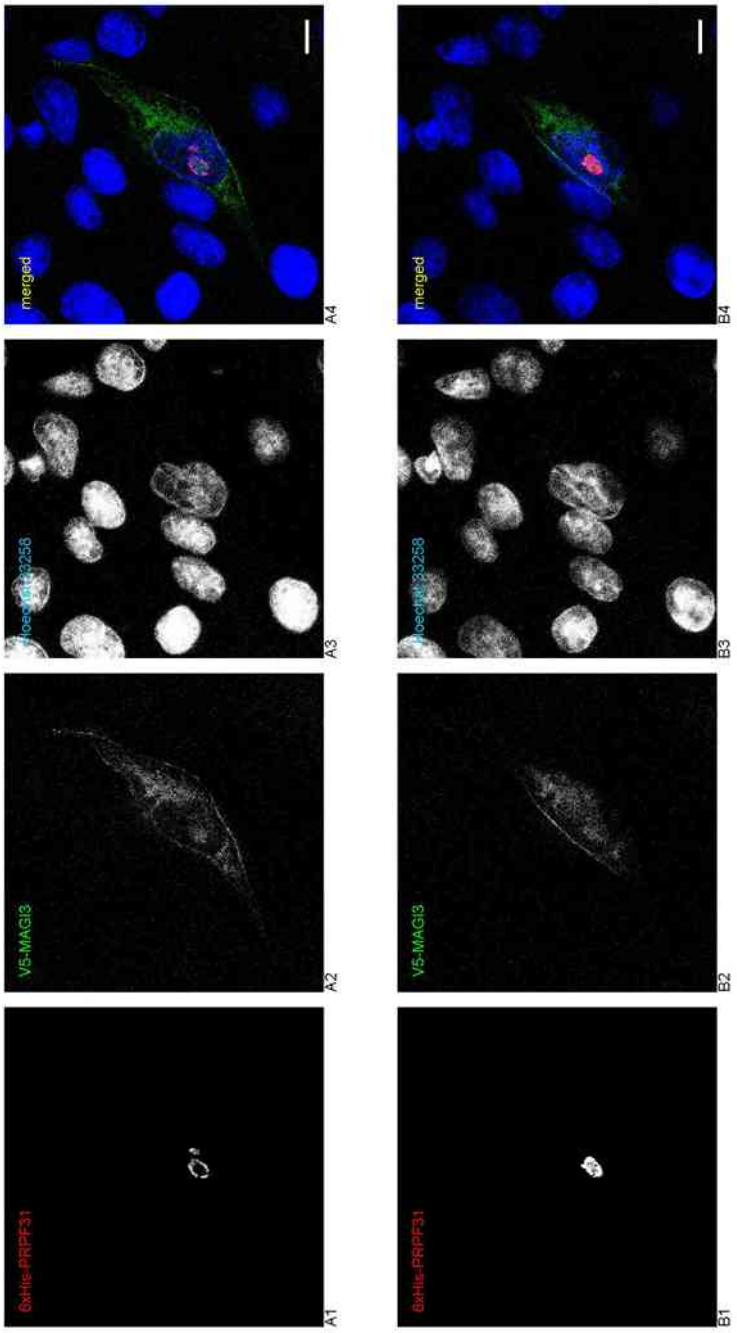


Figure 5.8. Localisation of 6xHis-PRPF31 and V5-MAGI3 in MDCK cells. Co-expression of 6xHis-PRPF31 (in A1, B1; Cy3) and V5-MAGI3 (in A2, B2; FITC), the proteins show the second type of localisation pattern: mainly cytoplasmic and junctional for V5-MAGI3, and "nucleolar" for 6xHis-PRPF31. Scale bars, 10µm. The two sets of panels represent two z sections (0.50µm) of the same area. Colour is indicated by legend on each black and white panel.

5.2.6 Co-localisation of 6xHis-PRPF31 and V5-TRAK2

Figures 5.9, 5.10 and 5.11 show three examples of the localisation observed for 6xHis-PRPF31 and V5-TRAK2, co-expressed in MDCK cells. In this case, 6xHis-PRPF31 exhibited a nuclear (speckles and aggregates) and a cytoplasmic presence. Interestingly, when co-expressed with V5-TRAK2, in a number of cells 6xHis-PRPF31 was detected at the tip of spike-like structures on cellular extensions (Figure 5.11 provides an example). V5-TRAK2 was detected in the cytoplasm, and it was frequently observed that its signal was mainly detected on one side of the cell. Furthermore, the V5-TRAK2 protein was detected on nuclear aggregates, co-localising with 6xHis-PRPF31 with a pattern similar to V5-MAGI3 (*i.e.* an inner staining of V5-TRAK2 and PRPF31 on the outside). Interestingly, V5-TRAK2 co-stained with 6xHis-PRPF31 on the aforementioned cellular extension (as indicated by white arrows in panel A4 of Figure 5.11).

The most prominent pattern of co-localisation observed was (on 100 co-transfected cells for each of three independent transfections):

- perinuclear signal for V5-TRAK2 with nuclear aggregates, and 6xHis-PRPF31 detected throughout the entire cell, and more strongly in nuclear aggregates (~58% of co-expressing cells counted).

The rest showed a weak staining to no staining of V5-TRAK2 in the nucleus, and 6xHis-PRPF31 present especially in nuclear aggregates.

The presence of 6xHis-PRPF31 in the cellular extensions was unexpected, as the single expression of the 6xHis-PRPF31 did not present such localisation. As was discussed in the previous section 5.2.4, the single expression of V5-TRAK2 showed this type of localisation; thus TRAK2 over-expression may alter the localisation of 6xHis-PRPF31, or it may enhance its presence in these regions (as it may be that there is a basal presence of the wild type splicing factor which is not detectable).

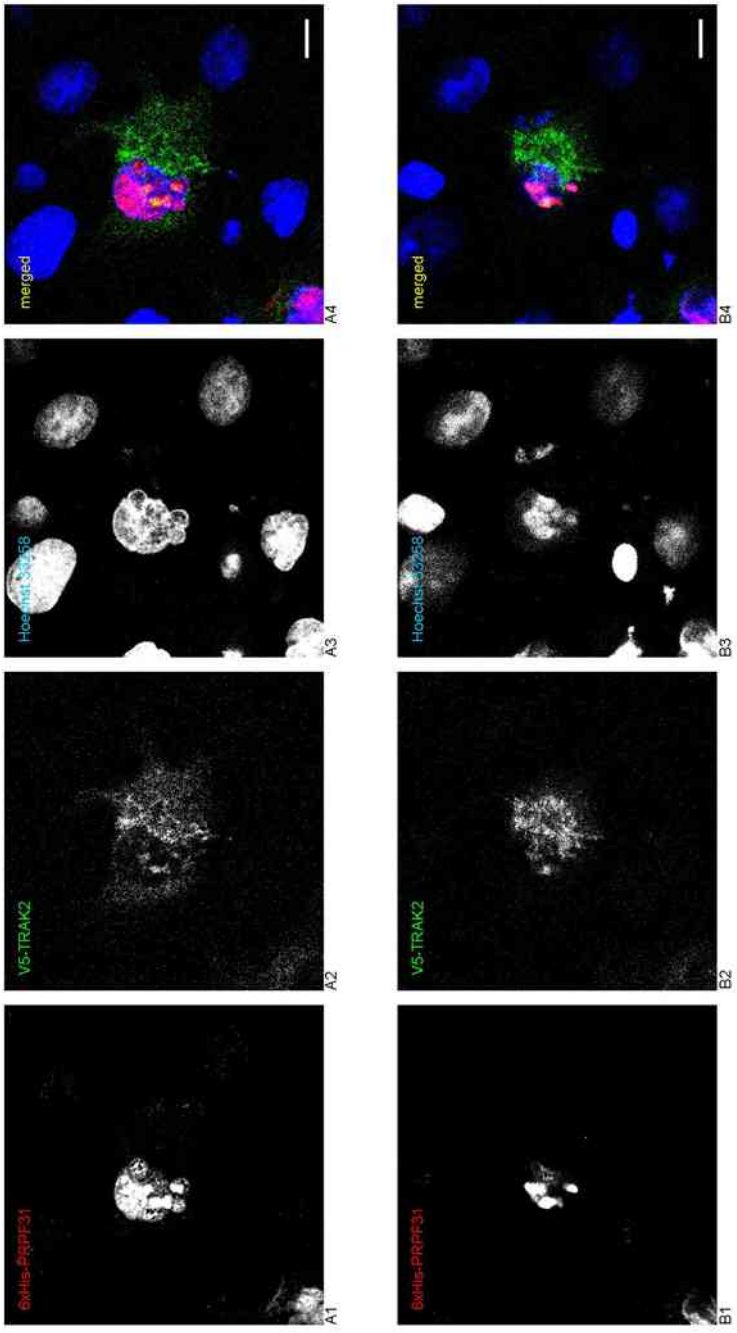


Figure 5.9. Co-localisation of 6xHis-PRPF31 and V5-TRAK2 in MDCK cells. Co-expression of 6xHis-PRPF31 (A1, B1; Cy3) and V5-TRAK2 (in A2, B2; FITC): the proteins co-localised to the nucleus, in round aggregates (yellow signal in A4, B4). Scale bars, 10µm. The two sets of pictures represent two *z* sections (0.45µm) of an identical area. Colour is indicated by legend on each black and white panel.

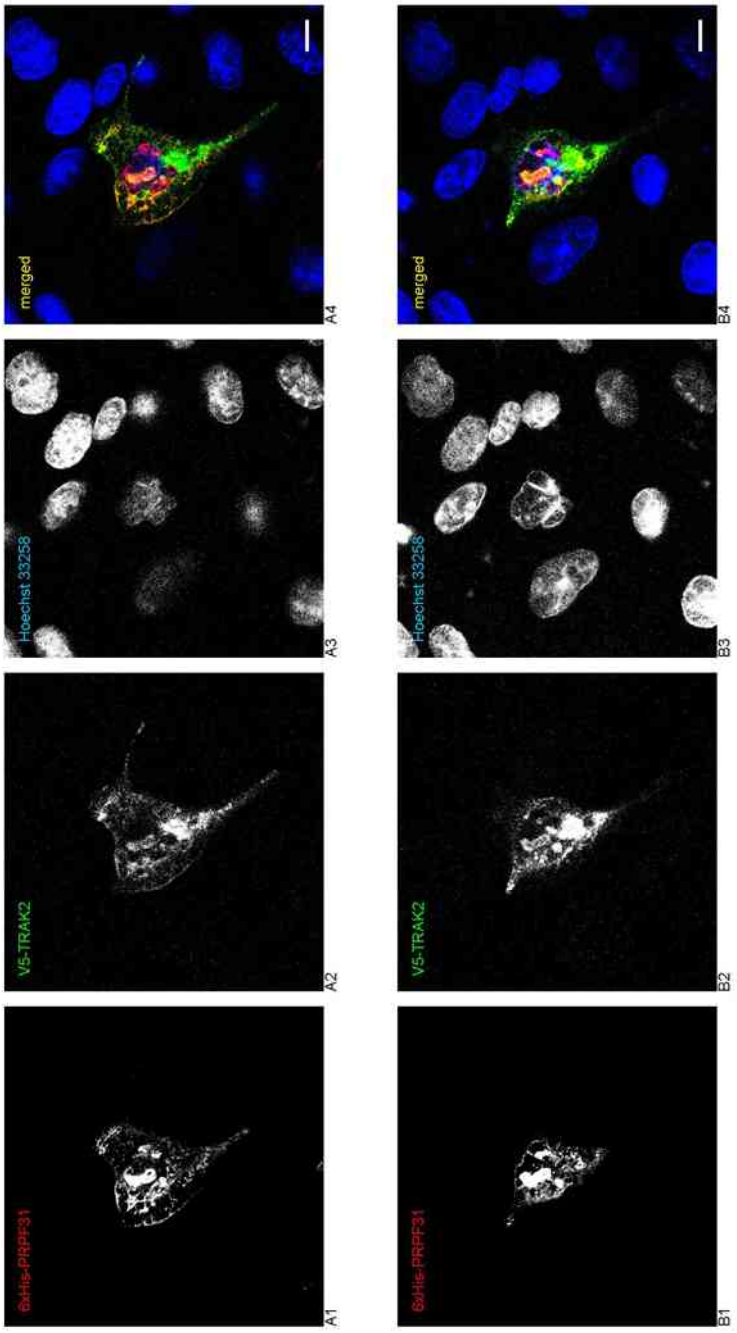


Figure 5.10. Co-localisation of 6xHis-PRPF31 and V5-TRAK2 in MDCK cells. Co-expression of 6xHis-PRPF31 (in A1, B1; Cy3) and V5-TRAK2 (in A2, B2; FITC), both proteins co-localised to the nucleus, resembling aggregates, and also to the cytoplasm (yellow signal in A4, B4). Scale bars, 10µm. The two sets of pictures represent two *z* sections (0.70µm) of the same area. Colour is indicated by legend on each black and white panel.

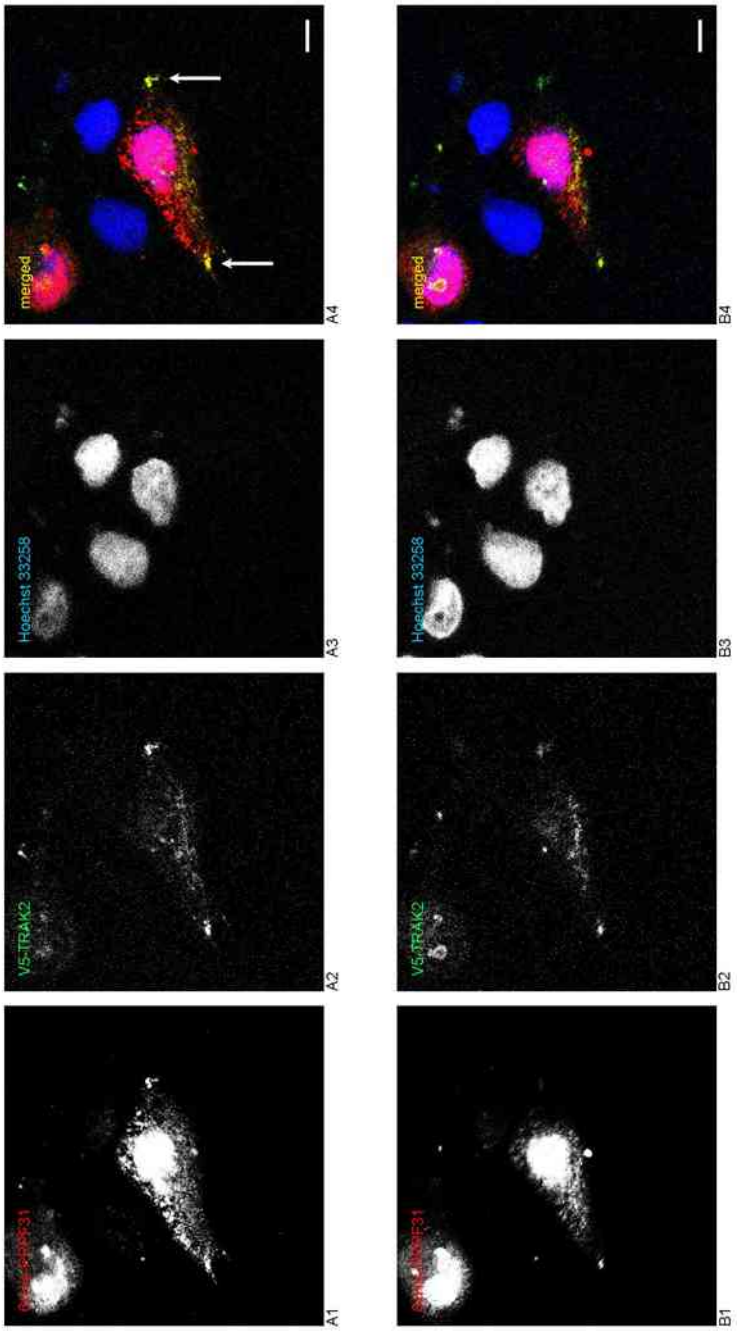


Figure 5.11. Co-localisation of 6xHis-PRPF31 and V5-TRAK2 in MDCK cells. Co-expression of 6xHis-PRPF31 (in A1, B1; Cy3) and V5-TRAK2 (in A2, B2; FITC), both proteins co-localised to the cytoplasm and at the tip of cellular extensions (yellow signal in A4, B4; indicated by white arrows in A4). Scale bars, 10µm. The 2 sets of pictures represent two z sections (0.30µm) of an identical area. Colour is indicated by legend on each black and white panel.

5.3 Exclusion of Cross-reaction between anti-PRPF31$_{484-497}$ and anti-V5 Antibodies

To ensure no cross-reaction between the antibodies and the epitopes was occurring, a set of control experiments were carried out for both 6xHis-PRPF31 + V5-MAGI3 transfection and 6xHis-PRPF31 + V5-TRAK2 transfection. Figure 5.12 summarises the controls: panels A, B and C are dedicated to 6xHis-PRPF31 + V5-MAGI3 transfection, whereas D, E and F are for 6xHis-PRPF31 + V5-TRAK2 transfection:

1. A, D: double-transfected cells were incubated with primary anti-PRPF31$_{484-497}$ only, followed by incubation with anti-rabbit-Cy3 and anti-mouse-FITC antibodies;
2. B, E: double-transfected cells were incubated with primary anti-V5 antibody only, followed by incubation with anti-rabbit-Cy3 and anti-mouse-FITC antibodies;
3. C, F: double-transfected cells were incubated with anti-rabbit-Cy3 and anti-mouse-FITC secondary antibodies (performed following a mock primary antibody incubation).

Any unexpected staining was not detected: the anti-PRPF31$_{484-497}$ immunoreactivity corresponded to a nuclear/nucleolar staining due to 6xHis-PRPF31 (red signal in A and D; due to the image processing, any cytoplasmic staining is not visible on the picture). Also, anti-V5 immunoreactivity revealed only V5-interactants: as in panel B, V5-MAGI3 signal (in green) was detected at the junctions, and in the nucleus (nucleoli-like structures) along with cytoplasm; panel E exhibits the pattern of localisation previously observed for V5-TRAK2 (in green), which was weakly nuclear with a strong cytoplasmic signal (a faint signal from nuclear aggregates was also observed). The staining of the double-transfected cells with a combination of the secondary antibodies was insignificant, as shown by panels C and F.

The results described above support the validity of the data presented in sections 5.2.3 and 5.2.4. Moreover, panels B and E are in accordance with the description provided in section 5.2.1: 6xHis-PRPF31 was not detected by the anti-V5 antibody (Figure 5.2, panel C; section 5.2.1), nor it was detected this instance.

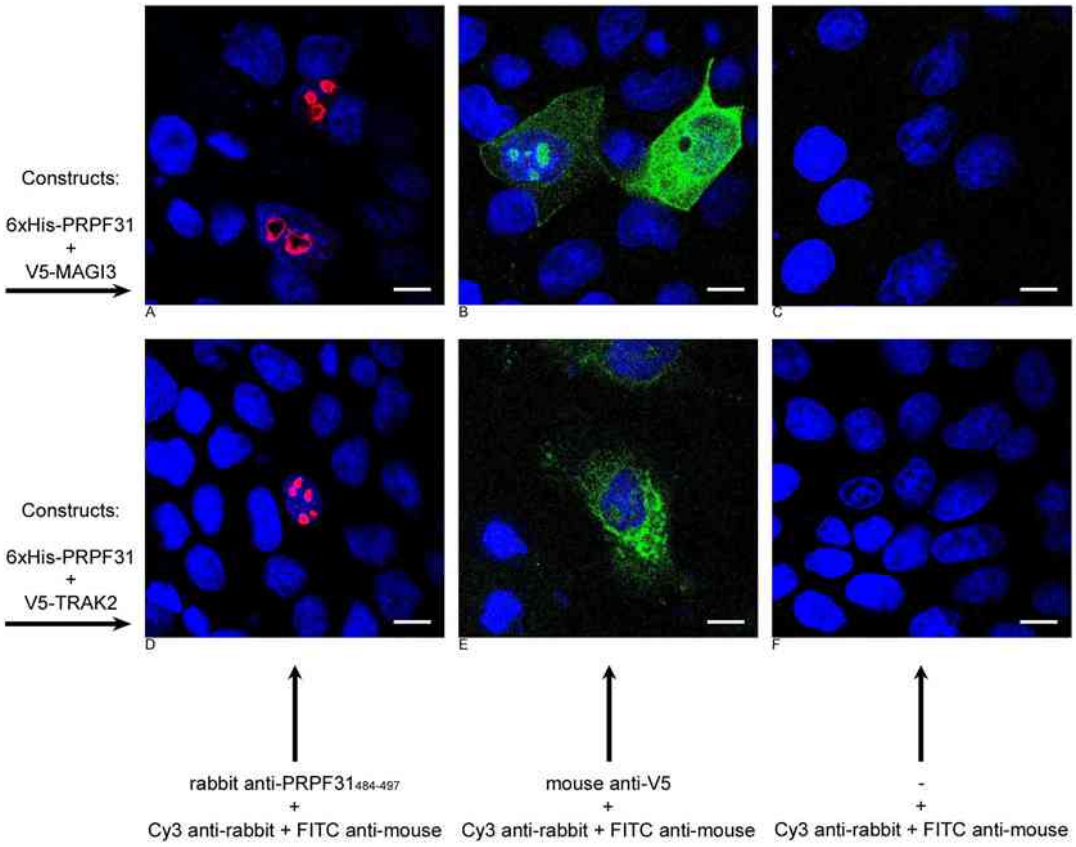


Figure 5.12. Control experiments on MDCK cells expressing 6xHis-PRPF31 and V5-interactant.
A, B and C: cells transfected with 6xHis-PRPF31 and V5-MAGI3; D, E and F: cells transfected with 6xHis-PRPF31 and V5-TRAK2. The anti-PRPF31$_{484-497}$ antibody did not reveal any fluorescence pattern resembling the V5 recombinant protein (panels A and D, 6xHis-PRPF31 in red). The anti-V5 antibody did not reveal any fluorescence pattern belonging to 6xHis-PRPF31 (panels B and E, V5-tagged proteins in green). No signal was detected upon incubation with the two secondary antibodies only (panels C and F) Nuclei were stained with Hoechst 33258 (blue signal). Scale bars, 10µm.

5.4 Localisation of Wild Type PRPF31 and Wild Type MAGI3 in MDCK Cells

As demonstrated by the yeast two-hybrid, immunoprecipitation and localisation experiments, mounting evidence suggests an association between the two interactants, MAGI3 and TRAK2, with PRPF31. However, it is important to validate the co-localisation of the endogenous proteins. The techniques adopted so far have provided promising perspectives; however the use of over-expressed proteins has made them considerably artificial, as the amount of protein expressed may not correlate closely enough to the *in vivo* environment. For specialised cells such as photoreceptor cells it is not practical to work *in vivo*, hence the use of models is preferred which simulate basic *in vivo* conditions. Furthermore, the results shown thus far suggest a thorough investigation of the role of PRPF31 should be undertaken prior to speculating the role of PRPF31 and interacting partners in the retina. That is why we continued work on MDCK cells to test whether or not the endogenous PRPF31 and MAGI3 proteins co-localise (for TRAK2, see section 5.5).

5.4.1 PRPF31 and MAGI3 Overlap at Cell-cell Junctions

Wild type PRPF31 and MAGI3 were detected with the anti-PRPF31$_{484\text{-}497}$ and the anti-MAGI3 antibodies respectively.

As shown by panel A1 in Figure 5.13, anti-PRPF31$_{484\text{-}497}$ immunoreactivity was predominantly nuclear, in speckles. Cytoplasmic staining for PRPF31 was negligible; however, a distinct staining at cell-cell junctions was observed (arrows on panel A1). This staining pattern appeared to intensify depending on cell density: when cells were sparsely populated, the staining was not as intense as when the cells became more confluent. In the latter case however, the signal strength was exceeded by that of the nuclear counterpart, although one could still detect it at a different focus plane. This signal appeared to define the shape of the cell as if the protein was localised along intercellular junctions; the staining appeared to be very grainy and in speckles. Compared to the nuclear fluorescence, the junctional signal was very faint, as the

majority of PRPF31 appeared to localise to the nucleus. The exposure was adjusted to compensate for the low junctional staining in order to detect it when recording images.

MAGI3 (panel A2, Figure 5.13) was observed mainly at intercellular junctions (possibly tight junctions; Wu *et al.*, 2000; Adamsky *et al.*, 2003), where it appeared as an intermittent signal; it also was localised to the cytoplasm, in a perinuclear fashion, and very weakly to the nucleus, reminiscent of the localisation pattern of V5-MAGI3. Generally, anti-MAGI3 immunoreactivity was fainter than that of PRPF31.

In epithelial cells, the PRPF31 and MAGI3 signals showed no co-localisation. However, a partial overlap was seen principally at cell-cell contacts (in panel A4, as indicated by white arrow). Furthermore, the co-staining of PRPF31 and MAGI3 revealed that PRPF31 does not localise to the tight junctions, as no yellow signal was observed (indicating co-localisation of the two proteins). Still, the partial overlap indicates that the proteins may be recruited to a common site and possibly interact.

Wild type MDCK cells were also subjected to a mock primary antibody incubation, followed by incubation with secondary antibodies. As shown in Figure 5.13, panels B1 and B2 illustrate an area of the specimen, laser-excited to detect any red and green fluorescence respectively, for donkey anti-rabbit Cy3-conjugated and donkey anti-mouse FITC-conjugated antibodies. No significant signal was observed.

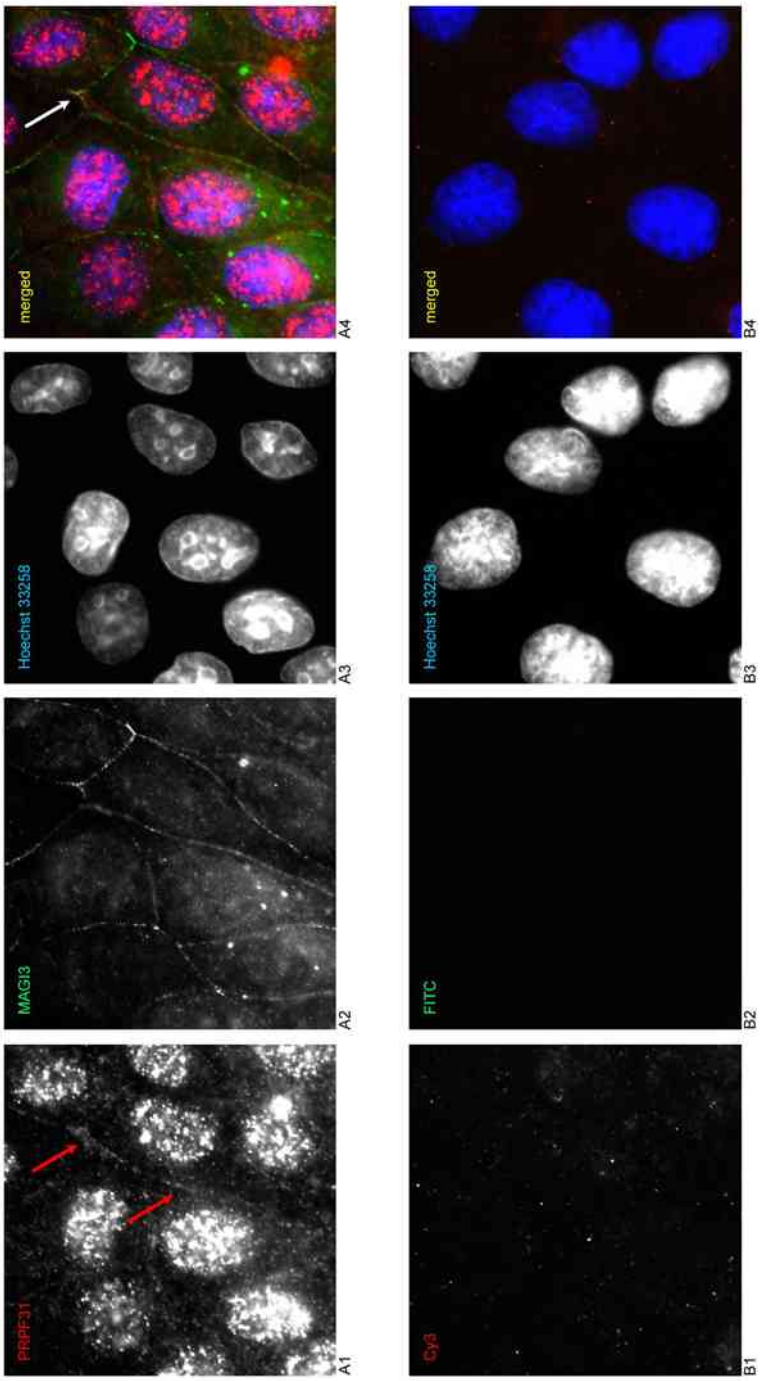


Figure 5.13. Localisation of PRPF31 and MAGI3 in MDCK cells. The above panels show the immunofluorescent staining of MDCK cells with the anti-PRPF31$_{484-497}$ (panel A1; Cy3) and the anti-MAGI3 (panel A2; FITC) antibodies. The localisation of wild type PRPF31 is predominantly nuclear, but also defines the cell shape at the junctions. Wild type MAGI3 is mostly detected at the tight junctions, with a weak perinuclear and nuclear presence.
The B panels illustrate wild type MDCK cells incubated with the secondary antibodies used for the detection of primary antibodies for PRPF31 (B1) and MAGI3 (B2). No significant fluorescence was observed. B1 is for Cy3, B2 is for FITC, B3 is for nuclear stain and B4 is merged. Colour is indicated by legend on each black and white panel.

5.4.2 Detection of PRPF31 with a Second anti-PRPF31 Antibody in MDCK Cells

The junctional localisation observed for PRPF31, described in the previous section 5.4.1, was unexpected. PRPF31 is a well-known splicing factor and therefore presumed to localise specifically to the nucleus. To verify that the junctional localisation of PRPF31 was consistent and not caused by non-specific binding of the anti-PRPF31$_{484-497}$ antibody, the splicing factor was detected with a second antibody. For this purpose, the commercial anti-PRPF31$_{140-154}$ antibody from Abcam was used.

Anti-PRPF31$_{140-154}$ immunoreactivity was detected mostly in the nucleus (Figure 5.14); a little fluorescence was observed in the cytoplasm, conferring a reddish colour to the cells. A junctional localisation pattern was observed (indicated by red arrows in panels A1 and B1 in Figure 5.14), therefore implying that PRPF31 localises at intercellular junctions in epithelial cells. Also this antibody did not appear to detect PRPF31 in the nucleoli, as the comparison with the chromatin staining reveals (compare with panels A2 and B2, Figure 5.14).

Figure 5.14 illustrates also wild type MDCK cells fixed with methanol which underwent a mock primary antibody incubation (panels C), followed by incubation with donkey anti-goat Cy3-conjugated antibody, used to detect PRPF31 with the anti-PRPF31$_{140-154}$ antibody. A faint cytoplasmic fluorescence was observed (panel C1), which did not appear to interfere with the detection of the signal from the primary antibody.

5.4.3 Western Blot Analysis of anti-PRPF31$_{484-497}$, anti-PRPF31$_{140-154}$, anti-MAGI3 Antibodies on MDCK Cell Extracts

MDCK cell extracts were electrophoresed, blotted on nitrocellulose, and the membranes were probed using the antibodies to detect endogenous PRPF31 and MAGI3. Figure 5.15 shows the blots for the two antibodies used to detect PRPF31 (blots A and B for anti-PRPF31$_{484-497}$ and anti-PRPF31$_{140-154}$ respectively) and the antibody for MAGI3

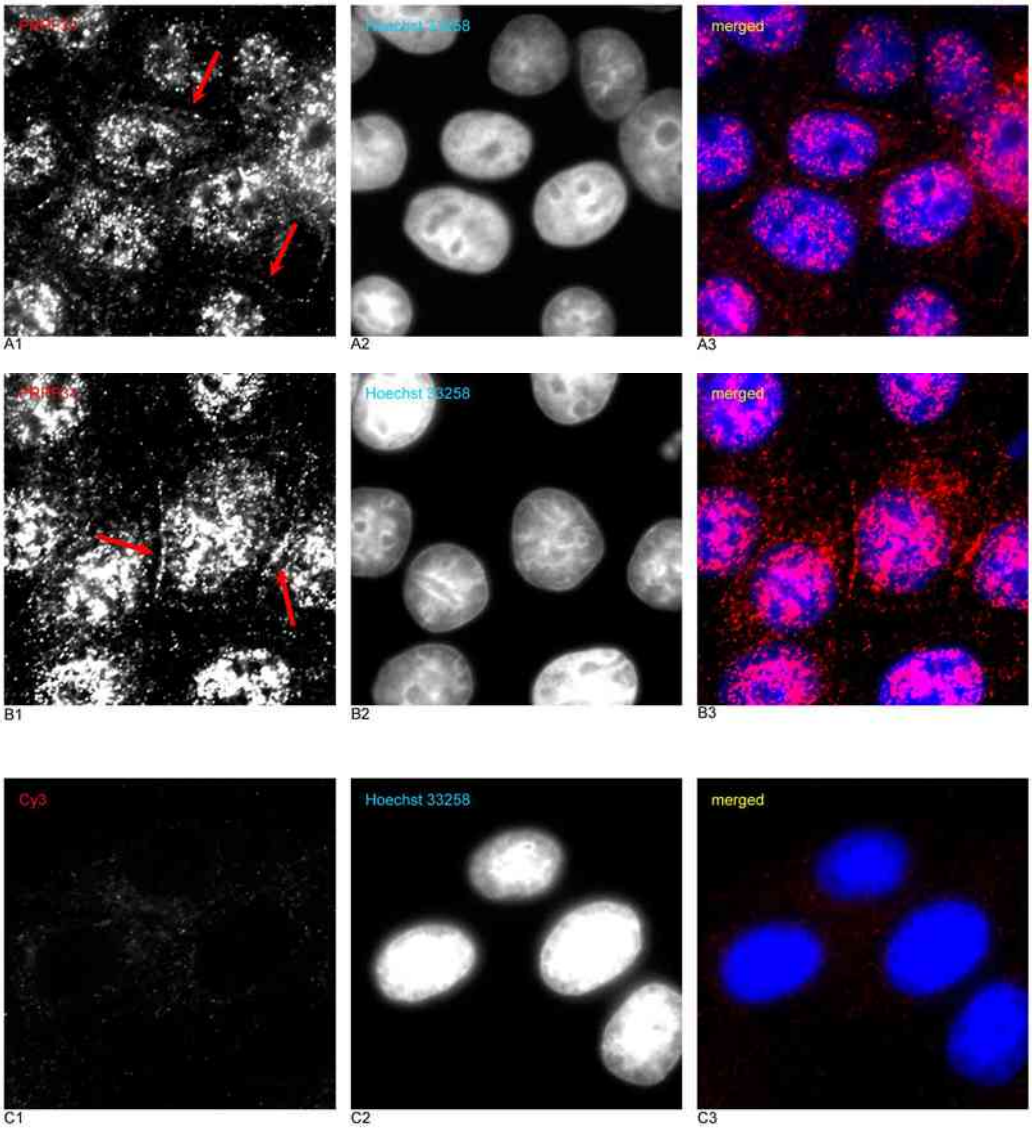


Figure 5.14. Nuclear and junctional localisation of PRPF31 in wild type MDCK cells, using the anti-PRPF31$_{140\text{-}154}$ antibody.
Panels A and B illustrate two different areas of cells incubated with the anti-PRPF31$_{140\text{-}154}$ antibody (Abcam; in Cy3): the antibody revealed a similar localisation pattern as the anti-PRPF31$_{484\text{-}497}$ antibody. Red arrows indicate the intercellular junctions.
The panels C illustrate epithelial cells incubated with the donkey anti-goat secondary antibody used for the detection of anti-PRPF31 antibody$_{140\text{-}154}$. No significant fluorescence was observed.
Colour is indicated by legend on each black and white panel.

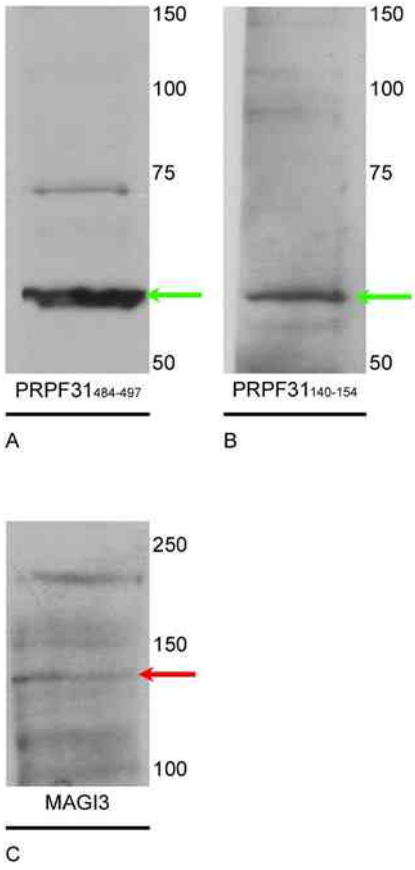


Figure 5.15. Immunoblots showing expression of PRPF31 and MAGI3 in MDCK cells.
In blots A and B, PRPF31 is detected using two different antibodies: the anti-PRPF31$_{484-497}$ antibody detects a major band at ~61 kDa (pointed by the green arrow in A), and a minor band at ~75 kDa. The anti-PRPF31$_{140-154}$ antibody detects only the ~61 kDa (green arrow in blot B). Blot C shows MAGI3 detected by anti-MAGI3 antibody, showing ~140 kDa (red arrow); a second band at ~170 kDa may correspond to an isoform (indicated by upper red arrow). The positions of molecular mass standards (kDa) are shown on the right of each blot.

(blot C). Anti-PRPF31$_{484-497}$ immunoreactivity shows two major bands: one at ~61 kDa, which corresponds to the splicing factor, and another one at ~75 kDa, termed as non-specific (Makarova *et al.*, 2002). The anti-PRPF31$_{140-154}$ antibody from Abcam revealed a band at ~61 kDa, likely corresponding to PRPF31. The anti-MAGI3 antibody displayed a major band at ~140 kDa, likely corresponding to MAGI3. Each relevant band is indicated by an arrow.

The blots from MDCK cells showed that the anti-PRPF31$_{484-497}$ antibody displays two bands. As Makarova and co-workers (2002) reported a non-specific band at ~75 kDa, a correlation between this band and the junctional staining may be hypothesised. However, the second antibody for PRPF31 (anti-PRPF31$_{140-154}$, Abcam) detects only the expected band for PRPF31 (at ~65 kDa), and was also able to detect the junctional staining, hence supporting the validity of the immunofluorescent signals observed for PRPF31 described in the previous section.

5.5 Localisation of Wild Type PRPF31 and Wild Type TRAK2 in SK-N-SH Cells

TRAK2 has been described as enriched in neuronal tissues (Brickley *et al.*, 2005), hence MDCK cells were not suitable for its localisation, since they are a kidney epithelial cell line. For this reason, SK-N-SH cell line (a human neuroblastoma cell line) was used for this purpose. Cells were fixed with methanol, followed by co-staining with the anti-PRPF31$_{484-497}$ and anti-TRAK2 antibodies. Figure 5.16 provides an example of the observed localisation. PRPF31 was observed in nucleus and cytosol: in the nucleoplasm, the signal was detected in small speckles and approximately one third to half of the cells exhibited a weaker signal than MDCK cells; some cells also showed a uniform signal throughout nucleus and cytoplasm, forming a *continuum* (red arrow on panel A1 of Figure 5.16). The cytoplasmic labelling was more pronounced than MDCK cells, and the protein was observed to localise around the nucleus. TRAK2 showed a localisation pattern similar to the over-expressed protein:

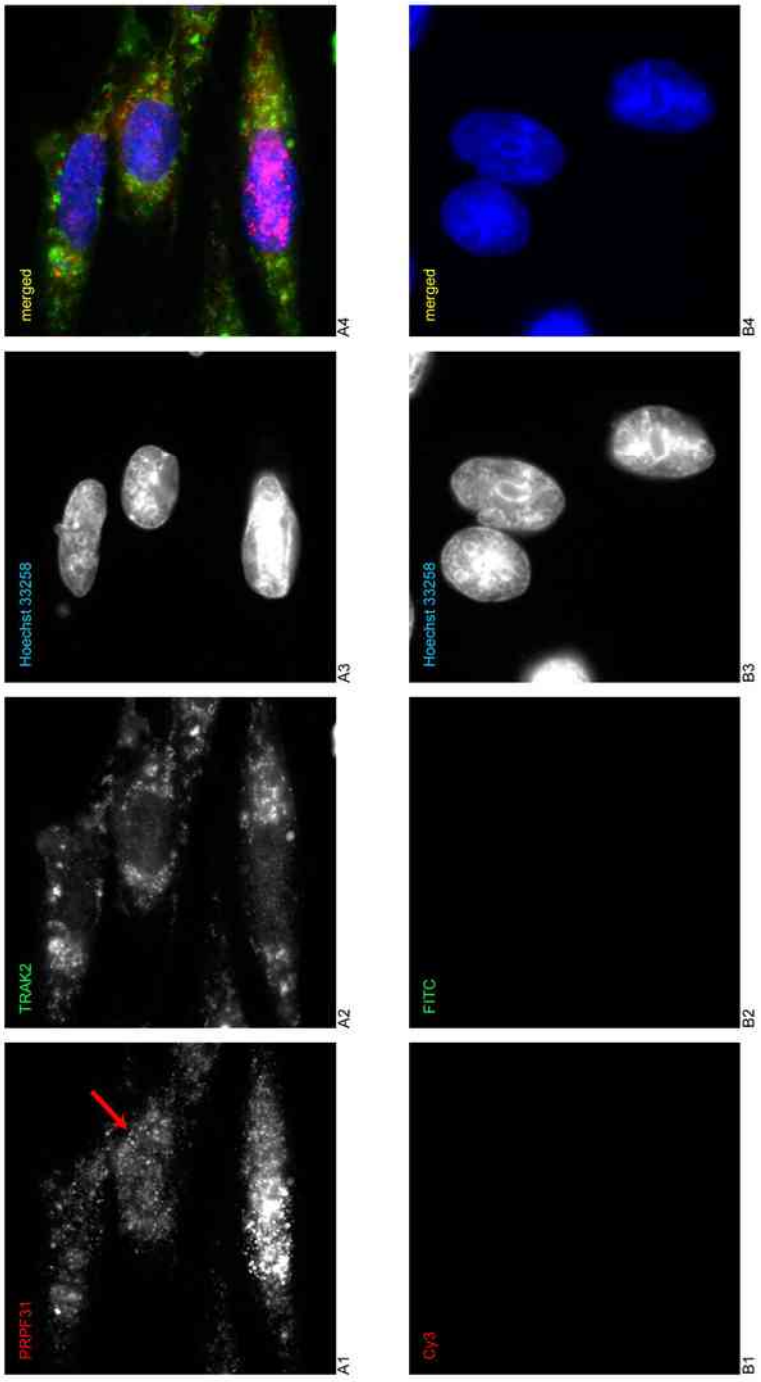


Figure 5.16. Localisation of PRPF31 and TRAK2 in SK-N-SH cells. The panels show immunofluorescent staining of neuroblastoma cells using the anti-PRPF31$_{484-497}$ and the anti-TRAK2 antibodies. Panel A1 shows the localisation of wild type PRPF31 (Cy3), which is both nuclear and cytoplasmic; the red arrow indicates the *continuum* of cytoplasm and nucleus. Panel A2 illustrates wild type TRAK2 (FITC), mostly detected in a perinucler fashion.
The staining of SK-N-SH cells with secondary antibodies only is illustrated in the lower panels. Panels B1 and B2 show the fluorescence from Cy3 and FITC chromophores respectively, without significant signals. B3 is for the nuclear stain. B4 illustrates the merge signal. Colour is indicated by legend on each black and white panel.

it was detected predominantly in the cytoplasm, in a perinuclear fashion with a mottled signal. The spike-like structures on cellular extensions witnessed when MDCK cells expressed V5-TRAK2 were not observed. When the signals from the two proteins were merged together, no clear-cut co-localisation (indicated by yellow) was observed.

SK-N-SH cells which contacted other cells showed localisation for PRPF31 which resembled the one observed at the junctions in MDCK cells. However the reproducibility in print was not representative due to the weakness of the signal, nonetheless one example is illustrated for completeness (Figure 5.17).

Methanol-fixed SK-N-SH cells were also subjected to a staining with secondary antibodies only, to ensure no interfering signals were present. Cells were incubated with donkey anti-rabbit Cy3-conjugated and donkey anti-goat FITC-conjugated antibodies. No significant labelling was present, as shown in panels B in Figure 5.16.

SK-N-SH crude lysates were electrophoresed and blotted on nitrocellulose to assess the anti-TRAK2 antibody; also the anti-PRPF31$_{484\text{-}497}$ antibody was tested. Figure 5.18 illustrates the immunoblot: an anti-TRAK2 immunoreactive band was detectable at ~100 kDa, likely corresponding to TRAK2 (indicated by light blue arrow on panel A); PRPF31 is represented by the band at ~61 kDa (indicated by red arrow in panel B).

5.6 Conclusion

The data regarding the interactions observed by yeast two-hybrid experiments (Chapter 3) and co-immunoprecipitation experiments (Chapter 4) is now supported by the data generated from localisation in cells.

V5-MAGI3 localisation and distribution was in accordance with what has been reported in previous publications (Wu *et al.*, 2000; Adamsky *et al.*, 2003), *i.e.* nuclear but mainly cytoplasmic, in particular surrounding the nucleus and at the tight junctions. V5-TRAK2 was present mostly in the cytoplasm, however a little nuclear staining was also detected; interestingly, V5-TRAK2 was strongly present in punctuate structures at cellular extensions. Most significantly, over-expressed 6xHis-PRPF31 appeared to be strongly cytoplasmic as well as nuclear, forming large nuclear aggregates which resembled the nucleoli.

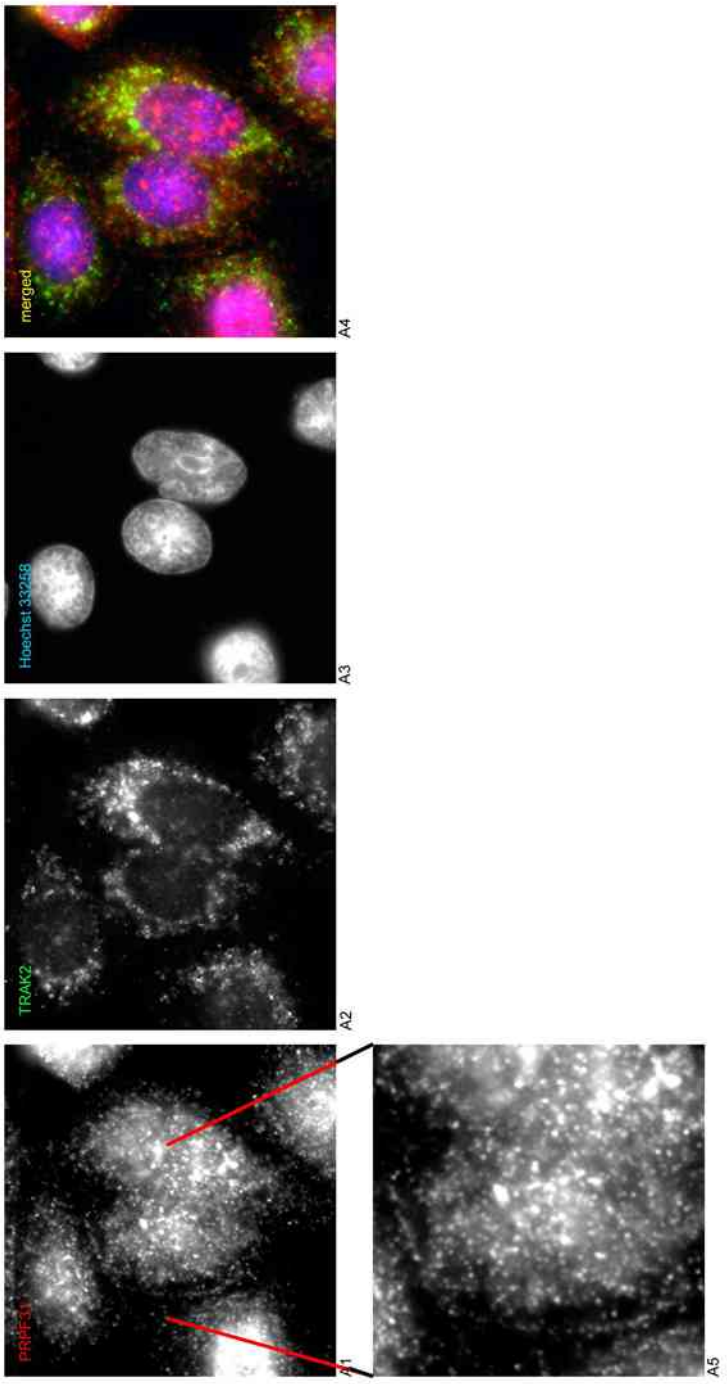


Figure 5.17. Magnification of PRPF31 localisation in SK-N-SH cells. Panels A1 and A2 illustrate cells incubated with antibodies to detect PRPF31 (Cy3) and TRAK2 (FITC) respectively. A3 shows the nuclear staining (Hoechst 33258). The signals are merged in panel A4. Panel A5 offers a magnification of an area in which PRPF31 signal appeared to be towards the border of the cell. Colour is indicated by legend on each black and white panel.

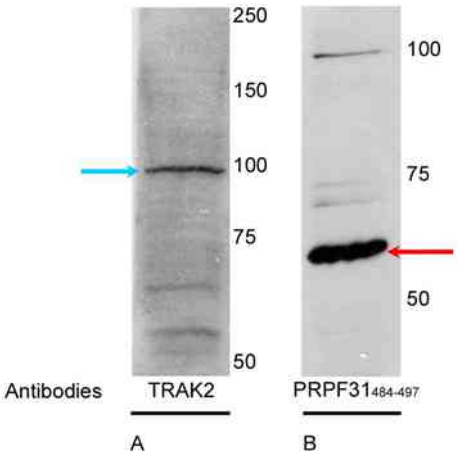


Figure 5.18. Expression of TRAK2 and PRPF31 in SK-N-SH cells.
The anti-TRAK2 antibody was used to probe a crude lysate of neuroblastoma SK-N-SH cells, and it detected a band at ~100 kDa likely corresponding to TRAK2, indicated by the blue arrow on panel A. Panel B shows the immunoblot using the anti-PRPF31$_{484-497}$ antibody (PRPF31 is pointed out by the red arrow). The positions of molecular mass standards (kDa) are shown on the right of each blot.

The co-expression of 6xHis-PRPF31 and V5-MAGI3 resulted in the concentration of PRPF31 in nuclear aggregates, and the protein co-localised with MAGI3 within these aggregates. No other sites of co-localisation were detected. The same pattern of "nucleolar" co-localisation was also observed with V5-TRAK2: the two proteins co-localised mostly to nuclear aggregates, and also to the cytoplasm, especially to spike-like structures.

The observations concerning the co-localisation of tagged proteins were not misinterpreted because no cross-reaction between the anti-V5 antibody and the 6xHis tag was observed (as opposed to the co-IPs, see Chapter 4). The likelihood of a cross-reaction was ruled out by experiments which required the use of the anti-V5 antibody on 6xHis-PRPF31 expressing cells. Also, secondary antibodies did not exhibit any cross-reaction.

However, experiments implying the overexpression of a protein may give only an approximation of the *in vivo* situation, since the protein is present with an additional tag and at a higher level than the normal levels.

Endogenous PRPF31 localised mainly to the nucleus in MDCK cells, with a faint presence in the cytoplasm; it appeared not to be present in the nucleoli. Wild type MAGI3 displayed the same pattern of localisation as the over-expressed protein: it was detected mostly at the tight junctions and around the nucleus, and weakly inside the nucleus. The overlays showed that the signals from the two proteins did not merge, suggesting no co-localisation; however, the signals overlapped at the junctions, giving a yellow signal in forms of little spots. Interestingly, wild type PRPF31 was found at the junctions (likely not tight junctions as otherwise co-localisation with MAGI3 would have been observed). The intercellular junction staining was also observed when using a different antibody against PRPF31.

Experiments to co-localise wild type PRPF31 and TRAK2 were conducted on SK-N-SH cells: SK-N-SH cell line is derived from neuroblastoma cells, thus supposed to express TRAK2 in an amount adequate for detection (TRAK2 has been described as enriched in neuronal tissues; Brickley *et al.*, 2005). The staining revealed a pattern for TRAK2 which resembled the localisation of the tagged TRAK2 protein, *i.e.* it was mainly cytoplasmic and surrounding the nucleus. No spike-like structures were observed,

although this may be due to the type of cell used. Regarding PRPF31 localisation, many cells did not display an intense fluorescence in the nucleus; conversely they showed smaller speckles than in the MDCK cells. It appeared that many cells (30-50%) did not exhibit a distinction between the nuclear and the cytoplasmic signal: a poor labelling of the nucleoplasm but a distinct labelling of the cytosol around the nucleus.

6. PRPF31 in MDCK Cells and Retina

As PRPF31 appeared to localise to intercellular junctions in MDCK cells, the localisation of the splicing factor in this cell line was pursued further, as this may link to recent findings of splicing occurring outside the nucleus (Denis *et al.*, 2005; König *et al.*, 2007). Since MDCK cells show a stable differentiation, allowing inferring conclusions on observations, we opted to continue with their use in our investigation of the role of PRPF31. As retina is a highly stressed tissued, we wanted to investigate whether stress events elicit any change in the pattern of localisation of PRPF31: for this purpose, cells MDCK were cultured under different conditions, altering the levels of stress.

Furthermore, since the junctional localisation posed an interesting question regarding the localisation of PRPF31, we performed staining of PRPF31 in wild type murine retina sections.

The results reported in this chapter reveal an interesting scene for PRPF31, however are far from elucidating the role of PRPF31 in MDCK cells and in mouse retina. More in-depth studies addressing PRPF31 in epithelial cells and also studies targeting retina have to be pursued.

6.1 PRPF31 Localises to the Adherens Junctions in MDCK Cells

Chapter 5 (section 5.4) presented an unexpected localisation of the PRPF31 splicing factor in MDCK cells. The aforementioned section, dedicated to the concomitant localisation of PRPF31 and MAGI3, suggested that PRPF31 not only localises to the nucleus, but also at intercellular junctions (localisation not previously reported).

Cells were cultured on glass coverslips, fixed with methanol and then stained with the appropriate antibodies (procedures in sections 2.12.4 and 2.14.1 respectively).

Figure 6.1 illustrates these observations: the splicing factor was abundant in nuclear speckles, in a granular and mottled fashion; comparing with the chromatin staining (panel A2), the nucleoli appear to be largely unstained (white arrows in panel A3).

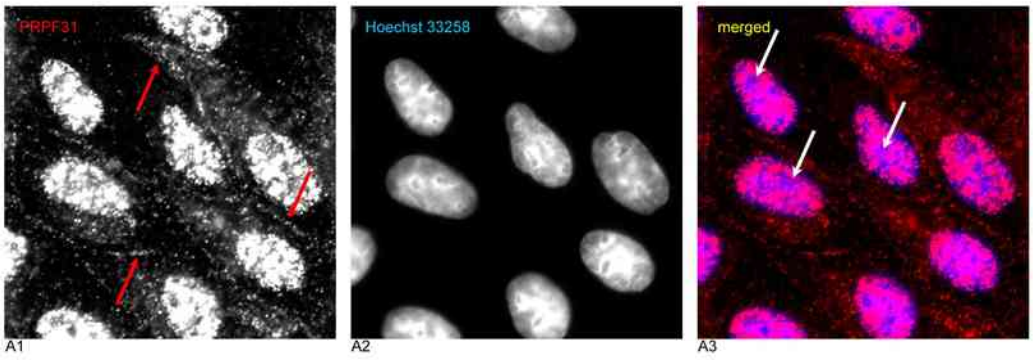


Figure 6.1. PRPF31 localises to the nucleus and at cell-cell junctions in wild type MDCK cells. PRPF31 presence at intercellular junctions is indicated by red arrows in A1. PRPF31 does not appear to localise to the nucleoli, as pointed by white arrows in A3. Nuclei were stained with Hoechst 33258 (blue signal), panel A2. Colour is indicated by legend on each black and white panel.

A little fluorescence was detected in the cytoplasm; however, a more abundant PRPF31 signal was observed at the periphery of the cell, likely corresponding to intercellular junctions (Figure 6.1, as indicated by red arrows in A1).

As already stated, the labelling of the protein was observed at the junctions at certain cell densities. When cells were grown sparsely, the signal from the cell membrane was almost imperceptible. When the cells were grown to confluence, the signal became more prominent (as in Figure 6.1); however, the fluorescence of nuclear PRPF31 was stronger than the junctional, therefore the exposure was adjusted to compensate for the low junctional staining in order to detect it when recording images (hence explaining the brightness of the nuclei). At higher cell density, the signal was detected on a different plane of focus than the nuclear plane, and as its intensity was exceeded by that of the nuclear signal, it was difficult to take representative pictures.

Due to the fact that PRPF31 did not co-localise with MAGI3 (section 5.4.1), it was suggested that the splicing factor is not present in the tight junctions of this type of epithelial cell line, and with the conditions employed. We then hypothesised that PRPF31 localised at the adherens junctions (AJs): to investigate this hypothesis, PRPF31 was localised with E-cadherin, a protein known to localise to the AJs of MDCK cells (Sheibani *et al.*, 2000). As shown in Figure 6.2, PRPF31 and E-cadherin exhibit a similar distribution pattern at the AJs (compare panels A1 with A2) and the merge panel illustrates the resulting yellow signal (indicated by white arrows in panel A4), suggestive of co-localisation.

Cells were also co-stained for PRPF31 and SC-35. SC-35 is a non-snRNP splicing factor of the SR family of proteins, and it is known to be part of the splicing complex and hence it can be used as a marker for splicing nuclear speckles (as in Comitato *et al.*, 2007). In panel A4 of Figure 6.3, the yellow signal in the nucleus illustrates that the mottled pattern shown by PRPF31 (Cy3) merges with the dotted presence of SC-35 (FITC). However, SC-35 appears to be less abundant, having fewer speckles, whereas PRPF31 is observed throughout the nucleoplasm, likely representing PRPF31 associated with snRNPs, but not with the spliceosomes. PRPF31 also appears to be absent from the nucleoli, as SC-35 protein is not present in these structures, supporting the results reported by Makarova and co-workers (2002). However, it would be desirable to

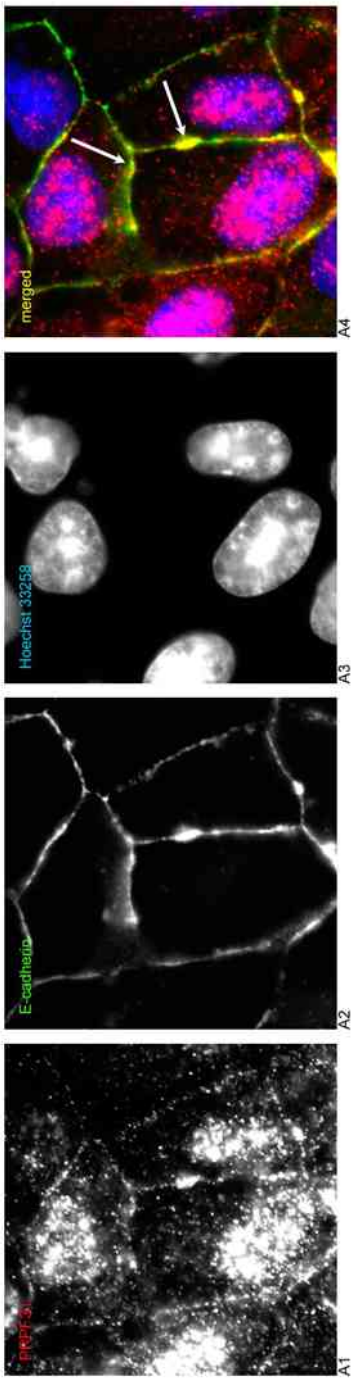


Figure 6.2. PRPF31 localises to adherens junctions with E-cadherin in wild type MDCK cells. Panel A1 illustrates PRPF31 (Cy3); panel A2 illustrates SC-35 (FITC); Hoechst stain (blue) is illustrated in panel A3. Panel A4 represents the overlay of the three signals. Colour is indicated by legend on each black and white panel.

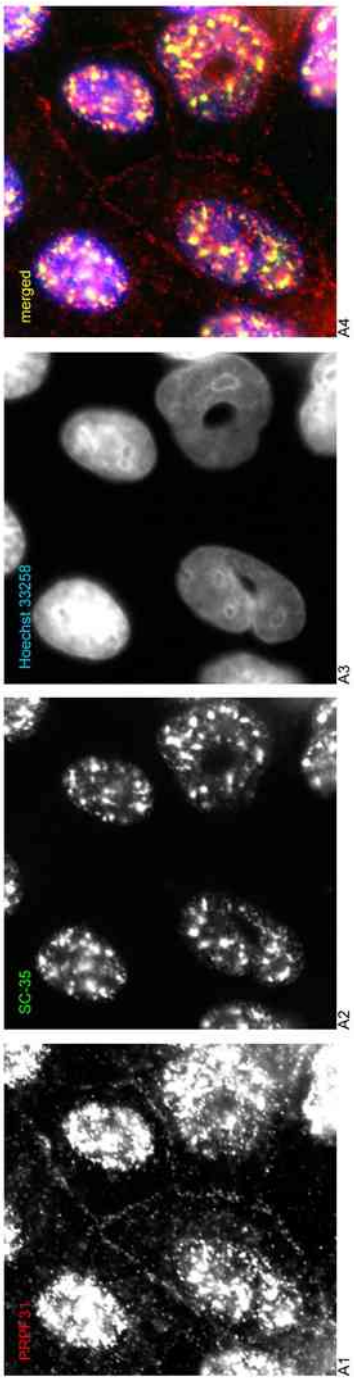


Figure 6.3. PRPF31 localises with SC-35 to nuclear speckles in wild type MDCK cells. PRPF31 is shown in panel A1 (Cy3); SC-35 is shown in panel A2 (FITC); Hoechst stain (blue) is shown in panel A3. Panel A4 is the merge of the three signals. Colour is indicated by legend on each black and white panel.

perform a co-labelling with a nucleolar marker (fibrillarin for example). In Figure 6.3 the adherens junction staining of PRPF31 is also visible.

6.2 PRPF31 Localisation in G0/G1 Arrested Cells

To investigate PRPF31 localisation in epithelial cells according to the normal proliferative state of mammalian cells, MDCK cells were arrested at the G0/G1 phase. To specifically target cell proliferation, cells were treated with CDK inhibitors (protocol described in section 2.12.5) to arrest the cell cycle.

Two inhibitors were used: olomoucine, which is specific for CDK2, and "CDK4 inhibitor", which is specific for CDK4. After 24 hours of treatment, cells were fixed with methanol, stained for PRPF31 using the anti-PRPF31$_{484\text{-}497}$ antibody (for procedures, see sections 2.12.4 and 2.14.1) and examined at the microscope. Results are shown in Figure 6.4.

Olomoucine-treated cells grew in islands; they presented round nuclei, similar to the the untreated cells (compare panel A2 with C2); PRPF31 was detected at the junctions, although its signal was more attenuated and weaker than the wild type (signal indicated by the red arrows in panel A1), nonetheless it still delineated the cell borders. Olomoucine-treated cells appeared to be sparser and more separated than the untreated cells.

Cells treated with CDK4 inhibitor (panels B) were sparser, compared to both the olomoucine-treated cells and the untreated. They also appeared to grow in smaller patches compared to the untreated cells, and the nuclei were not as round as the untreated control. PRPF31 localised both to the nucleus and to the intercellular junctions (red arrow in panel B1), and the junctional signal appeared to be more prominent than the olomoucine-treated cells.

In both cases, the protein did not appear to be present in the nucleoli, according to the chromatin staining with Hoechst (white arrows in panels A3 and B3).

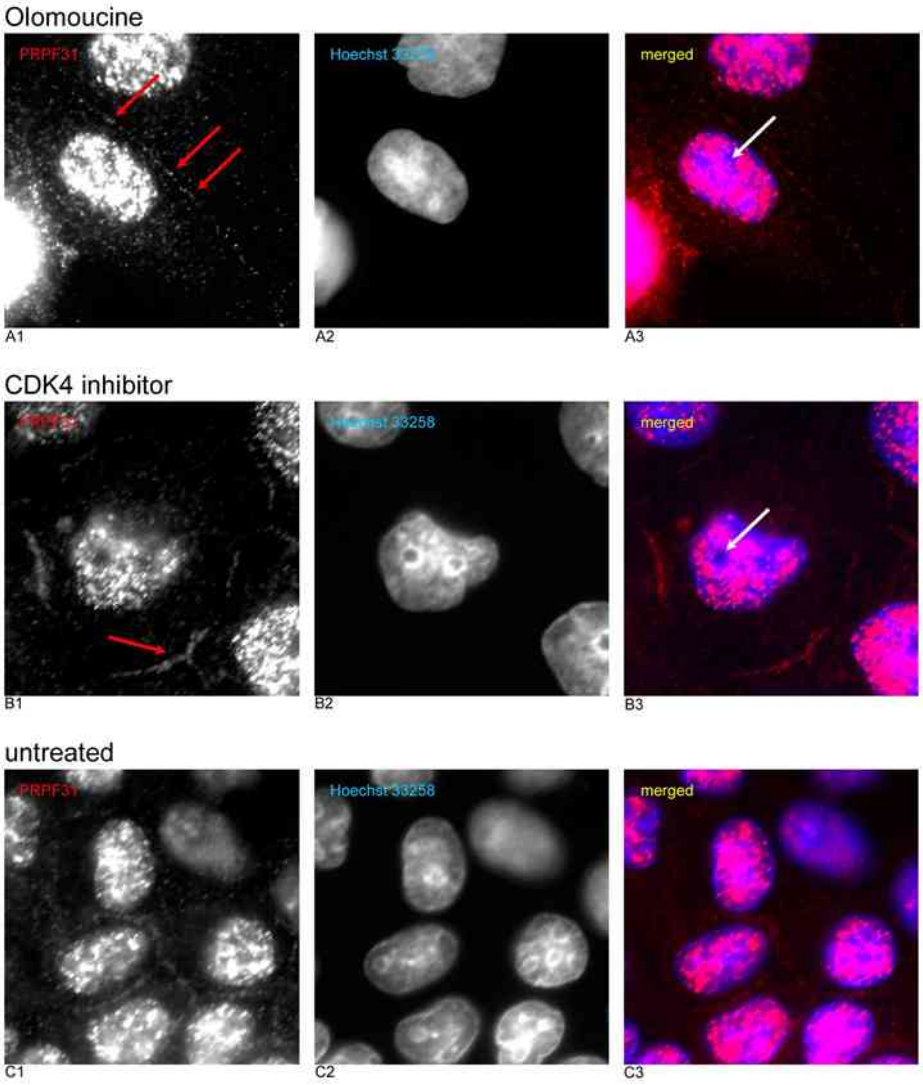


Figure 6.4. PRPF31 localises to the nucleus and to intercellular junctions during G0/G1 phase in MDCK cells.
Panels A refer to Olomoucine-treated cells; panels B to CDK4-inhibitor-treated cells; panels C to the untreated cells. PRPF31 localisation is illustrated in panels A1, B1 and C1 (Cy3); Hoechst 33258 in A2, B2 and C2 (blue signal); A3, B3 and C3 represent the overlays. Red arrows in A1 and B1 indicate the junctional staining for PRPF31 in treated cells; white arrows in the overlays A3 and B3 indicate the nucleoli of the treated cells, according to the Hoechst chromatin stain. Colour is indicated by legend on each black and white panel.

6.3 PRPF31 Localisation under Stress Conditions

To investigate whether the distribution of the splicing factor is affected during stress conditions, PRPF31 localisation in MDCK cells was examined after stress treatments of the cells. Two different methods were applied to insult the cells: a temperature stress, heat-shocking the cells at 43°C for two hours (procedure in section 2.12.6), and an osmotic stress, varying the concentration of salts in the medium (section 2.12.7). Cells were then fixed with methanol and incubated with the anti-PRPF31$_{484\text{-}497}$ antibody (procedures 2.12.4 and 2.14.1 respectively), together with the anti-SC-35, anti-E-cadherin (AJs marker) and anti-MAGI3 (one interacting partner) antibodies. The results are described in the next two sections.

6.3.1 Heat Shock

PRPF31 localisation on cells subjected to heat-shock treatment appeared to be similar to the untreated cells: as shown in Figures 6.5, 6.6 and 6.7, PRPF31 immunoreactivity is detected at the AJs and in the nucleus, predominantly in speckles (panels A1 of the aforementioned figures). However, it appeared that the intensity of the signal from the AJs due to PRPF31 was stronger than the untreated sample (compare panels A1 and B1 of Figures 6.5, 6.6 and 6.7, where panels A1 illustrate the treated sample and panels B1 illustrate the untreated sample).

Antibody against SC-35 was used to localise the spliceosomes, and its signal was detected mainly in nuclear speckles, along with a faint halo throughout the nucleus (panel A2 of Figure 6.5), resembling the untreated cells (panel B2). The simultaneous localisation of PRPF31 and SC-35 revealed co-localisation (yellow signal in A4 and B4) of the two proteins in splicing nuclear speckles, in both treated and untreated cells.

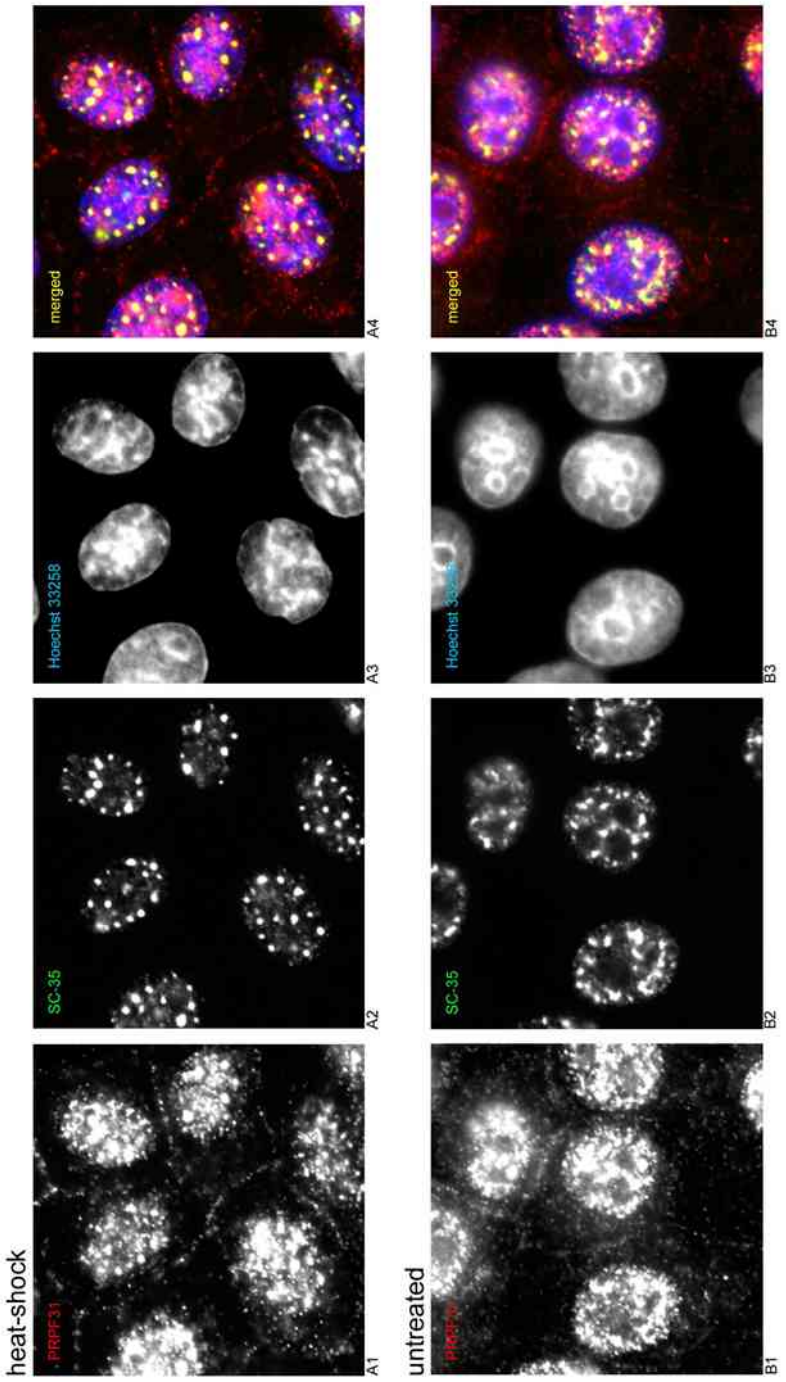


Figure 6.5. PRPF31 localises with SC-35 to nuclear speckles in heat-shocked MDCK cells. Panels A represent heat-shocked cells, and panels B the untreated cells. PRPF31 is shown in panels A1 and B1 (Cy3), and localises to nucleus, and intercellular junctions; SC-35 is shown in panels A2 and B2 (FITC); panels A4 and B4 are the merge of the three signals. Nuclei were stained with Hoechst 33258 (blue signal in panels A3 and B3). Colour is indicated by legend on each black and white panel.

As shown on Figure 6.6 (panel A4), the signal from the anti-PRPF31$_{484-497}$ antibody poorly localises with the signal from the anti-E-cadherin antibody in the heat-shocked cells. Nonetheless it appears that the junctional pattern of PRPF31 is similar to the pattern of E-cadherin. It may be that the procedures were inadvertently altered while processing the cells, hence the poorer localisation of PRPF31 to the junctions compared to different samples of treated cells.

A third set of coverslips were incubated with the anti-PRPF31$_{484-497}$ and the anti-MAGI3 antibodies (Figure 6.7): in heat-shocked cells, PRPF31 localised to the nucleus and to the AJs (panel A1), and MAGI3 localised mainly to the TJs (panel A2 of Figure 6.7). The localisation pattern given by the anti-PRPF31$_{484-497}$ antibody appeared towards the inward part of the membrane, whereas the pattern for MAGI3 delineated the outer part of the plasma membrane. The overlay from PRPF31 and MAGI3 signal does not appear yellow, however a partial overlap on the lower left part of the plasma membrane is observed (as indicated by the white arrow in panel A4).

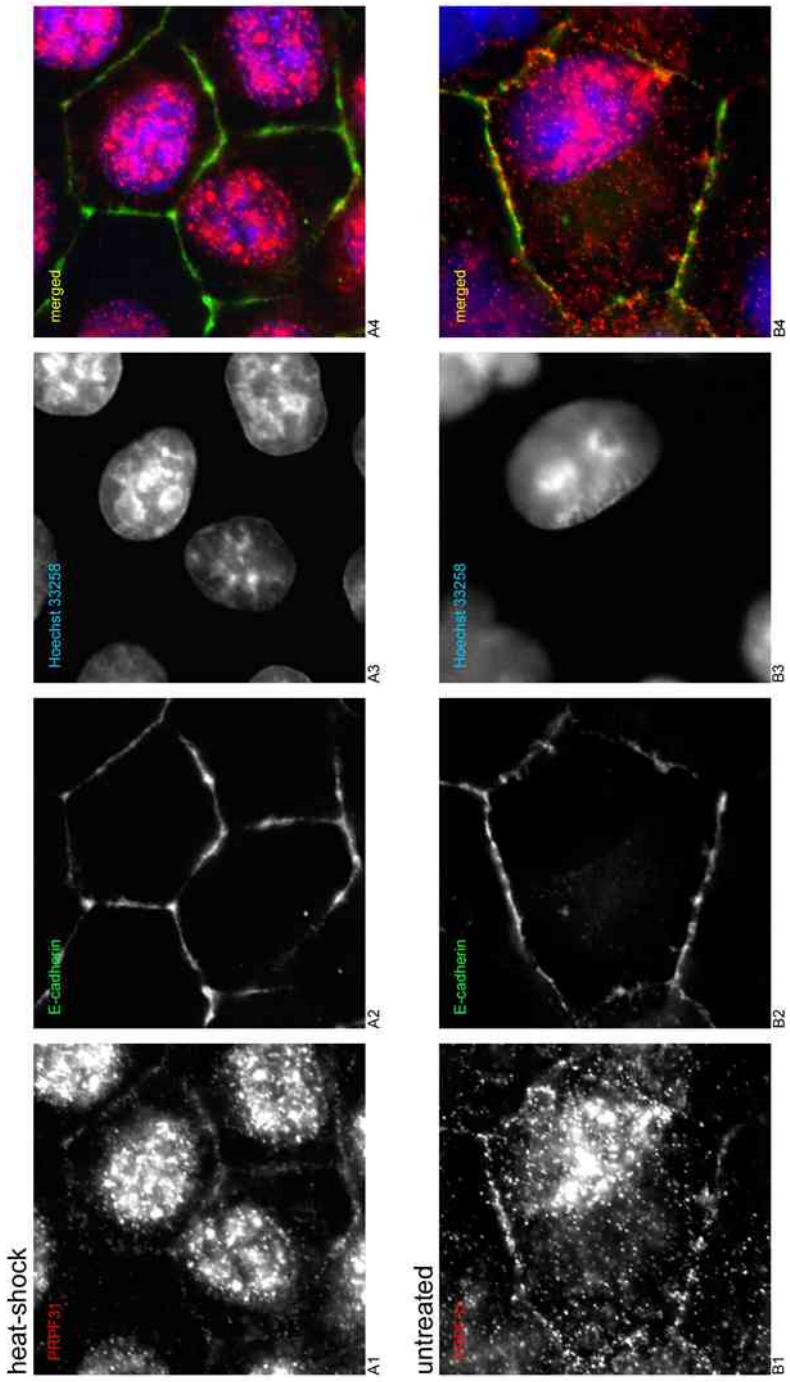


Figure 6.6. PRPF31 co-localises with E-cadherin to adherens junctions heat-shocked MDCK cells. Panels A represent heat-shocked cells, and panels B the untreated sample. PRPF31 is illustrated in panels A1 and B1 (Cy3), and localises to nucleus and adherens junctions. E-cadherin localisation is illustrated in panels A2 and B2 (FITC); panels A4 and B4 illustrate the overlays. Nuclei were stained with Hoechst 33258 (blue signal in panels A3 and B3). Colour is indicated by legend on each black and white panel.

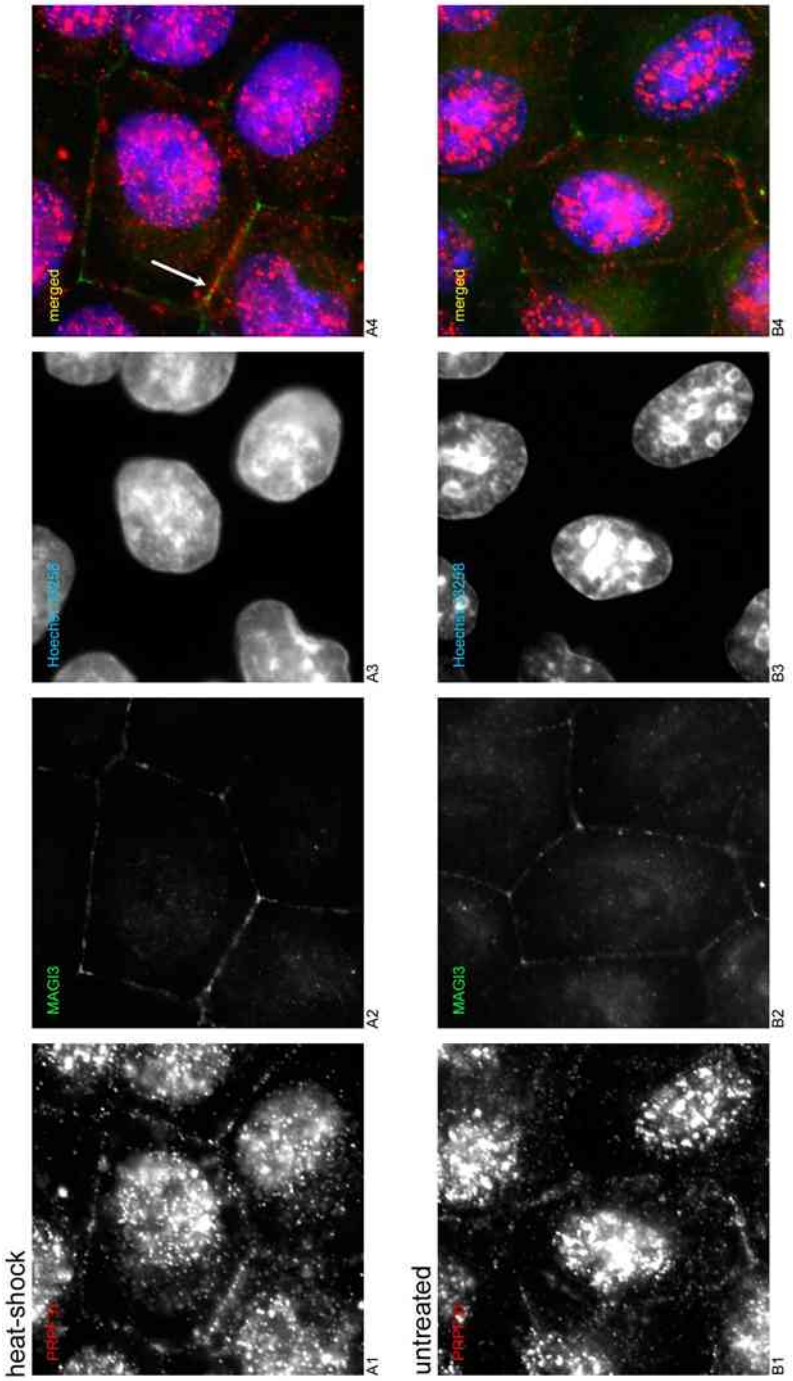


Figure 6.7. PRPF31 and MAGI3 localisation in heat-shocked MDCK cells. Panels A represent heat-shocked cells, and panels B the untreated sample. Panels A1 and B1 represent PRPF31 localisation (Cy3), which appears to localise to nucleus and intercellular junctions. Panels A2 and B2 represent MAGI3 localisation (FITC) to the tight junctions. Panels A4 and B4 illustrate the overlays: the partial overlap is indicated by a white arrow in A4. Nuclei were stained with Hoechst 33258 (blue signal in panels A3 and B3). Colour is indicated by legend on each black and white panel.

6.3.2 Osmotic Stress

Cells were put under osmotic stress by altering the physiological osmolarity (300 mOsm; procedure in section 2.12.7) of the culturing medium to:

- 600 mOsm for a hyperosmotic stress;
- 150 mOsm for a hypoosmotic stress.

Two sets of cells were treated separately with the two media for 2 hours. Cells were then fixed with methanol and stained with the anti-PRPF31$_{484\text{-}497}$ antibody, together with the anti-SC-35 antibody for detection of splicing nuclear speckles, the anti-E-cadherin antibody for the adherens junctions and the anti-MAGI3 antibody for the interacting partner MAGI3 (protocols described in sections 2.12.4 and 2.14.1).

As the heat-shock and osmotic stress experiments were conducted on the same batch of cells and at the same time, the images of the untreated cells reported in this section correspond to the images of the untreated cells of the heat-shock experiment.

Figures 6.8, 6.9 and 6.10 show the cells stained for PRPF31, together with SC-35 (Figure 6.8), E-cadherin (Figure 6.9), and MAGI3 (Figure 6.10). Each figure illustrates cells treated with hyperosmotic- and hypoosmotic-medium and the untreated control.

In cells which underwent treatment with hyperosmotic medium, PRPF31 (all panels A1 of Figures 6.8, 6.9 and 6.10) appeared to be in the nucleus and also at the AJs, as in the untreated cells. PRPF31 presence appeared to be more granular than the untreated cells in both the nucleus and the junctions; SC-35 signal was also observed in larger granules (panel A2, Figure 6.8). In addition, PRPF31 appeared to be present in the nucleoli, a localisation which was not observed in the untreated cells. The nucleolar staining is visible in Figure 6.8, and is indicated by a red arrow in panel A1 and a green arrow in panel A2 (SC-35), and white arrows in the overlay panel A4. Figure 6.9 shows that the anti-E-cadherin fluorescence pattern is shared with the anti-PRPF31$_{484\text{-}497}$ fluorescence pattern: PRPF31 appeared to be prominent at the adherens junctions. Panel A4 of Figure 6.9 shows the merge image, with an intense yellow signal along the adherens junctions, indicative of co-localisation of the two proteins. MAGI3 was present at the TJs (Figure 6.10, panel A2), and in the nucleus; in the latter the fluorescence was diffuse, resembling a halo, and showed increased intensity towards the centre (green arrow, panel B2), likely

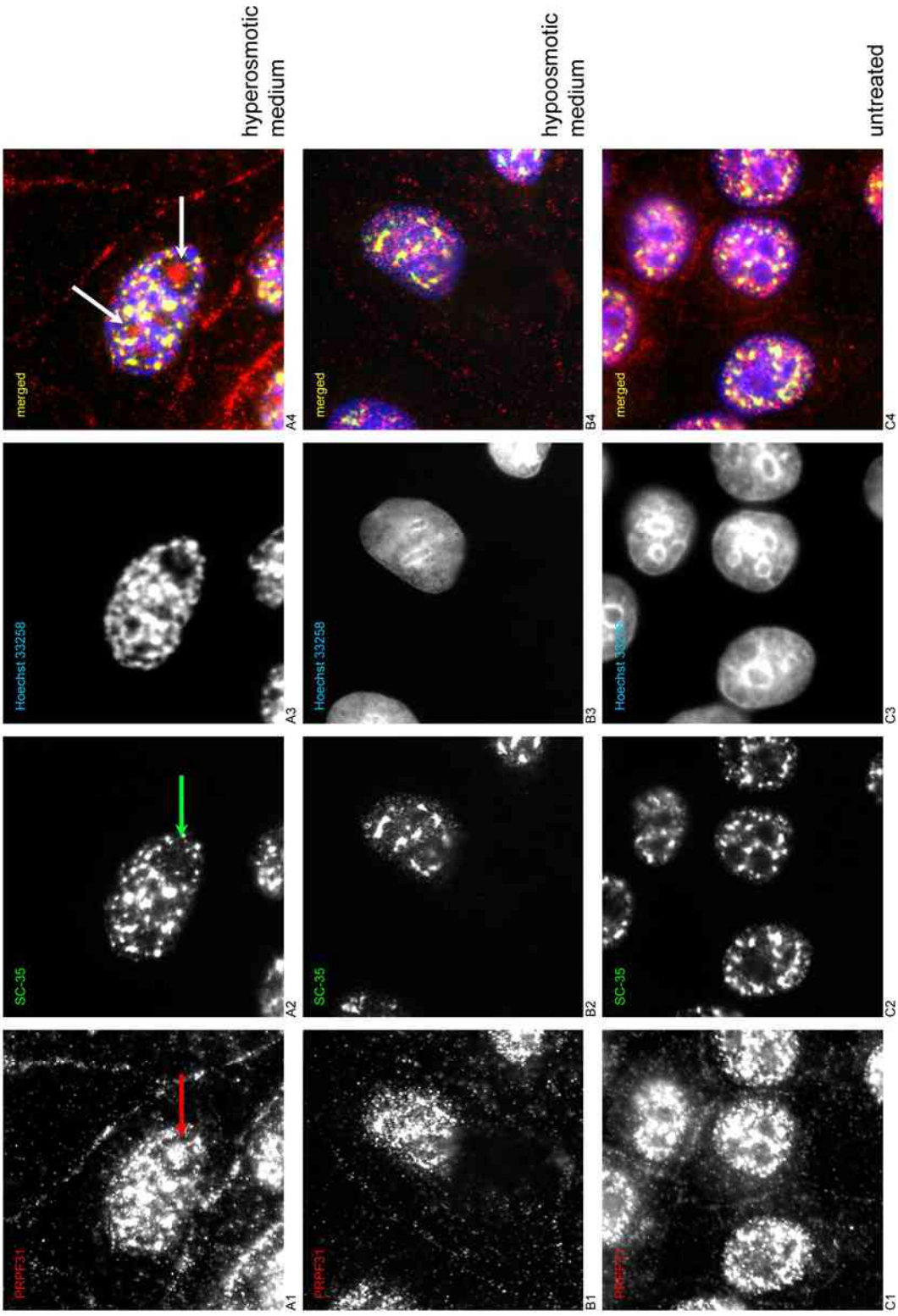


Figure 6.8. Localisation of PRPF31 and SC-35 in MDCK cells maintained in hyperosmotic (panels A), hypoosmotic (panels B) and isoosmotic (panels C) media. A1, B1, C1: PRPF31 (Cy3); A2, B2, C2: SC-35 (FITC); A3, B3, C3: Hoechst 33258; A4, B4, C4: overlays. Red arrow in A1, green arrow in A2, white arrows in A4 indicate nucleoli. Colour is indicated by legend on each black and white panel.

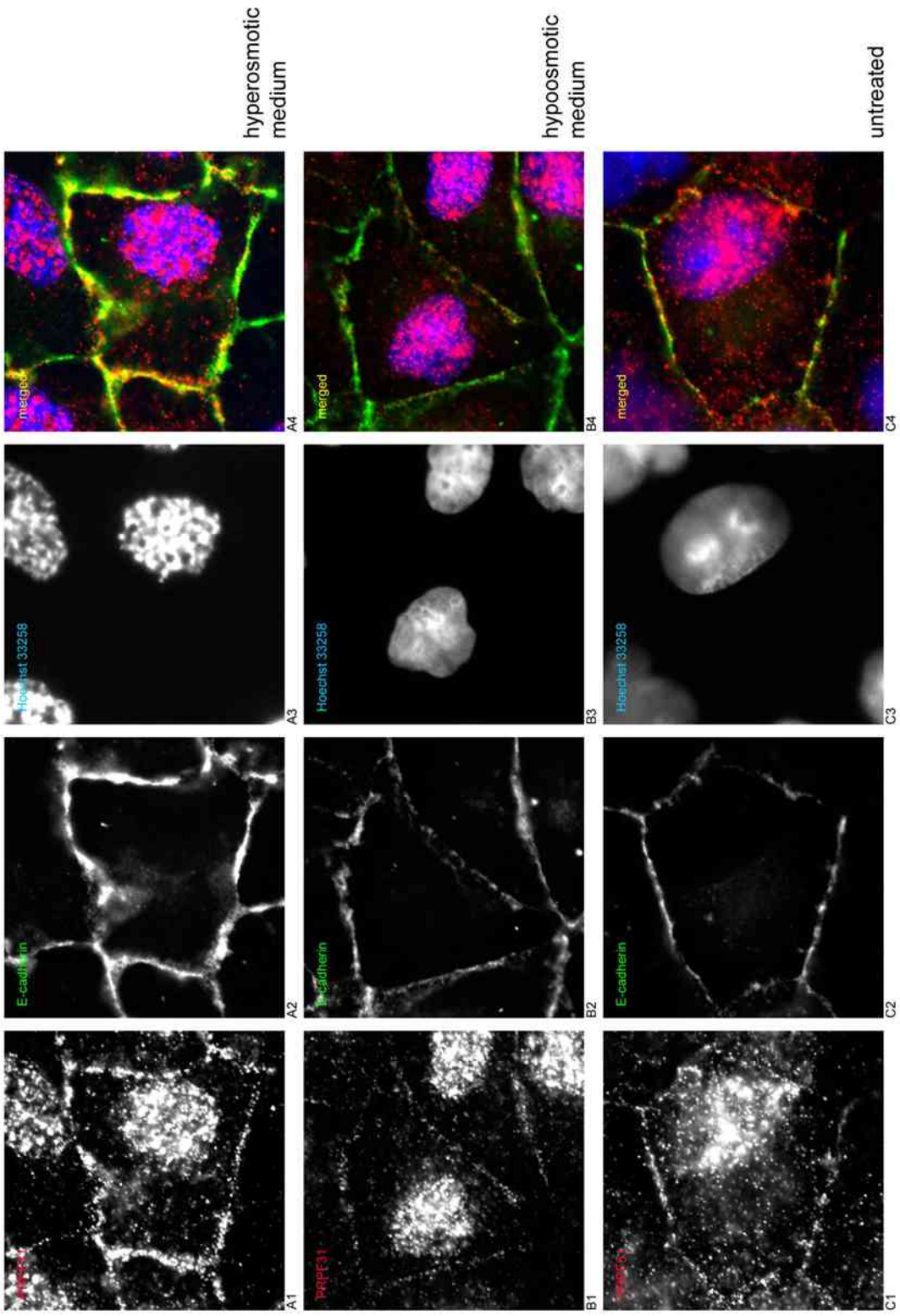


Figure 6.9. Localisation of PRPF31 and E-cadherin in MDCK cells maintained in hyperosmotic (panels A), hypoosmotic (panels B) and isoosmotic (panels C) media. A1, B1, C1: PRPF31 (Cy3); A2, B2, C2: E-cadherin (FITC); A3, B3, C3: Hoechst 33258; A4, B4, C4: overlays. Colour is indicated by legend on each black and white panel.

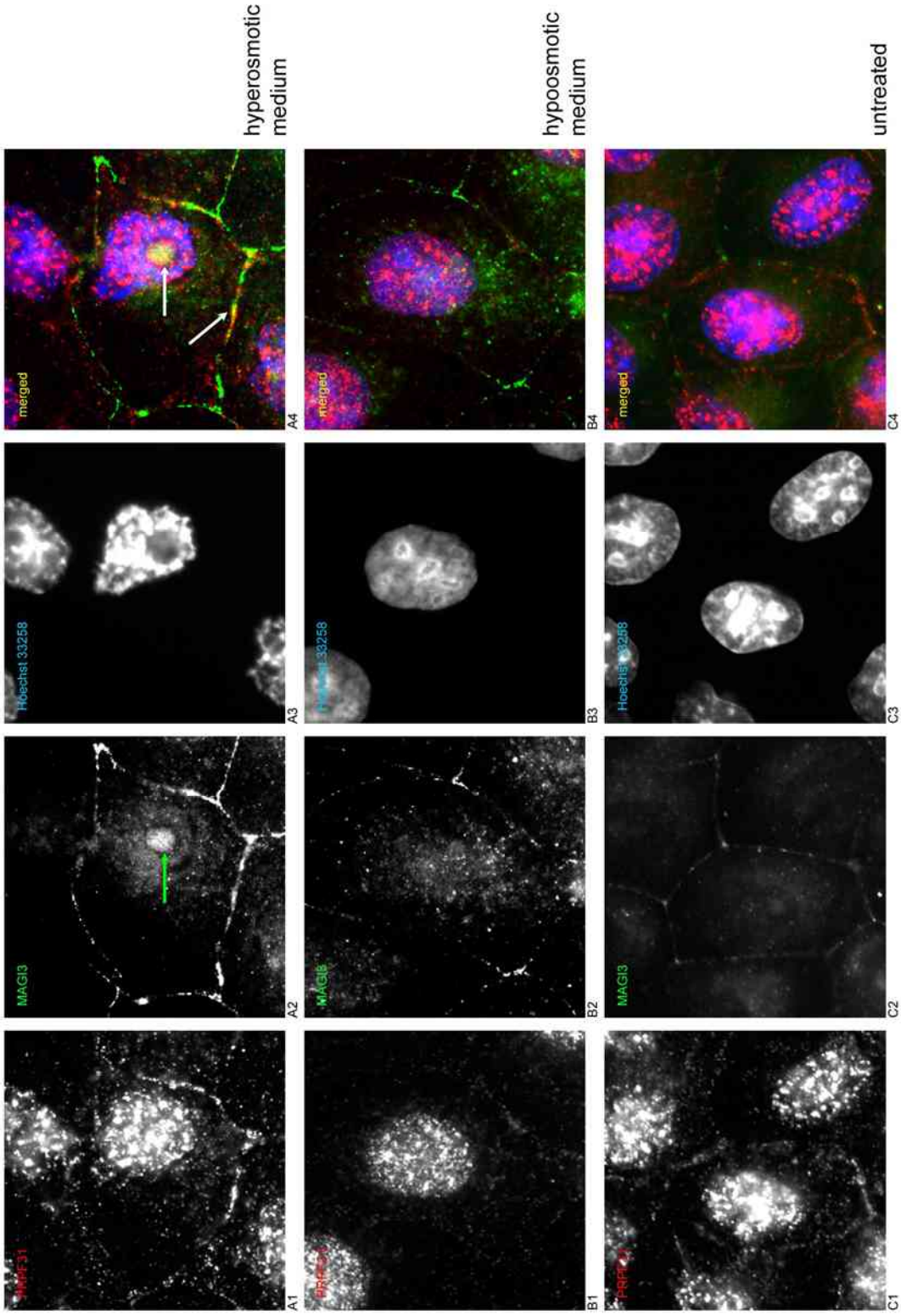


Figure 6.10. Localisation of PRPF31 and MAGI3 in MDCK cells maintained in hyperosmotic (panels A), hypoosmotic (panels B) and isoosmotic (panels C) media. A1, B1, C1: PRPF31 (Cy3); A2, B2, C2: MAGI3 (FITC); A3, B3, C3: Hoechst 33258; A4, B4, C4: overlays. Green arrow in A2 indicates a nucleolus; in A4, upper white arrow indicates a nucleolus, lower white arrow indicates co-localisation of PRPF31 and MAGI3. Colour is indicated by legend on each black and white panel.

a nucleolus, as identified by the Hoechst staining (Figure 6.10, panel A3). The nuclear and nucleolar signals for MAGI3 were more prominent in the cells treated for hyperosmotic stress than for the hypoosmotic stress. Furthermore, the pattern resembled the one observed when V5-tagged MAGI3 was expressed in MDCK cells (compare with Figures 5.6 and 5.7 in section 5.2.5 of Chapter 5). Merging the three signals together – PRPF31, MAGI3 and Hoechst, (red, green and blue respectively) – reveals a co-localisation of PRPF31 and MAGI3 to the nucleolus and partially to the junctions (yellow signal in panel A4, indicated by white arrows).

In cells treated with hypoosmotic-medium, PRPF31 presence was detected mainly in the nucleus, and did not appear to be in the nucleoli as the co-staining for the SC-35 protein revealed (Figure 6.8, panels B1 and B2, to be compared with the nuclear staining B3). Also, the fluorescence appeared to be more evenly distributed in the nucleoplasm, with smaller speckles than the hyperosmotic stressed cells. The junctional staining was weak and barely visible, almost absent in places. Also, the faint cytoplasmic staining usually seen for PRPF31 in wild type cells, was not detectable, leaving the cytoplasm almost unlabelled in comparison.

The markers SC-35 (Figure 6.8, panel B2) and E-cadherin (Figure 6.9, panel B2) also produced a fainter signal compared to untreated cells and to hyperosmotic-medium treated cells, however co-localisation with PRPF31 was still observed.

The hyperosmotic stress generally appeared to affect cell morphology, as proteins and chromatin detection revealed grainy patterns. However, PRPF31 still localised to the adherens junctions, as well as to the nucleus, and strongly to structures resembling the nucleoli (as revealed by co-localisation with SC-35).

The co-localisation between PRPF31 and MAGI3 reinforces what was observed in the co-localisation experiments performed using the tagged proteins (compare with section 5.2.5 and Figures 5.6 and 5.7 of Chapter 5). However, the “nucleolar” co-localisation observed between MAGI3 and PRPF31 under stress conditions may be considered either an artefact due to abnormal conditions, such as transfecting reagents and change in medium composition, or may be considered a “normal response” which occurs in cells under certain physiological stresses (exemplified by a transfecting reagent and an altered medium in this case).

6.4 Localisation of PRPF31 in Murine Retina

We performed immunohistochemistry using the anti-PRPF31$_{484-497}$ antibody on sections of retina from a light-adapted mouse. Sections from this type of retina were extensively used for following labellings; however dark-adapted mouse retina was also used, although only to detect PRPF31.

In retina from a light-adapted mouse, anti-PRPF31$_{484-497}$ immunoreactivity was not extensively observed in the photoreceptor nuclei. As Figure 6.11 shows, the signal from PRPF31 was detected mostly in the inner segment of photoreceptors and in the outer plexiform layer. Also, PRPF31 was detected in, but mostly around, the nuclei of the cells composing the inner nuclear layer, just below the signal from the outer plexiform layer in panel A1. Anti-PRPF31$_{484-497}$ also weakly labelled the inner plexiform layer and ganglion cell layer. As mentioned, the splicing factor was observed to localise in strips bordering the nuclei of both the inner and the outer nuclear layers. This signal was more prominent in the inner nuclear layer. A retina section incubated with a secondary antibody (Cy3-conjugated, used to detect the anti-PRPF31$_{484-497}$ antibody) did not show any significant staining that may influence the interpretation of PRPF31 localisation (panel B1 of Figure 6.11).

To ascertain whether PRPF31 translocates in response to light, similar to arrestin and transducin (arrestin found in outer segments in light adapted retinas and proximal to the connecting cilium in dark adapted retina, the opposite is true for transducin; Reidel *et al.*, 2008), a retina section from a dark-adapted mouse was labelled for PRPF31. Figure 6.12 illustrates that the distribution of PRPF31 is similar to the light-adapted retina: in fact, fluorescence due to anti-PRPF31$_{484-497}$ immunoreactivity is strongly detected in the photoreceptor inner segment and in the outer plexiform layer. The signal appeared to be stronger in the IS than in the OPL as compared to the light-adapted retina (in which the opposite was observed, *i.e.* weaker in the IS and stronger in the OPL), however, this is a qualitative assessment and a detailed investigation may be required. Furthermore, this observation may be an artefact produced by the procedure of the preparation of the retina. The signal was also detected in the inner nuclear layer and in the inner plexiform layer to a lesser extent. Dark-adapted mouse retina incubated with Cy3-conjugated secondary antibody did not show any significant staining (panel B1 of Figure 6.12).

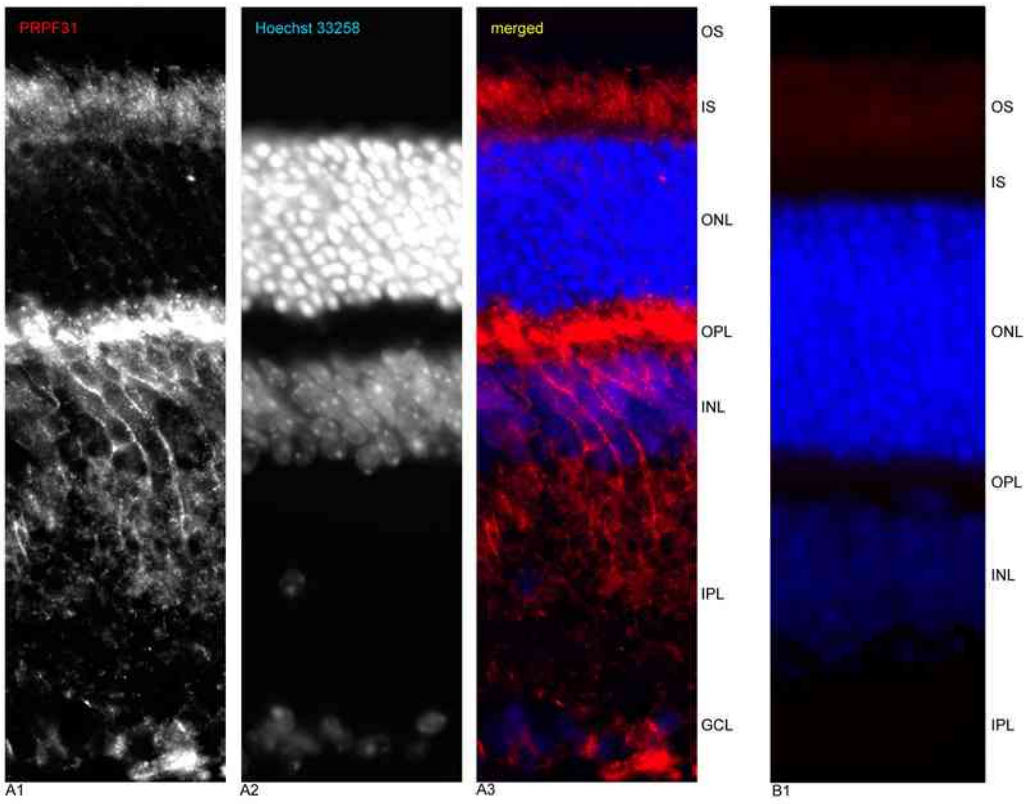


Figure 6.11. PRPF31 localisation in light-adapted mouse retina.
PRPF31 (A1; Cy3) localises to the inner segment, to the outer plexiform layer and to the inner nuclear layer but not within the nuclei of the outer nuclear layer in light-adapted murine retina. A2: Hoechst 33258 stain; A3: overlay. B1: section incubated with Cy3-conjugated secondary antibody and Hoechst 33258. Colour is indicated by legend on each black and white panel.
(OS: outer segment; IS: inner segment; ONL: outer nuclear layer; OPL: outer plexiform layer; INL: inner nuclear layer; IPL: inner plexiform layer; GCL; ganglion cell layer).

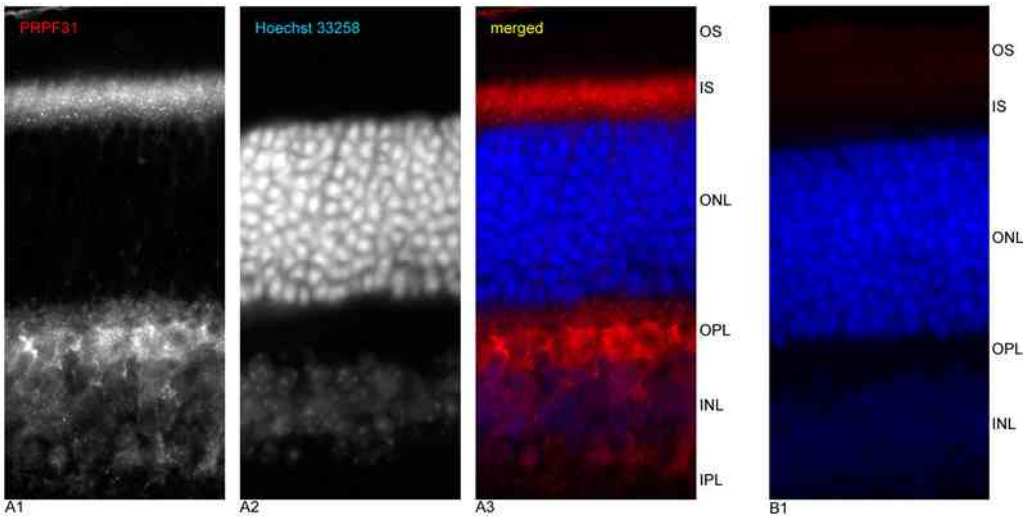


Figure 6.12. PRPF31 localisation in dark-adapted mouse retina.
PRPF31 (in A1; Cy3) localises to the inner segment, to the outer plexiform layer and to the inner nuclear layer but not within the nuclei of the outer nuclear layer in dark-adapted murine retina. A2: Hoechst 33258 stain; A3: overlay. B1: section from dark-adapted mouse incubated with Cy3-conjugated secondary antibody and Hoechst 33258. Colour is indicated by legend on each black and white panel.
(OS: outer segment; IS: inner segment; ONL: outer nuclear layer; OPL: outer plexiform layer; INL: inner nuclear layer; IPL: inner plexiform layer).

An anti-Rhodopsin antibody, a marker of the rod photoreceptor outer segments, was used to ascertain whether PRPF31 was present in the rod outer segment. Figure 6.13 illustrates that PRPF31 does not appear to localise to the outer segments of rod photoreceptors, as no yellow signal (indicative of co-localisation) is observed in the merge image with Rhodopsin (panel A4 of Figure 6.13).

From the previous observation, we proceeded to investigate the localisation of PRPF31, using the anti-TOPORS antibody. The TOPORS protein has been detected at the base of the connecting cilium of the rod photoreceptors (Dr C. Chakarova personal communication), hence the anti-TOPORS antibody was used to evaluate PRPF31 presence in the connecting cilium. Figure 6.14 illustrates the localisation of PRPF31 and TOPORS, showing that PRPF31 does not appear to localise to the connecting cilia, as no distinct yellow signal is present. The partial overlap in panel A4 may be due to a cutting plane not parallel to the plane of the retina; however no co-localisation is observed, as shown by the overlay B4 (from confocal microscopy).

Figure 6.15 shows a murine retina section stained with the anti-PRPF31$_{484\text{-}497}$ antibody and peanut agglutinin (PNA), which is a marker for cone cells. PNA detects carbohydrates found only in the extra cellular matrix sheaths of cone photoreceptors, allowing the rod and cone photoreceptors to be distinguished (van der Spuy *et al.*, 2002). This staining should highlight cone photoreceptors and may provide a better appreciation of PRPF31 presence in the retina. It is expected that PRPF31 is present in both rods and cones, as it is a ubiquitous protein; however, if localisation is limited to only the rod photoreceptors, it may be indicative of a disease mechanism underlying retinal degeneration, such as RP. The overlay A4 did not show a distinct yellow signal as the fluorescence from the cone marker was much higher than that of PRPF31. However, on closer inspection (Figure 6.15, overlay B4; from confocal microscopy), PRPF31 appears to be present in cone photoreceptors inner segments.

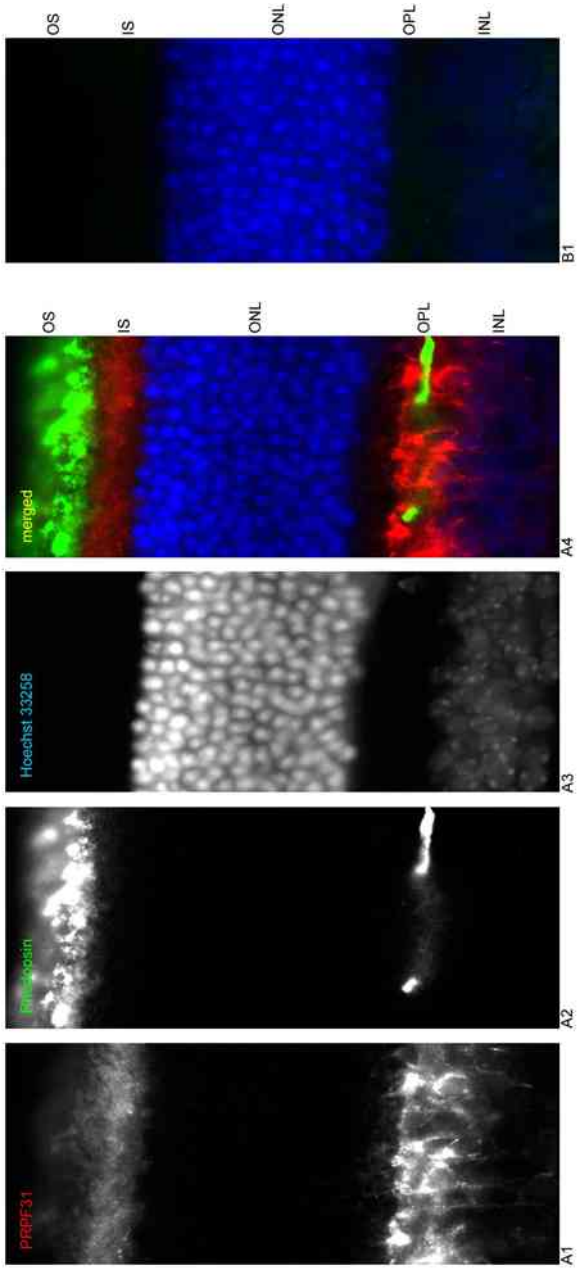


Figure 6.13. PRPF31 and Rhodopsin localisation in light-adapted murine retina.
PRPF31 (Cy3; in A1) localises to the inner segment, to the outer plexiform layer and to the inner nuclear layer, Rhodopsin localises to the outer segment only (FITC; in A2). A3: Hoechst 33258 stain; A4: overlay. B1: section from light-adapted mouse incubated with FITC-conjugated secondary antibody and Hoechst 33258. Colour is indicated by legend on each black and white panel.
(OS: outer segment; IS: inner segment; ONL: outer nuclear layer; OPL: outer plexiform layer; INL: inner nuclear layer).

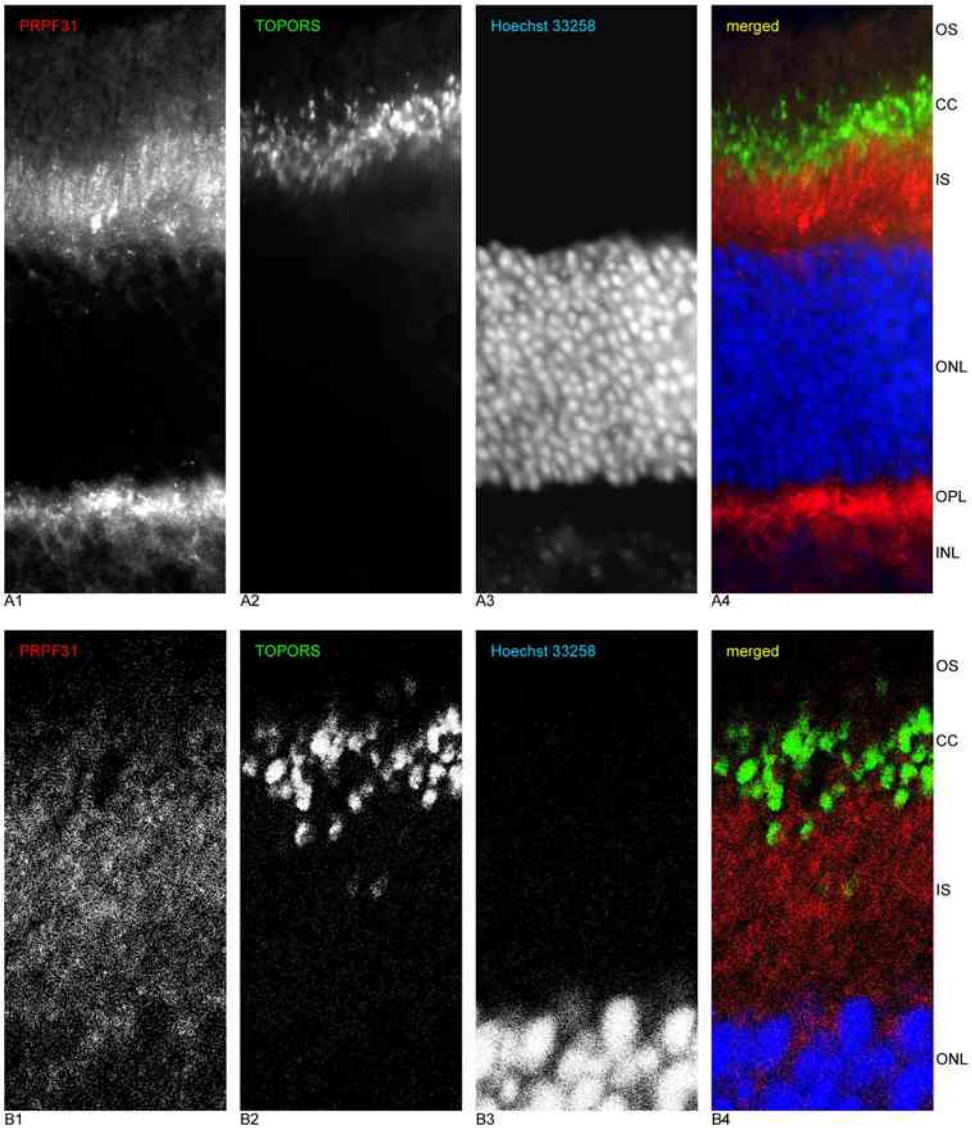


Figure 6.14. PRPF31 and Topors do not co-localise to the connecting cilium in murine retina.
Panels A illustrate the localisation of PRPF31 (Cy3; in A1 and B1) and TOPORS (FITC; in A2 and B2) in respect to the entire retina. Panels B (from confocal microscope) show an enlargement of the inner segment and connecting cilium area. A3, B3: Hoechst 33258 stain; A4, B4: overlays. Colour is indicated by legend on each black and white panel.
(OS: outer segment; CC: connecting cilium; IS: inner segment; ONL: outer nuclear layer; OPL: outer plexiform layer; INL: inner nuclear layer).

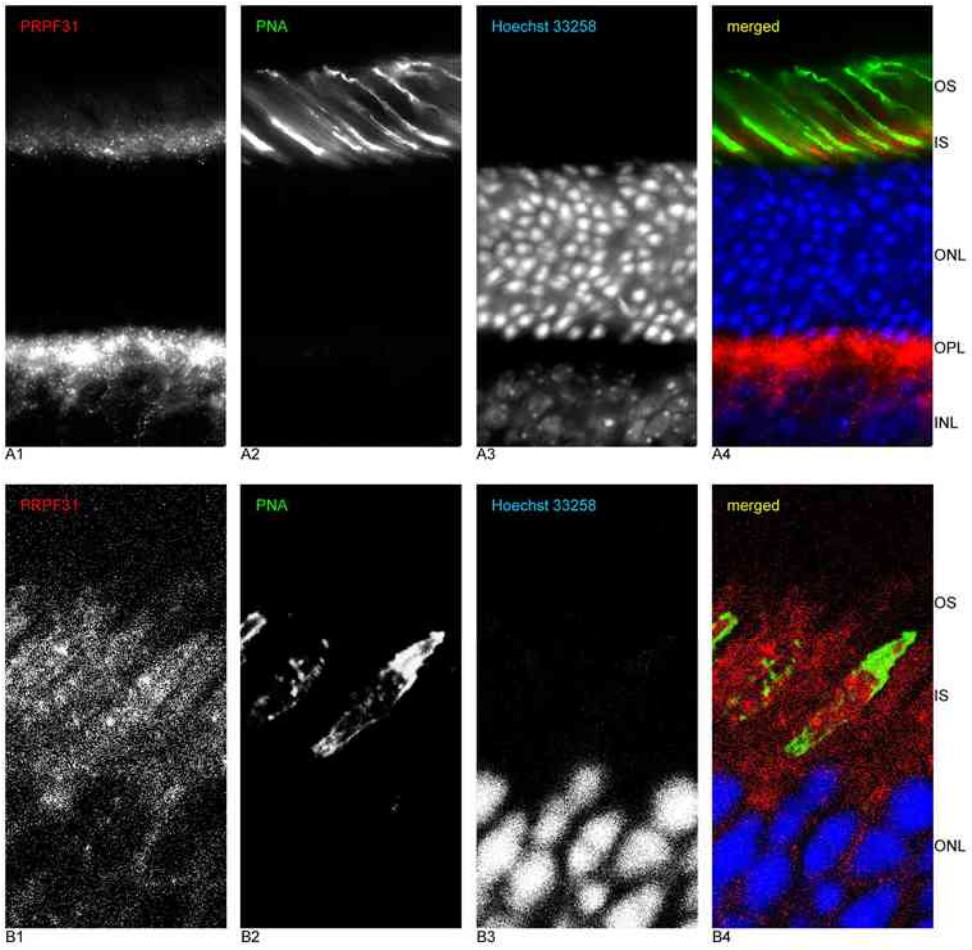


Figure 6.15. PRPF31 localises in cone photoreceptors in mouse retina.
Panels A illustrate the co-localisation of PRPF31 (Cy3; in A1 and B1) and PNA (Cy2; in A2 and B2). Panels B (from confocal microscopy) show an enlargement of the inner segment area, evidencing overlay between red and green signals. A3, B3: Hoechst 33258; A4, B4: overlays. Colour is indicated by legend on each black and white panel.
(OS: outer segment; IS: inner segment; ONL: outer nuclear layer; OPL: outer plexiform layer; INL: inner nuclear layer).

A double labelling with the anti-PRPF31$_{484-497}$ and the anti-MAGI3 antibodies was also performed, to ascertain whether PRPF31 and MAGI3 co-localise in retina. Figure 6.16 illustrates what has been previously shown for PRPF31, which localised to the inner segment, to the inner and outer plexiform layer, and the inner nuclear layer (panel A1 of Figure 6.16). Anti-MAGI3 immunoreactive fluorescence was detected in the outer segment, marking the photoreceptor outer segments; a weaker signal was also present in the inner segment. The signal also labelled the outer and inner plexiform layers (Figure, 6.16; panel A2). Therefore, it is possible that PRPF31 and MAGI3 overlap in the plexiform layers and partly to the inner segment (overlay shown in panel A4 of Figure 6.16). Panels B show a magnification of a cell from the inner nuclear layer which was observed to be labelled around the edge by both antibodies. A number of inner nuclear layer cells, bordering the inner plexiform layer and presenting this pattern were observed; however they were not abundant and hence may be a particular type of cell.

We were also interested in the second interacting partner, TRAK2, for which we had available two antibodies: the anti-TRAK2 antibody used for staining the SK-N-SH cells did not give a significant signal in retina. We therefore performed immunohistochemistry using the second antibody, anti-GRIF1$_{8-633}$, which was raised in rabbit, and hence could not be coupled with the anti-PRPF31$_{484-497}$ antibody.

Both the goat anti-TRAK2 and the commercial antibody goat anti-PRPF31$_{140-154}$ did not give a significant signal in retina, which may be due to the type of fixation used for the mouse tissue.

Nonetheless, a single staining with the anti-GRIF1$_{8-633}$ antibody was performed to investigate TRAK2 localisation and infer whether or not it co-localised with PRPF31.

From Figure 6.17, it appears that TRAK2 was detected strongly amid the nuclei of the photoreceptor cells (outer nuclear layer), in the upper portion of the outer plexiform layer and in the ganglion cell layer. The antibody also labelled the inner and outer segments, however the signal was weak. The TRAK2 localisation pattern suggests that if TRAK2 and PRPF31 interact, this may happen at the edges of their respective expression domains, such as the outer plexiform layer and the inner segment layer; however this speculation should be corroborated by performing the experiment using

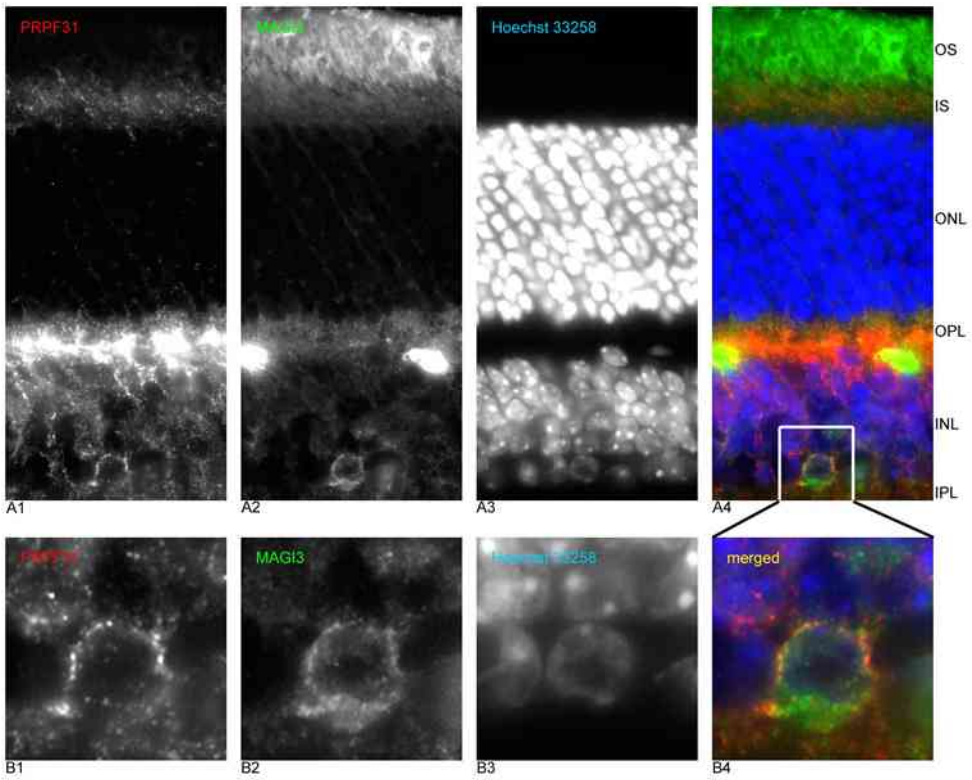


Figure 6.16. PRPF31 and MAGI3 signals overlap at the inner segment, outer plexiform layer and inner plexiform layer in mouse retina.
A1, B1: PRPF31 (Cy3); A2, B2: MAGI3 (FITC); A3, B3: Hoechst 33258; A4, B4: overlays. Colour is indicated by legend on each black and white panel.
(OS: outer segment; IS: inner segment; ONL: outer nuclear layer; OPL: outer plexiform layer; INL: inner nuclear layer; IPL: inner plexiform layer).

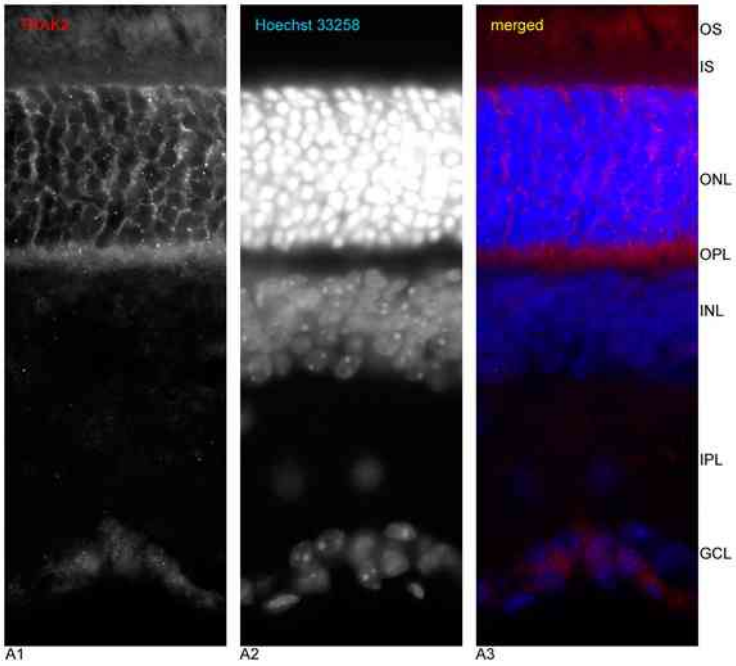


Figure 6.17. Localisation of TRAK2 in murine retina.
TRAK2 (in A1; Cy3) is predominantly detected in the outer nuclear layer, in the outer plexiform layer and in the ganglion cell layer. A2: Hoechst 33258; A3: overlay. Colour is indicated by legend on each black and white panel.
(OS: outer segment; IS: inner segment; ONL: outer nuclear layer; OPL: outer plexiform layer; INL: inner nuclear layer; IPL: inner plexiform layer; GCL: ganglion cell layer).

antibodies raised in different species in order to achieve a simultaneous detection of the proteins.

The localisation of three other splicing factors was investigated in retina: PRPF3, PRPF8 and PRPF16. *PRPF3* and *PRPF8* genes are involved in tri-snRNP assembly (Liu *et al.*, 2006) and known to be mutated in RP (Chakarova *et al.*, 2002; McKie *et al.*, 2001). PRPF16 is involved in the later stages of splicing, in structural remodeling of the spliceosome (Wahl *et al.*, 2009); to date, no mutations causing retinal degeneration have been reported in *PRPF16*.

Co-staining with PRPF31 was not possible as the antibodies were raised in the same species (rabbit).

Figures 6.18 and 6.19 show the labelling of PRPF3 and PRPF8 respectively. The two splicing factors shared a similar pattern of localisation as PRPF31: PRPF3 and PRPF8 were localised to the inner segment and outer plexiform layer; they were observed to weakly surround the nuclei of the photoreceptors, in streaks amid the nuclei as PRPF31; also, they did not appear to be present within the photoreceptor nuclei. Interestingly, both PRPF3 and PRPF8 antibodies showed a stronger signal from the inner nuclear layer than PRPF31. However, to ascertain whether there is any difference among the splicing factor distributions, a co-staining on the same retina section would be of great value, allowing a direct comparison of the localisation signals. All three splicing factors, PRPF31, PRPF3 and PRPF8, weakly labelled the ganglion cell layer.

Conversely, PRPF16 localisation appeared to be different to that of PRPF31, PRPF3 and PRPF8. The PRPF16 splicing factor localised to some photoreceptor nuclei and also around them, in a circular pattern (ONL layer in panel A1 of Figure 6.20). The anti-PRPF16 antibody also weakly labelled the outer plexiform layer, the inner nuclear layer with a granular signal, more prominent in the nucleoplasm than the other splicing factors tested. PRPF16 also appears to be present in the nuclei of the ganglion cell layer, and in the nuclei of the RPE cells. The inner segment of photoreceptor cells appeared weakly labelled.

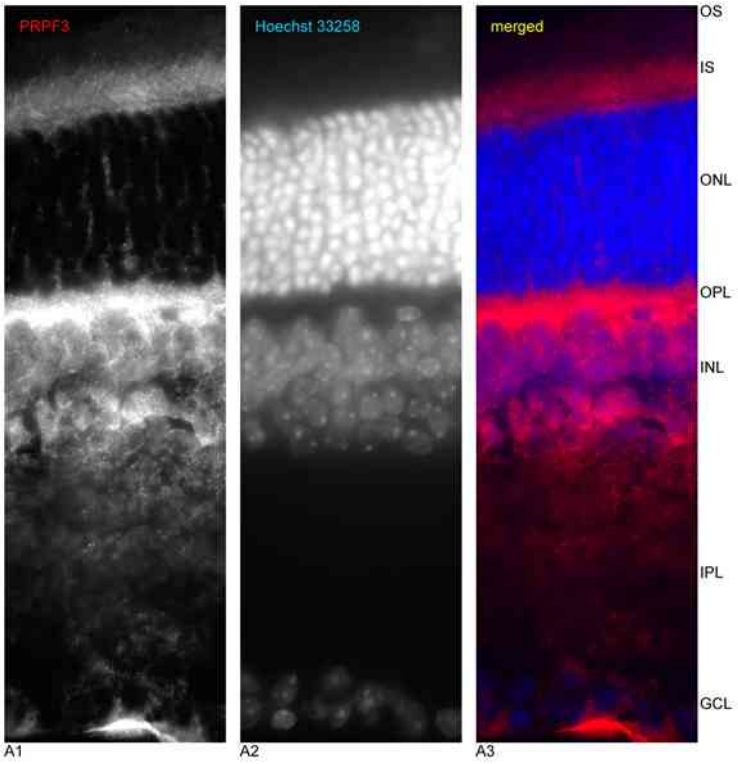


Figure 6.18. PRPF3 shows a similar pattern of localisation to PRPF31 in mouse retina.
PRPF3 (A1; Cy3) is prominently found in the inner segment, the outer plexiform layer and in the inner nuclear layer. A2: Hoechst 33258; A4: overlay. Colour is indicated by legend on each black and white panel.
(OS: outer segment; IS: inner segment; ONL: outer nuclear layer; OPL: outer plexiform layer; INL: inner nuclear layer; IPL: inner plexiform layer; GCL: ganglion cell layer).

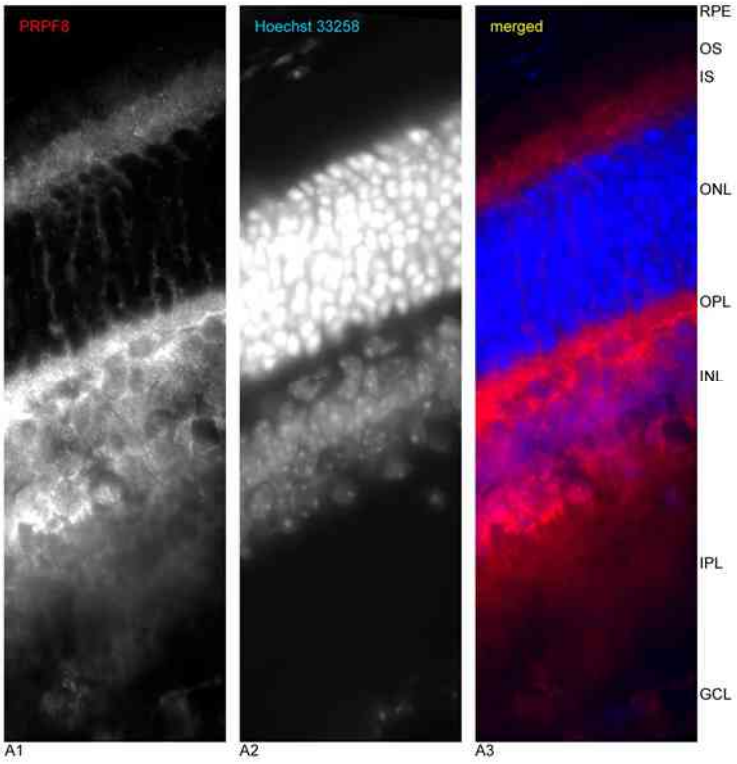


Figure 6.19. PRPF8 shows a localisation pattern comparable to PRPF31 in mouse retina.
Also PRPF8 (Cy3; in A1) localises to inner segment, outer plexiform layer and inner nuclear layer predominantly. A2: Hoechst 33258; A4: overlay. Colour is indicated by legend on each black and white panel.
(RPE: retinal pigment epithelium; OS: outer segment; IS: inner segment; ONL: outer nuclear layer; OPL: outer plexiform layer; INL: inner nuclear layer; IPL: inner plexiform layer; GCL: ganglion cell layer).

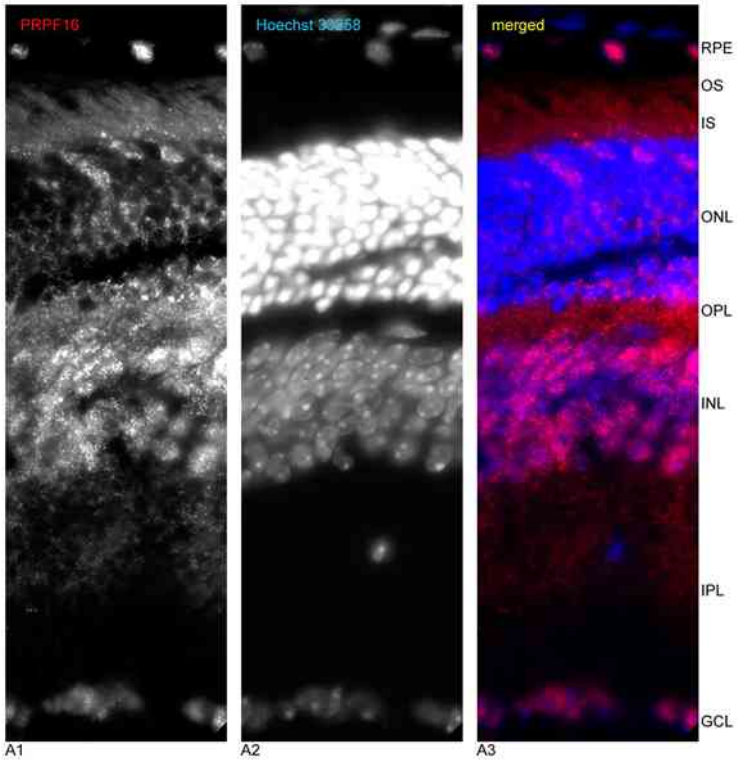


Figure 6.20. PRPF16 localisation in murine retina differs from the tri-snRNP associated splicing factors PRPF31, PRPF3 and PRPF8.
PRPF16 (Cy3; in A1) localises mostly to the inner nuclear layer and the ganglion cell layer, and weakly to the outer nuclear layer and the outer plexiform layer. It is also detected in the RPE nuclei. A2: Hoechst 33258; A4: overlay. Colour is indicated by legend on each black and white panel.
(RPE: retinal pigment epithelium; OS: outer segment; IS: inner segment; ONL: outer nuclear layer; OPL: outer plexiform layer; INL: inner nuclear layer; IPL: inner plexiform layer; GCL: ganglion cell layer).

6.5 Conclusion

As observed in Chapter 5, PRPF31 had a significant presence at the junctions in MDCK cells. In this chapter we proceeded with additional labellings to ascertain whether PRPF31 presence at intercellular junctions was affected by changes in the physiological conditions of the cells.

In wild type MDCK cells, PRPF31 was localised to the adherens junctions as the co-staining with E-cadherin revealed. PRPF31 was also shown to be absent from the nucleoli (from SC-35 co-staining), as was previously described (Makarova *et al.*, 2002), but strongly present in the rest of the nucleus as expected. However, the pattern of localisation at the junctions appeared to be grainy and mottled and not as uniform as that of E-cadherin. At times it was interrupted, appearing in different planes of the cell, thus suggesting a fluctuation of the proteins' localisation (however, it may be also due to the cells not being completely horizontal).

Treatment with CDK inhibitors to arrest cells in G0/G1 phase moderately altered the distribution of PRPF31: olomoucine treatment retained the PRPF31 signal mostly in the nucleus, and the junctions were not as intensely labelled as the untreated cells. There were fewer cells for which PRPF31 was present at the junctions and the observed signal was much weaker. On the other hand, cells treated with the CDK4 Inhibitor, still showed nuclear and junctional staining, although the latter was quite intermittent.

Heat-shocked cells and cells treated with hyperosmotic-medium showed PRPF31 at the junctions, but interestingly, its presence was more prominent than in the untreated cells. Furthermore, hyperosmotic-stressed cells showed a peculiar localisation of PRPF31 in the nucleoli, which was not observed in untreated samples. The hypoosmotic-stressed cells showed a weaker, almost imperceptible staining at the junctions for PRPF31.

Concerning the murine retina, PRPF31 would be expected to be found mostly in the nucleus of retina cells, however, not all the nuclei showed this localisation: in fact, only the inner nuclear layer and the ganglion cells (although weakly) show PRPF31 in the nucleus. The splicing factor was then observed in the inner and outer plexiform layers (layers of neuronal synapses from ganglion cells and bipolar cells, and from photoreceptor cells and horizontal and bipolar cells respectively), and in the photoreceptor inner segments, of both rods and cones. Interestingly, PRPF31 was found

at the borders of the cells on inner and outer nuclear layer. PRPF31 localisation did not appear to be greatly affected by light and dark adaptation.

MAGI3 signal in retina was predominantly observed in the outer segment, delineating the outer segments of photoreceptors; in the remaining layers of the tissue MAGI3 signal was generally faint, labelling weakly the inner segment and the outer plexiform layer. Interestingly, its presence was observed in few cells along the lower section of the inner nuclear layer, overlapping with the signal from PRPF31. From the localisation sites of MAGI3, it is possible to reason that the signals for the splicing factor and MAGI3 may overlap in the outer plexiform layer and in the region between the inner nuclear layer and the inner plexiform layer.

Results from the PRPF3 and PRPF8 labellings supported the observations regarding PRPF31: these splicing factors were also detected in the inner segments, the outer plexiform layer, the inner nuclear layer and ganglion cell layer, in a similar manner as PRPF31, with streaks lining the nuclei of photoreceptors cells and inner nuclear layer, and in the ganglion cells with a weak signal from the nuclei.

PRPF16 showed a different distribution in retina: the splicing factor was observed in the nuclei of the inner nuclear layer, and also in some nuclei of the photoreceptor cells, and in that same layer also it surrounded the nuclei in a circular fashion. Opposed to PRPF31 (and PRPF3 and PRPF8), PRPF16 was observed in some nuclei of the retinal pigment epithelium.

7. Retina Isoforms and Mutation Screening

The results discussed in the previous chapters suggested MAGI3 and TRAK2 as novel interacting partners of PRPF31. However, *MAGI3* and *TRAK2* genes are not retina-specific but are instead expressed in a wide range of tissues. It may be that a retina-specific transcript is produced, whose protein interacts with PRPF31: Part I of this chapter is dedicated to the study conducted on human retina RNA to verify the possible existence of retina-specific transcripts, for both *MAGI3* and *TRAK2*.

If the two interactants are involved in PRPF31 pathway, it is possible that a mutation in their genes may lead to retinal degeneration. Therefore the genes may be considered candidates for screening of mutations in patients affected by RP. Part II addresses screening of patients, affected by autosomal dominant RP, in order to find mutations in *MAGI3*, potentially linked to retinal degeneration.

Part I Isoform Study on Retina for *MAGI3* and *TRAK2*

PRPF31 is a splicing factor and ubiquitously expressed protein, however mutations in *PRPF31* only result in retinal degeneration. As described in Chapter 3, the interacting proteins found from the bovine library corresponded to human MAGI3 and TRAK2, later proven to be associated with PRPF31 by co-immunoprecipitation and co-localisation studies (Chapters 4 and 5). Both MAGI3 and TRAK2 are expressed in a wide range of tissues, (see Chapter 3, sections 3.6.1, 3.6.3 and 3.7). It was proposed that these proteins may have a retina-specific transcript, which interacts with PRPF31 and produces eye specific function. Therefore, the amplification of the transcripts as predicted by the Ensembl database was performed, using human retina RNA. *PRPF31* transcript was not included in the experiments, since the presence of alternative, retina-specific isoforms of *PRPF31* was recently excluded by Tanackovic and Rivolta

(2009): the authors have analyzed PRPF31 transcripts by PCR on cDNA derived from human retina, a whole eye as a positive control, and HeLa cells as a negative control. The detected changes affected the 5' UTR region, but not the coding region. This was also proven by western blot, showing the presence of a single band, corresponding to the full-length protein.

7.1 Amplification of *MAGI3* and *TRAK2* Coding Sequences from Human Retina Total RNA

Human retina total RNA was used to prepare retina cDNA, following the procedure described in section 2.9.6. Oligonucleotides (listed in section A.3 and A.4 for *MAGI3* and *TRAK2* respectively, in Appendix) were designed to allow the amplification of DNA fragments of approximately ~1kb in length, using KOD Hot Start DNA polymerase. The aim was to investigate whether a different protein isoform was expressed in retina.

7.1.1 *MAGI3*

Figure 7.1 is a representation of the four transcripts of *MAGI3* (Ensembl database, accessed January 2009).

MAGI3-001 is considered the reference sequence in Ensembl: it shows the longest 5' UTR and 3' UTR (see Figure 7.2 A and B), and comprises 21 exons. MAGI3-002 differs from MAGI3-001 by an extra exon (exon 7a, 75 bp long), in between exons 7 and 8 of the reference transcript. MAGI3-003 shares the same exon architecture as MAGI3-001, although it presents a shorter 5' UTR and a portion of the 3' UTR is missing, generating an extra exon, 21a. MAGI3-004 shows a shorter 5' UTR and shares the same exon architecture as MAGI3-001, but displays a shorter exon 21.

Intron 1 is almost 160 kb, intron 2 is ~31 kb. Also introns 6, 9 and 16 are long (~19, kb, ~20 kb and ~12 kb respectively). The gene spans ~295 kb.

The experimental design was aimed to verify which isoform(s) is(are) expressed in retina, and if there are new, alternatively-spliced retina-specific transcripts.

The lower part of Figure 7.1 provides a schematic view of the portions of transcripts designed to be amplified. The approach implied the design of eight overlapping fragments: in the figure, a black bar is assigned to each of them (eight bars in total,

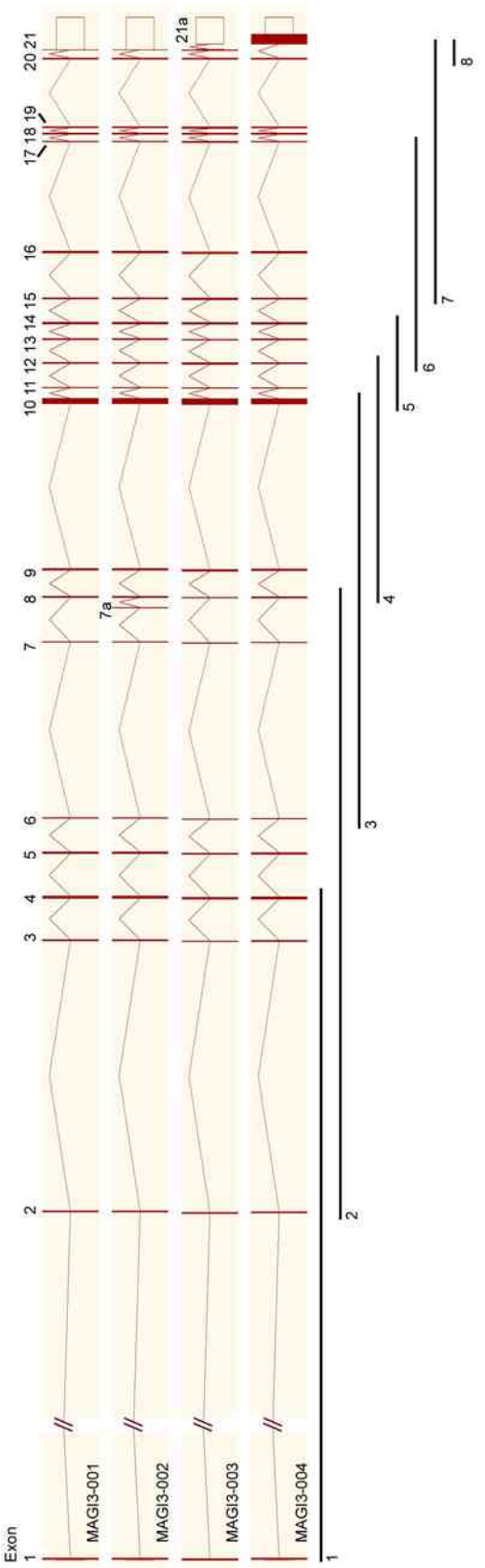


Figure 7.1. Overview of *MAGI3* transcripts.
The figure represents the alignment of the four transcripts generated by *MAGI3*. Each vertical bar represents an exon, whereas a bar linking two exons represents an intron. The lower section of the figure shows eight bars spanning the adjacent exons to be amplified (see Table 7.1). Image adapted from www.ensembl.org (accessed January 2009).

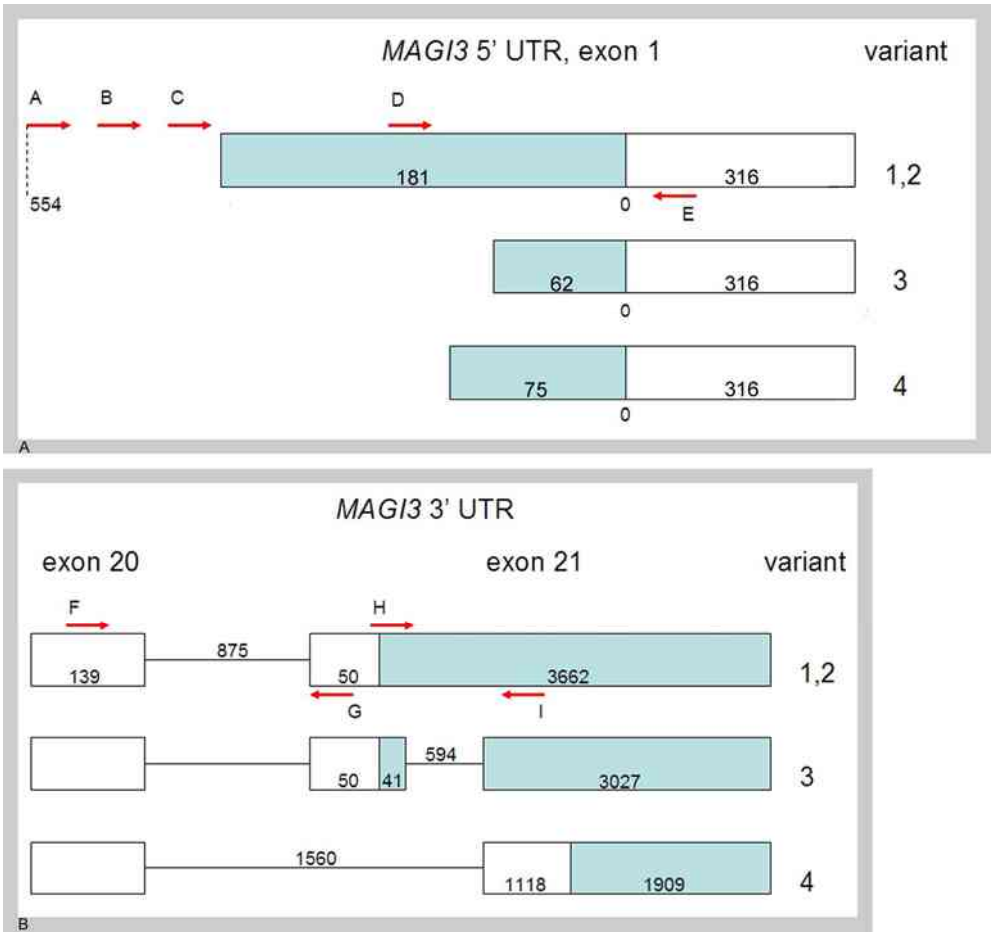


Figure 7.2. Diagram to show *MAGI3* UTRs.
In A: *MAGI3* 5'UTR; in B: *MAGI3* 3'UTR. Rectangles represent exons (light blue: UTR; white: translated; connecting lines: introns - not to scale). Numbers 1, 2, 3 and 4 on the right of the transcripts represent the variant (1 is for MAGI3-001, and forth). Numbers inside the exons represent the length in base pairs. Red arrows denote the position of primers, marked by capital letters.

numbered 1 to 8), spanning several adjacent exons, and overlapping by two exons (for further details on the length, see Table 7.1).

Although our interest was mainly focused on the coding region of the transcripts, the UTRs were included in the experiments to ensure that any potential exons lying within these regions were not excluded.

Figure 7.2 provides a diagrammatic representation of the 5' and 3' untranslated regions of *MAGI3*. For the 5' region, the approach consisted of amplifying regions which were assumed to be transcribed (hybrids of translated and untranslated regions), and regions which were assumed not to be transcribed (according to Ensembl): hence, a set of three forward primers (A, B and C in Figure 7.2 A) were designed on the upstream intronic region of the 5' UTR, encompassing <600 bp; a unique forward primer was designed to anneal on the UTR region (primer D in Figure 7.2 A, position -89). The four forward primers were coupled with a unique reverse oligonucleotide (E) pairing to the translated sequence, after the ATG start site.

For the 3' region, the approach involved the generation of amplicons spanning part of the penultimate exon and part of the last exon (forward primers F and H respectively, Figure 7.2 B), using different 3' end points (reverse primers G and I, Figure 7.2 B).

Following are three tables, which summarise the oligonucleotide pairs used, the expected length of the corresponding products, and the transcript versions of the gene. Table 7.1 is for the coding region of the *MAGI3* mRNAs. Tables 7.2 and 7.3 are for the UTR region. Table 7.2 and 7.3 also feature small drawings of the expected products, based on the drawing of Figure 7.2.

Table 7.1. Expected fragment sizes for *MAGI3* mRNA sequence (translated region).

#	1	2	3	4	5	6	7	8
Exons	1-4	2-8	6-10	8-12	10-14	12-17	15-21	20-21
Length	763	855 / 930 (vv 1,3,4 / 2)	1028 / 1103 (vv 1,3,4 / 2)	1079	980	920	938	1057 (vv 1,2) / 463 (v 3) / 372 (v 4)

Summary of the fragments to be amplified: the exon interval, the length in bp is indicated for each fragment (1 to 8 as in Figure 7.1); for fragments 2, 3 and 8, the sizes according to the isoforms are reported (for example, "v1" is for variant MAGI3-001).

Table 7.2. Expected fragments for *MAGI3* 5' UTR.

Primers 5'	MAGI3-001	MAGI3-002	MAGI3-003	MAGI3-004	Genomic
A + E (9)	n.e.	n.e.	n.e.	n.e.	607
B + E (10)	n.e.	n.e.	n.e.	n.e.	413
C + E (11)	n.e.	n.e.	n.e.	n.e.	281
D + E (12)	142	142	n.e.	n.e.	142

The oligonucleotides refer to the Figure 7.2 A. The table provides the expected product sizes (bp) according to the pair of oligonucleotides used and the transcript version, plus genomic DNA. N.e. stands for "not expected", as no amplicons is expected. The rectangles depict the expected products, referring to Figure 7.2 A (light blue: UTR; white: translated - not to scale).

Table 7.3. Expected fragments for *MAGI3* 3' UTR.

Primers 3'	MAGI3-001	MAGI3-002	MAGI3-003	MAGI3-004	Genomic
F + G (13)	138	138	138	n.e.	1013
F + I (14)	1024	1024	430	339	1899
H + I (15)	883	883	289	n.e.	883

Primer pairs according to Figure 7.2 B. The table presents the expected product sizes (bp) according to the primer pair, the transcript and genomic DNA. N.e. stands for "not expected". The rectangles illustrate the expected products, referring to Figure 7.2 B (light blue: UTR; white: translated - not to scale).

Figure 7.3 shows the products of the amplification of *MAGI3* coding sequence: the primers were designed to amplify the fragments listed in Table 7.1. Genomic DNA (*g* lanes) was included in the amplification. Lanes corresponding to the fragments are marked by progressive numbers 1-8 (see Table 7.1 for expected sizes):

- lane 1: exons 1-4 generated a fragment lower than 800 bp (expected size: 763 bp);
- lane 2: exons 2-8 produced a ~800 bp fragment (expected 855 or 930 bp), whose sequence matched to the 855 bp version. A higher size product (~1000 bp) was observed, suspected to correspond to MAGI3-002 (which shows an extra exon, 7a, of 75 bp): after purification and sequencing, it appeared to be a non-specific amplicon (not corresponding to MAGI3-002 or to any genomic sequence);
- lane 3: exons 6-10 generated a fragment at ~1000 bp (expected 1028 or 1103 bp). Its sequencing revealed that it was the 1028 bp product. The gel did not show any extra band suggesting the presence of the extra exon 7a, hence supporting that the upper band on lane 2 was an artefact;
- lane 4: exons 8-12 generated a fragment higher than 1000 bp, corresponding to the expected 1079 bp;
- lane 5: exons 10-14 produced an amplicon higher than 1000 bp, which sequencing revealed was the expected 980 bp size. Abnormalities in the agarose gel may have compromised a proper sizing. The two lower, weak bands were extracted but their sequences failed to align to *MAGI3*, thus they were considered non-specific products;
- lane 6: exons 12-17 generated a fragment at ~1000 bp, which corresponded to the expected 920 bp product;
- lane 7: exons 15-21 product is a band at ~1000 bp, which corresponded to the expected 938 bp;
- lane 8: exons 20-21 generated products at 463 bp and 372 bp. Also, a ~2000 bp band was present in the genomic DNA lane (red arrow), hence proving that the conditions were optimal to allow the synthesis of a high molecular weight product (the 1057 bp product for MAGI3-001 and MAGI3-002).

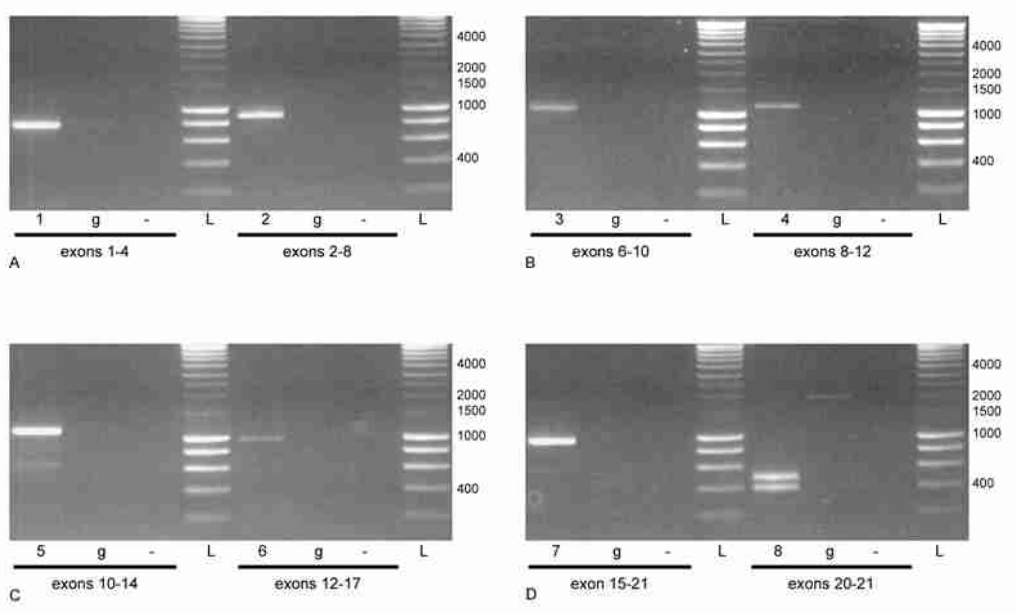


Figure 7.3. Agarose gel electrophoresis of *MAGI3* CDS amplification products.
Gel lanes are numbered according to fragments corresponding to Figure 7.1 and to Table 7.1. Lane 1: fragment 1 (763 bp); lane 2: fragment 2 (855 bp); lane 3: fragment 3 (1028 bp); lane 4: fragment 4 (1079 bp); lane 5: fragment 5 (980 bp); lane 6: fragment 6 (920 bp); lane 7: fragment 7 (938 bp); lane 8: fragment 8 (463 bp and 372 bp). L is for ladder (bp), g is for genomic DNA, - is for negative control.

Figure 7.4 shows the amplification products of the 5' UTR of *MAGI3*:

- lane 9: primer pair A+E did not generate any visible product; however, the low intensity of the product in the corresponding genomic lane may suggest that further optimisation of the experimental conditions would be desirable.
- lanes 10, 11: no amplicon was expected for primer pairs B+E and C+E; however bands of molecular size comparable to the genomic DNA product (each indicated by a red arrow, in lanes 10 and 11 respectively) were observed;
- lane 12: primer pair D+E generated a fragment <200 bp which corresponded to the expected (142 bp).

Figure 7.5 shows the amplification products of the 3'UTR of *MAGI3* (see Table 7.3 for a reference on product sizes):

- lane 13: for F+G pair, a band of 138 bp was present as expected;
- lane 14: F+I primers generated two amplicons above and below 400 bp, in agreement with the product sizes expected for transcript versions 003 and 004 (430 bp and 339 bp respectively). No band at ~1000 bp denoting versions 001 and 002 was observed. A weak 2000 bp band is present in the genomic DNA lane, red arrow);
- lane 15: H+I pair generated a unique fragment at ~300 bp, corresponding to the product size expected for transcript 003.

The products from primer pairs F+I and H+I suggest that the retina transcripts of *MAGI3* do not show the longer UTR described for transcripts 001 and 002.

Reactions performed using "cDNA" made from RNA withouth the reverse transcriptase enzyme did not show any significant band suggestive of genomic DNA contamination (panels C in Figures 7.4 and 7.5).

All products were gel-extracted, sequenced and compared to the deposited *MAGI3* UTR sequence.

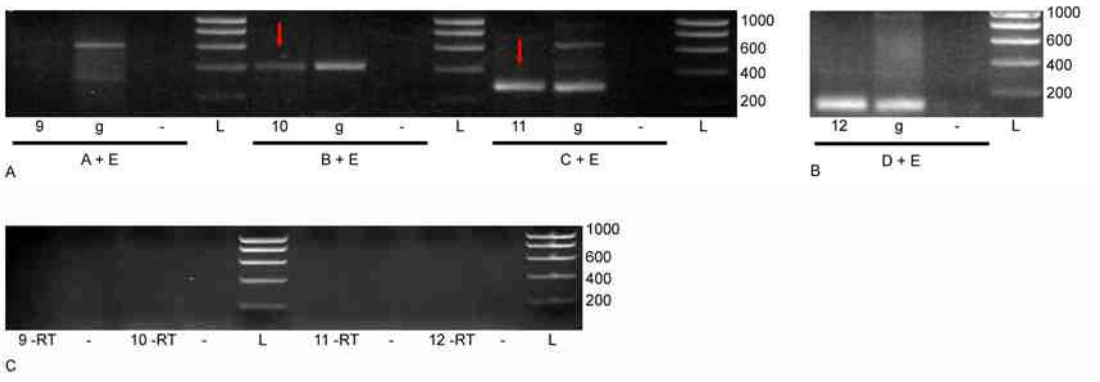


Figure 7.4. Agarose gel electrophoresis of *MAGI3* 5'UTR amplification products.
Gel lanes are numbered according to numbers in brackets in Table 7.2. Lanes 10, 11 and 12 show bands at ~400bp, ~300bp and ~150bp respectively. Panel C shows the reactions using cDNA prepared without reverse transcriptase (-RT). L is for ladder (in bp), g is for genomic DNA, - is for negative control.

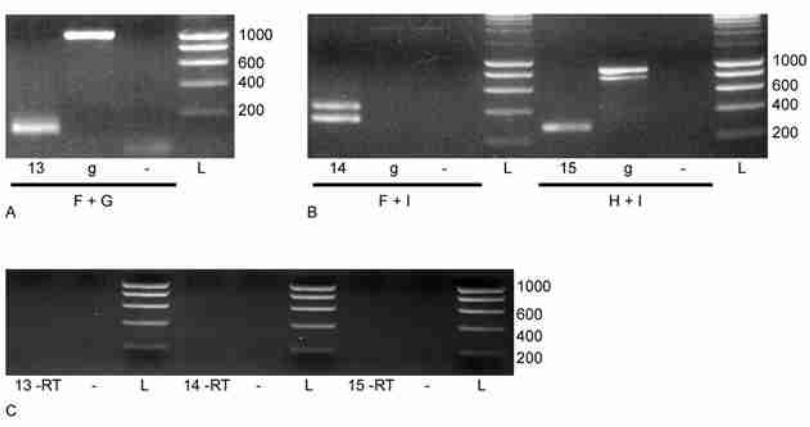


Figure 7.5. Agarose gel electrophoresis of *MAGI3* 3'UTR amplification products.
Gel lanes are numbered according to numbers in brackets in Table 7.3 (primers' ordinate). Lane 13 shows an expected band at ~138 bp; lane 14 shows two bands at 430 bp and 339 bp, expected for transcripts 003 and 004; lane 15 shows a band at 289 bp, expected for transcript 003. Panel C shows the reactions using cDNA prepared without reverse transcriptase (-RT). L is for ladder, g is for genomic DNA, - is for negative control.

The results so far described suggest that no new alternatively spliced transcript for *MAGI3* is present in retina. However, the findings determined that the retina transcripts of *MAGI3* may be hybrids between MAGI-001 and MAGI-003/MAGI-004. In fact, the retina transcript showed a 5'UTR similar to MAGI-001, but longer; this experimental data was also supported by EST data (GenBank accession numbers: CN286170.1 from embryonic stem cells, and BX333323.2 from neuroblastoma, 394 and 405 bp long upstream the beginning of 5' UTR respectively; retrieved through human ESTs database search using exon 1 and part of the region upstream as a query). Concerning the 3' region, it appeared that two isoforms (MAGI3-003, MAGI3-004) out of four are expressed in retina (see lane 8, panel D of Figure 7.3), as only the 463 and 372 bp products were amplified, and not the longer 1057 bp, which would have identified isoform 001. The accuracy of the amplification was proven by a genomic band at 1932 bp, which corresponded to the expected genomic DNA amplicon. Additionally, the extra exon 7a, which characterises MAGI-002, was not detected in two different reactions, likely suggesting that it is not transcribed in the retina.

In conclusion it appears that the protein MAGI3 is present in two different species in retina: they differ at their C-terminal, as one corresponds to the translation of MAGI-003 and the longer one to the translation of MAGI-004.

7.1.2 *TRAK2*

According to Ensembl database (accessed January 2009), the human *TRAK2* gene generates one unique transcript, TRAK2-001, shown in Figure 7.6: it encompasses 16 exons, of which the first exon and part of the second are not translated (UTR); also, most of exon 16 represents untranslated region. Introns 1 and 2 are quite large, being ~30 kb long and ~12 kb long respectively.

Figure 7.6 includes also four bars referring to the portions of transcript designed to be amplified: as for *MAGI3*, also in this case the approach implied the design of fragments overlapping by two exons, and spanning adjacent exons for a maximum size of ~1kb.

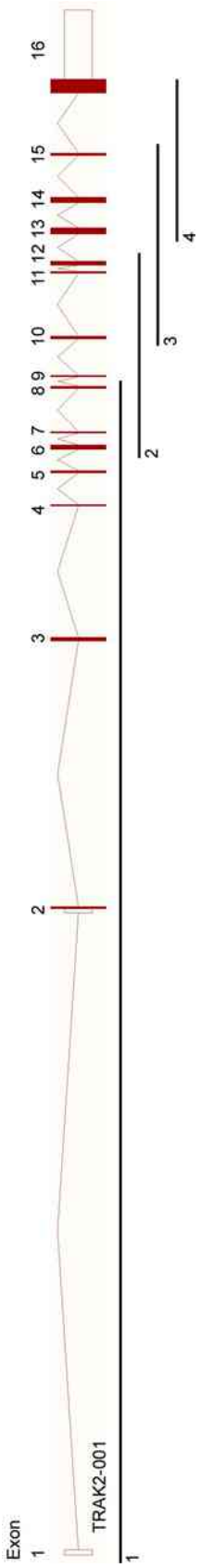


Figure 7.6. Ensembl overview of *TRAK2* transcript.
The above figure represents the mRNA generated by *TRAK2* (TRAK-001). Each vertical bar represents an exon, whereas a bar linking two exons represents an intron. The black bars span the adjacent exons to be amplified. Image adapted from www.ensembl.org (accessed January 2009).

Figure 7.7 provides a schematic view of the untranslated regions of *TRAK2*. The 5' UTR is <400 bp, and comprises exon 1 (182 bp) and part (199 bp) of exon 2, interrupted by a ~30kb-long intron. The method consisted of amplifying exons 1 and 2, using forward oligonucleotides priming the upstream intronic region and the 5' sequence of exon 1 (primers A and B respectively); the reverse primers were designed on the 5' and 3' regions of exon 2 (primer E and F). Other forward oligonucleotides were designed to anneal to the 5' sequence of exon 2 and the intronic region upstream exon 2 (primers D and C respectively).

Concerning the 3'UTR, exon 16 is ~4kb in length, of which only ~600 bp are translated and the remaining ~3.5 kb constitute the UTR. The fragments encompassed exon 15 and part of exon 16: forward primers were designed on the 5' area of exon 15 and 16 (primers G and H respectively), whereas the reverse primers were designed on the end of the translated sequence (primer I, exon 16) and at ~1 kb downstream the start of exon 16 (primer L).

The following tables 7.4, 7.5, and 7.6 summarise the characteristics of the fragments to be amplified: Table 7.4 corresponds to the coding sequence of the transcript: each number refers to the bars in Figure 7.6, providing the exon interval and the expected size. Tables 7.5 and 7.6 are dedicated to the various products of the 5' and 3' UTRs respectively.

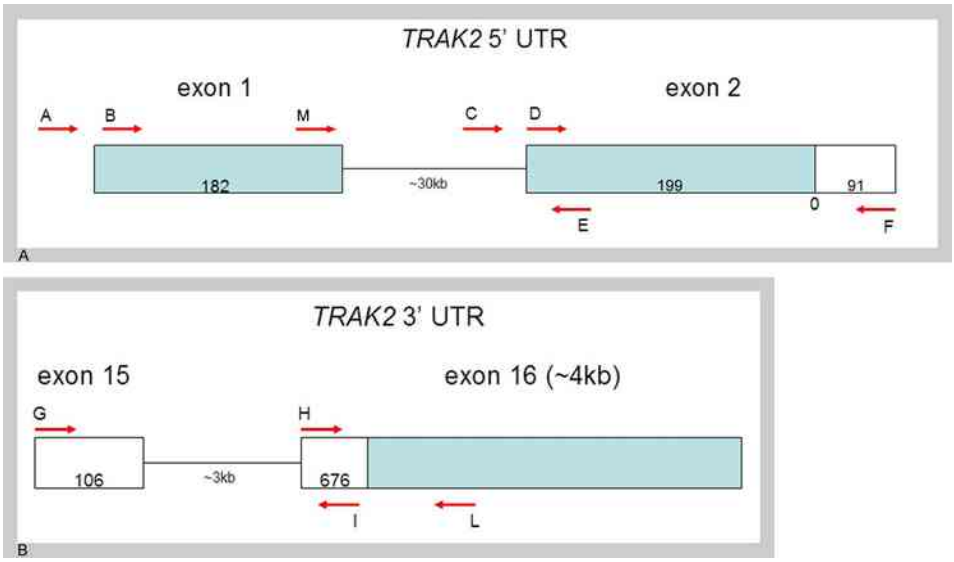


Figure 7.7. Drawing of *TRAK2* UTRs.
A: *TRAK2* 5'UTR; B: *TRAK2* 3'UTR. Rectangles represent exons (light blue: UTR; white: translated; connecting lines: introns - not drawn to scale). Numbers inside the exons represent the length in bp. Red arrows denote the position of primers used, marked by capital letters.

Table 7.4. Expected fragment sizes for *TRAK2* translated mRNA sequence.

#	1	2	3	4
Exons	1-8	6-12	10-15	13-16 (TGA)
Length	1281	917	1094	1348

Summary of the characteristics of each fragment to be amplified: for each segment 1 to 4 (to be compared with the black bars in Figure 7.6), the exon interval and its length in bp are indicated.

Table 7.5. Expected fragments for *TRAK2* 5' UTR.

Primer pairs	A + E (5)	B + E (6)	C + E (7)	B + F (8)	C + F (9)	D + F (10)
TRAK2-001	n.e.	247	n.e.	465	n.e.	290
Genomic DNA	n.e.	n.e.	338	n.e.	556	290

The oligonucleotides refer to the Figure 7.7 A. The table summarises the expected product sizes (in bp) according to the pair of oligonucleotides, the transcript and genomic DNA; numbers in brackets correspond to gel lanes in Figure 7.9. N.e. stands for "not expected", as either the region not transcribed, or the size of the amplicon is too large. The rectangles illustrate the expected products, referring to Figure 7.7 A (light blue: UTR; white: translated - not to scale).

Table 7.6. Expected fragments for *TRAK2* 3' UTR.

Primer pairs	G + I (11)	H + I (12)	H + L (13)
TRAK2-001	729	671	894
Genomic DNA	n.e.	n.e.	894

The primer pairs are in accordance with Figure 7.7 B. Expected product sizes (in bp; numbered as gel lanes in Figure 7.10) are in relation to transcript and genomic DNA. N.e. stands for "not expected", as the size of the amplicon is too large. The rectangles depict the expected products, referring to Figure 7.7 A (light blue: UTR; white: translated - not to scale).

Figure 7.8 illustrates the products obtained from the amplification of the coding sequence of *TRAK2*. Lanes corresponding to the designed fragments are numbered according to Table 7.4. The experiment also included amplification of genomic DNA (*g* lanes). The amplified products were comparable to the expected products:

- lane 1: exons 1-8 product was higher than 1000 bp, and the expected length was 1281 bp;
- lane 2: exons 6-12 generated a fragment lower than 1000 bp, and the expected product was 917 bp;
- lane 3: exons 10-15 generated a product higher than 1000 bp, corresponding to the expected product (1094 bp);
- lane 4: exons 13-16 presented a band higher than 1000 bp, corresponding to the expected 1348 bp.

All the bands were extracted and purified from the agarose gel, sequenced and successfully aligned to *TRAK2* reference sequence.

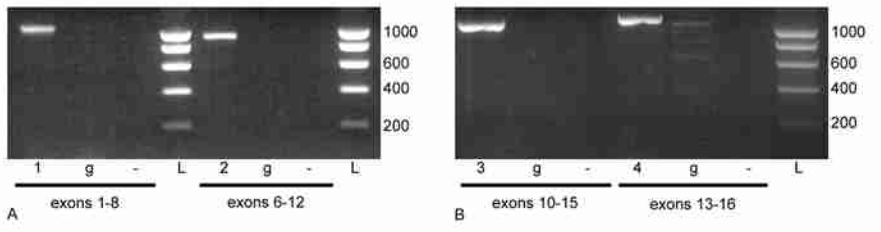


Figure 7.8. Agarose gel electrophoresis of TRAK2 CDS amplification products.
Gel lanes are numbered according to fragments corresponding to black bars of Figure 7.6 and to Table 7.4. Lane 1: fragment #1 (1281 bp); lane 2: fragment #2 (917 bp); lane 3: fragment #3 (1094 bp); lane 4: fragment #4 (1348 bp). L is for ladder (in bp), g is for genomic DNA, - is for negative control.

Figure 7.9 shows the amplification products for the 5' UTR of *TRAK2*:

- lane 5: primers pair A+E did not generate any product: as the forward primer anneals on a region upstream exon 1, and the reverse primer anneals on exon 2, no amplification is expected (either the region upstream exon 1 is not present in cDNA, or the amplicon is too large to be amplified, >30 kb for genomic DNA);
- lane 6: one band higher than 200 bp, corresponding to the product generated by B+E oligonucleotides (247 bp);
- lane 7: amplification product of C+E generated a band lower than 400 bp, observed in the genomic (*g*) lane as well (expected size: 338 bp). Both products were sequenced and corresponded to TRAK2, part intron 1 and part exon 2 (this result was not supported by any EST data). The product was not observed in a reaction using cDNA prepared without reverse transcriptase (7 –RT lane in panel F);
- lane 8: B+F generated a band at ~400 bp, corresponding to the expected product (465 bp);
- lane 9: C+F generated a band at ~500 bp, present also in the genomic DNA (*g*) lane (expected size: 556 bp). The sequencing of the two bands revealed it corresponded to TRAK2, part intron 1 and part exon 2, as observed for C+E primer pair (lane 7). The product was not detected in the reaction using cDNA prepared without reverse transcriptase (9 –RT lane in panel G);
- lane 10: primer pair D+F generated a band at ~300 bp, corresponding to the expected product size (290 bp, present on both cDNA and genomic DNA as expected).

Considering the findings from C+E and C+F primer pairs, a new forward primer was designed on the 3' end of exon 1 (primer M in Figure 7.7 A), to be coupled with reverse primer E. However, no amplicon was generated (not shown).

Also in this case, reactions using "cDNA" made from RNA without the reverse transcriptase enzyme were performed, not showing bands suggestive of genomic DNA contamination (panels F and G in Figure 7.9).

The amplification of the 3' UTR of *TRAK2* generated three amplicons (Figure 7.10), which corresponded to the deposited sequence:

- lane 11 (panel A): primers pair G+I generated a 729 bp product (band at ~800bp);
- lane 12: primers pair H+I generated a product of 671 bp (band at ~600bp);

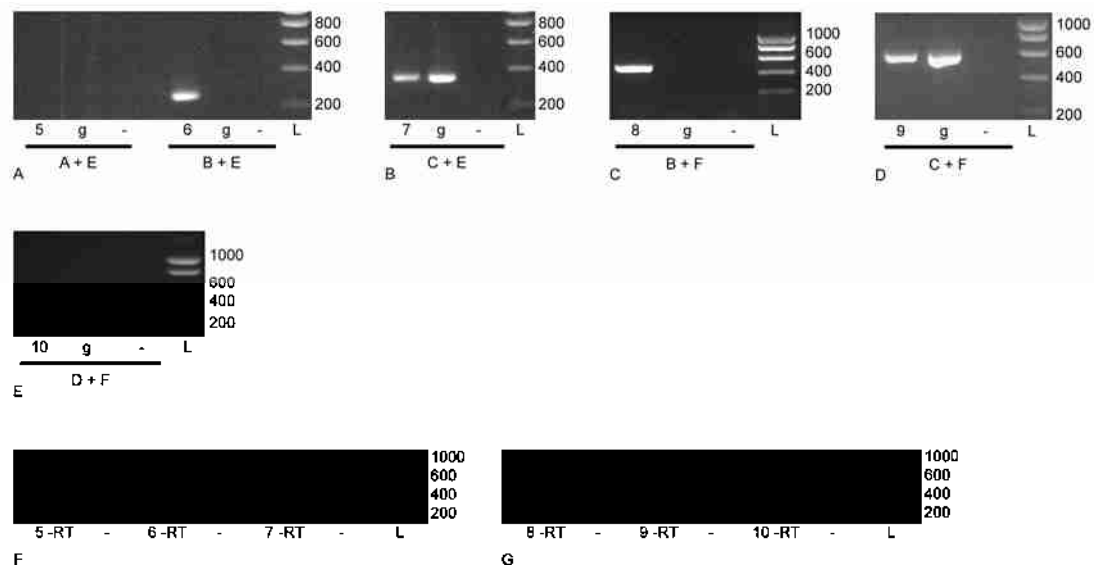


Figure 7.9. Agarose gel electrophoresis of *TRAK2* 5'UTR amplification products.
Gel lanes are numbered according to numbers in brackets in Table 7.5. Lanes 6, 7, 8, 9, 10 show bands at ~200bp, ~300bp, ~500bp, ~500bp and ~300bp respectively. To note, genomic lanes of C+E (panel B) and C+F (panel D) also show bands at ~300bp and ~500bp respectively. Panels F and G show the reactions using cDNA prepared without reverse transcriptase (-RT). L is for ladder (in bp), g is for genomic DNA, - is for negative control.

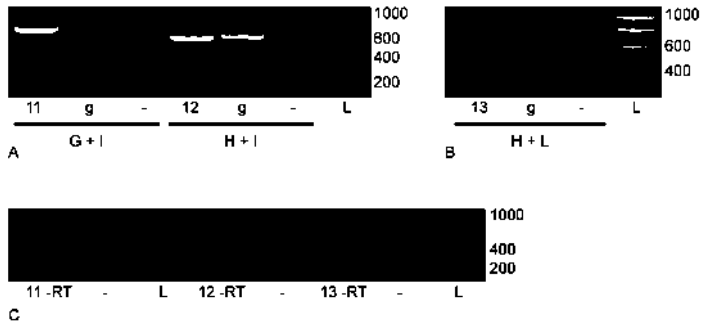


Figure 7.10. Agarose gel electrophoresis of *TRAK2* 3'UTR amplification products.
Gel lanes are numbered according to numbers in brackets in Table 7.6. Lanes 11, 12 and 13 show each a band at ~800bp, ~600bp and ~1000bp respectively. Panel C shows the reactions using cDNA prepared without reverse transcriptase (-RT). L is for ladder (in bp), g is for genomic DNA, - is for negative control.

- lane 13: primers pair H+L generated a product of 894 bp (band at ~1000bp).

No significant bands were observed in reactions using "cDNA" made from RNA without reverse transcriptase (panel C of Figure 7.10), as expected.

The amplification and subsequent sequencing of *TRAK2* transcript fragments suggested the presence of the unique transcript in the human retina, corresponding to the sequence deposited in Ensembl database. The amplification of the coding sequence did not reveal any variation, suggesting that the protein present in retina may correspond to the translation of the reference sequence. Also the amplification of the 3' UTR did not suggest any diversity from the deposited sequence: however, the entire length of the 3'UTR was not characterised (~3.5kb), but only ~300bp. Conversely, the study on the 5' UTR region revealed an extra portion upstream exon 2, adjacent to its 5' end: this experimental data was repeatedly found but was not supported by any EST data. It may be that the retinal *TRAK2* transcript possesses a longer exon 2 (not excluding an alternative upstream exon 1, or more exons), which could be better investigated with a different technique, such as RACE. Furthermore, this result should be examined in a range of tissues, to assess whether it is retina-specific or not. The experiments were not continued further due to time limitations.

Part II Mutation Screening of adRP Patients for *MAGI3*

A human disease gene may be identified in a large multigeneration family through linkage analysis. This approach relies on genetic markers, which help define a genomic region, since their position in the genome is known. Genetic markers lying next to each other tend to segregate during meiosis, and so they are inherited jointly (*i.e.* the likelihood of a crossover taking place within a small DNA fragment is low), allowing tracking of the inheritance pattern of an unknown gene. Genetic linkage is feasible when several affected and unaffected members of the same family tree are available, as this allows for mapping data to be determined. If no relevant mapping information is available, a position independent approach may be used. As *PRPF31* is the gene known

to be mutated in RP11, it is conceivable that gene products involved in PRPF31 pathway(s) can cause retinal degeneration too, such as adRP (RPGRIP is an example; Boylan & Wright, 2000; Roepman *et al.*, 2000a; Dryja *et al.*, 2001). This rationale runs parallel of a reverse genetic approach applied to discover new human RP genes, identifying mutant alleles in genes that are already known but independent from mapping data. In this case, gene products MAGI3 and TRAK2 are dependent on the direct physical interaction with PRPF31. The approach consisted of amplifying selected regions of genomic DNA of adRP patients; the regions corresponded to the exons of the gene. The genetic sequence of these regions was compared to a reference sequence for the detection of mutations. Due to limitation in time, only the *MAGI3* gene was screened, selected because of its localisation with PRPF31 (both nuclear and cytoplasmic in MDCK cells).

7.2 Primer Design

Primers (listed in section A.5 of the Appendix) were designed to amplify genomic regions corresponding to each exon, plus 100-250 bp of the upstream and downstream intronic regions, to identify any intronic mutation which may interfere with splicing. All four isoforms of *MAGI3* were included in the analysis.

7.3 Mutation Screening

A panel of DNAs from 95 patients affected by adRP was used as a sample population for the screening of the *MAGI3* gene. Each patient was representative of a family, for which the most common mutations in adRP genes had already been excluded (such as rhodopsin and *PRPF31*). Biotaq polymerase supplemented with *Pfu* polymerase (for proofreading) was used for amplification. A total of 23 fragments were amplified, purified using ExoSAP-IT and then sequenced (methods described in section 2.9). The sequences were analysed with Seqman software (DNASTAR Lasergene 8), aligning them to the corresponding exonic reference sequence (from Ensembl). The screening allowed the recognition of the known exonic SNPs (eight, according to Ensembl, accessed January 2009), of which one is represented in Figure 7.11 as an example (A/G change: c.2484A>G).

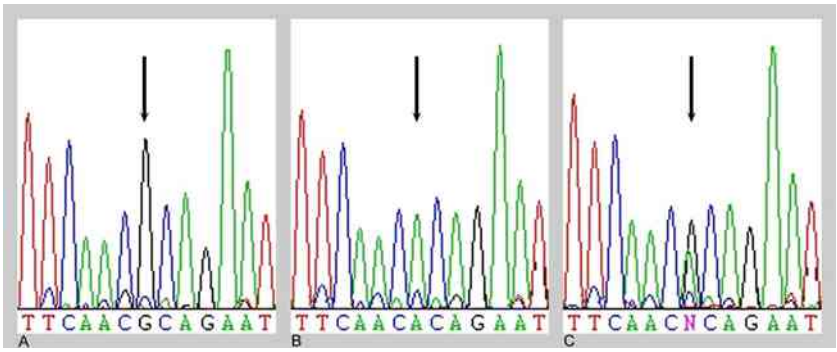


Figure 7.11. Known single nucleotide polymorphism in exon 15 of *MAGI3*.
The panels A, B and C represent a portion of the DNA sequence encompassing the c.2484A>G SNP in exon 15, indicated by a black arrow in each panel. The three sequences belong to three different individuals: two homozygous (G/G and A/A in panels A and B respectively) and one heterozygous (A/G in panel C).

The sequencing of the exons allowed the identification of four novel nucleotide changes (shown in Figure 7.12), in a heterozygous state:

1. c.793G>A, codon 265: non-synonymous change in exon 5 displayed by one patient (Aspartic acid to Asparagine; Figure 7.12, panel A);
2. c.853T>C, codon 285: non-synonymous change in exon 5 displayed by three patients (Tyrosine to Histidine; Figure 7.12, panel B);
3. c.2303G>A, codon 768: non-synonymous change in exon 14 displayed by two patients (Arginine to Histidine; Figure 7.12, panel C);
4. c.4074C>T, codon 1358: synonymous change in exon 21 displayed by one patient (Isoleucine to Isoleucine; Figure 7.12, panel D).

The sequences underwent a second amplification with fresh reagents: changes were re-confirmed after direct sequencing with both forward and reverse primers.

Sequence analysis of the remaining exons did not show any unknown alteration.

The three patients bearing the mutations were Caucasians, therefore 105 DNAs from unaffected Caucasian subjects were screened for the non-synonymous changes. Two subjects out of 105 presented a heterozygous change T>C (c.853T>C, #2 on the list above), suggesting the change may be a single nucleotide polymorphism. For changes #1 and #3, no alteration was observed in the control population. Therefore, the DNAs of the patients' relatives were analysed and sequenced:

- patient RR4691 (change #1): DNAs were available from two affected family members. The relatives' sequences corresponding to exon 5 did not show the change described in patient RR4691;
- patient KOS3396 (change #3): seven affected family members' DNAs were analysed for exon 14, however the change was not detected;
- patient PR3713 (change #3): DNA from one affected relative was analysed for exon 14, however the base-change was not detected.

The analysis suggested that the changes did not segregate within the family.

The comparison of the human MAGI3 protein sequence with other species than human (mouse, rat, pig, cow, gorilla, macaca, dog and chicken) showed that the aminoacids

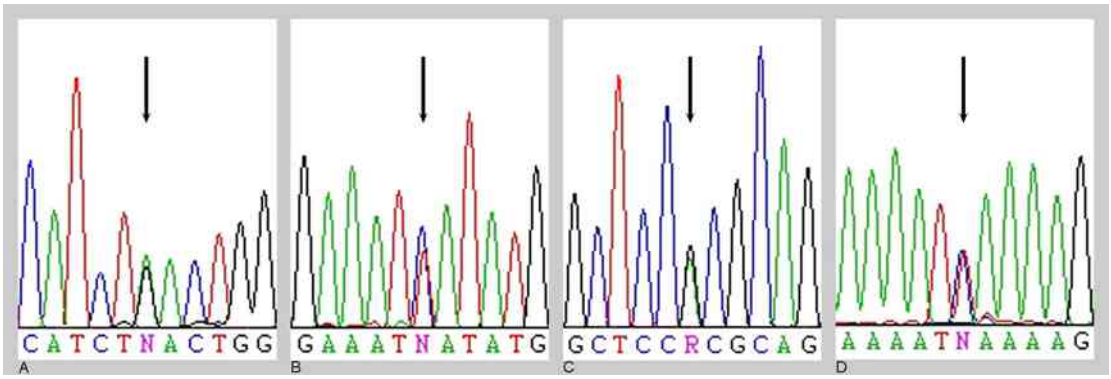


Figure 7.12. Electropherograms showing the nucleotide changes detected in *MAGI3*.
In each panel, the single nucleotide change is indicated by a black arrow. A: c.793G>A; B: c.853T>C; C: c.2303G>A; D: c.4074C>T. Nomenclature based on the RefSeq sequence NM_001142782 (March 2009).

generated by the triplets bearing the nucleotide changes are conserved. Change #1 represent an alteration from a negatively charged aminoacid (Aspartic acid) to a non-charged polar one (Asparagine); change #2 is a modification from a non-charged polar aminoacid (Tyrosine) to a positively charged one (Hystidine). To affirm whether they do not affect the protein folding and/or function, more investigation is needed; however, as the changes were not segregating within the families, we decided not to pursue further the study. Change #3 is instead a substitution of two positively charged aminoacids (Arginine to Hystidine), hence it may be that this change does not affect to a great extent the protein.

Finally, the nucleotide changes can still affect the codon usage during translation, and may perhaps influence the production of the protein.

7.4 Allele and Genotype Frequencies

The allele and genotype frequencies relative to the changes listed in the previous section were calculated according to section 2.8 of Chapter 2. As the changes were believed to be non-causative of adRP, the 95 patients were considered as part of a normal population. For changes #1-3, the population size was 200 individuals (95 patients + 105 controls), for change #4, the population was 95 individuals. The frequencies are summarised in the following tables.

Allele frequencies:

Table 7.7. Summary of allele frequencies for the changes detected in the *MAGI3* gene.

Change #	1.G>A	2.T>C	3.G>A	4.C>T
Allele p	0.9975 (G)	0.9875 (T)	0.995 (G)	0.995 (C)
Allele q	0.0025 (A)	0.0125 (C)	0.005 (A)	0.005 (T)

For each change, the allele is indicated in brackets.

Genotype frequencies:

Table 7.8. Summary of genotype frequencies within the population considered.

Change #	1.G>A	2.T>C	3.G>A	4.C>T
Genotype pp	0.995 (GG)	0.975 (TT)	0.99 (GG)	0.989 (CC)
Genotype qq	0 (AA)	0 (CC)	0 (AA)	0 (TT)
Genotype pq	0.005 (GA)	0.025 (TC)	0.01 (GA)	0.011 (CT)

For each change, the genotype is indicated in brackets following its frequency.

Genotype frequencies according to the Hardy-Weinberg equilibrium:

Table 7.9. Genotype frequencies calculated according to Hardy-Weinberg.

Change #	1.G>A	2.T>C	3.G>A	4.C>T
Genotype pp	0.9950062 (GG)	0.9751562 (TT)	0.990025 (GG)	0.990025 (CC)
Genotype qq	0.0000062 (AA)	0.0001562 (CC)	0.000025 (AA)	0.000025 (TT)
Genotype pq	0.0049874 (GA)	0.0246874 (TC)	0.00995 (GA)	0.00995 (CT)

For each change, the genotype is indicated in brackets.

The real genotype frequencies (Table 7.8) agree with the predicted frequencies (Hardy-Weinberg equilibrium, Table 7.9), inferring that there are no evolutionary mechanisms operating on this locus in the examined population.

7.5 Conclusion

The aim of the study presented in the first section of this chapter was to obtain an overview of the retinal transcripts generated by the interacting partners' genes. The purpose was to investigate whether any alternatively-spliced retina-specific isoform existed, leading to a different protein(s) being translated. For this reason, overlapping DNA fragments were amplified to be aligned to a reference sequence. We opted for this approach as both *MAGI3* and *TRAK2* genes showed a degree of characterisation according to the consulted database (Ensembl). However, no major differences from the deposited sequences were observed.

MAGI3 appears to be expressed principally in two isoforms: they both share the 5' UTR of MAGI-001/MAGI-002 (long 5' UTR), they do not present the extra exon 7a, but they

differ from these two splice variants for the 3' UTR: one shows the 3' UTR of variant 003 and the second shows an alternative exon 21 and a different 3' UTR (as variant 004). Therefore, two transcripts appear to be present, exhibiting a long 5' UTR, and an alternatively spliced exon (#21), generating two proteins which are similar for most of their length, but differ at the C-terminal region.

TRAK2 does not present any splice variant according to both literature and Ensembl database (accessed January 2009); however, it was observed that exon 2 extended towards the 5' upstream sequence, suggesting a different 5'UTR region.

In conclusion, the experiments showed that retina presents splice variants which were already documented or predicted, hence making the initial hypothesis of a retina-specific isoform less probable. Yet, the study does not exclude that a retinal isoform exists, which could not be detected with the chosen methodology. As the aim was to investigate transcript differences affecting the translated sequence, due to time limitations and to the degree of characterisation of the transcripts in the Ensembl database, the variations on the UTR regions for both genes were not investigated further.

The chromosomal localisation of *MAGI3* (chromosome 1, p13.2) and *TRAK2* (chromosome 2, q33.1) was compared to the position of mapped loci for non-syndromic and syndromic RP, however no references were found. Nevertheless, unrelated adRP patients were screened for mutation in the *MAGI3* gene, considering the interaction of MAGI3 with PRPF31 and their co-localisation in epithelial cells (overlap at intercellular junctions, co-localisation in nucleoli). From the analysis of the patients, the controls and the affected family members, it appears that none of the observed changes may be a cause for adRP. Three of the four changes resulted in amino acid changes; however, two out of three mutations were not present in the affected family members, thus ruling out each change as a cause of eye disease. Even though the screening did not allow the discovery of a new RP gene, these changes represent polymorphisms appearing at low frequencies.

8. Discussion and Further Perspectives

Retinitis Pigmentosa 11 (RP11) is caused by mutations in the *PRPF31* gene. *PRPF31* codes for a splicing factor expressed in many tissues (Vithana *et al.*, 2001), nevertheless the effects of the mutations are detrimental only in the retina, confining the phenotype to this tissue.

To explain RP11 aetiology, it was hypothesised that retinal degeneration arises either from 1) haploinsufficiency of the *PRPF31* gene and hence the levels of splicing activity in retina are not sufficient to maintain a healthy tissue, perhaps also influencing the splicing of a gene crucial for retina physiology; or 2) the disruption of an interaction between PRPF31 and a partner in retina; or 3) a different function of PRPF31 performed in retina. The study described in this thesis focused on the second hypothesis, using the yeast two-hybrid approach to screen for potential partners of PRPF31 in the retina.

During the time this research took place, mounting scientific evidence favours haploinsufficiency as the cause of RP11. Haploinsufficiency occurs when only one of the two copies of a gene is functional (with the other inactivated by a mutation, as the case for *PRPF31*), and its protein product cannot suffice the requirements, leading to an abnormal state. Hence it may be that RP11 is caused by an insufficiency in splicing function, and either some or all the proteins of the photoreceptor outer segment are not appropriately replenished. Lymphoblastoid cell lines from asymptomatic and symptomatic patients were used to analyse the expression of the mutant allele: the expression of the disease transcript showed a 50-80% decrease, likely due to nonsense-mediated decay. Such expression was lower than the asymptomatic individuals carrying the same mutations, suggesting the presence of *PRPF31* isoalleles, exhibiting either low or high expression (Rivolta *et al.*, 2006). At the time, the authors did not exclude the possibility of a retina-specific partner interacting with PRPF31 and perhaps with the other splicing-implicated proteins causing RP (PRPF3, PRPF8, PAP1, BRR2). Recently, Rio Frio and co-workers (2008b) showed that the level of total PRPF31 protein was decreased and no truncated protein was detected, suggesting degradation of *PRPF31* mutant transcripts through nonsense-mediated decay. The scientific team also showed

the existence of two genetic modifiers, modulating the expression of non-mutant *PRPF31* alleles: one is associated with the *PRPF31* gene, while the second is located on chromosome 14 (Rio Frio *et al.*, 2008a). Finally a recent work showed evidence that the *PRPF31* gene does not exhibit retina-specific isoforms and is not differentially expressed in retina (Tanackovic & Rivolta, 2009). Taken together, these results suggest that subtle variations of the pre-mRNA splicing process may lead to the development of retinitis pigmentosa in conjunction with mutations in the *PRPF31* gene. However, it is not yet understood whether retina requires a higher threshold of splicing activity than other tissues, threshold not met if one *PRPF31* allele is not expressed; or if the decreased level of PRPF31 does influence splicing activity, but still acceptable for the tissue to be healthy: in this case it may be that a transcript crucial for retina survival is not adequately spliced, and the consequent lower level of protein may lead to degeneration of the retina.

Besides, our research provided new insights into the PRPF31 protein, which will be discussed in the following sections. Experiments presented in Chapters 3 and 4 constitute the core of the project, as they show the identification and the partial validation of novel PRPF31 interacting partners. Chapters 5 support to some degree the findings of the previous chapters (Chapter 5). Chapter 6 provides observations regarding PRPF31 and its localisation in cells and retina. Finally Chapter 7 presents studies regarding the novel partners in terms of genetic screening and isoform analysis.

8.1 Yeast Two-Hybrid and Co-IP Experiments

In the first results chapter (Chapter 3), a screening of human and bovine retinal libraries for novel, direct interacting partners for PRPF31 was described. The methodology consisted of using yeast cells, engineered to facilitate the detection of interacting partners of a desired protein from a library of cDNA clones; this methodology is the yeast two-hybrid system. The approach involved screening the libraries using different baits of PRPF31 (four in total, three short fragments and one full length), to maximise chances of interactions.

The screening allowed the detection of three potential candidate partners from the bovine library, which corresponded to the human proteins MAGI3, TRAK2 and

ZNF143. MAGI3 belongs to the MAGUK (membrane-associated guanylate kinase) family of proteins, which are ubiquitous scaffolding proteins concentrated at cell junctions. MAGI3 encompasses a GK domain, two WW domains and six PDZ domains (Funke *et al.*, 2005) and was first discovered as a partner of PTEN/MMAC (Wu *et al.*, 2000). The role of scaffolding proteins is to organise macromolecular protein complexes, tethering receptors, intracellular signalling enzymes and adhesion molecules; therefore they are involved in specialised cellular adhesion and signal transduction mechanisms. MAGI3 was localised at the tight junctions of epithelial cells (Adamsky *et al.*, 2003). The rat orthologue of TRAK2 was discovered as an interacting partner of the $GABA_A$ receptor (Beck *et al.*, 2002). TRAK2 is believed to take part in membrane trafficking of vesicles and/or organelles, via motor proteins as it interacts with kinesin (a protein involved in anterograde transport). Thus, TRAK2 was proposed as an adaptor protein involved in motor-dependent trafficking of proteins to synapses (Brickley *et al.*, 2005). ZNF143 is a transcription factor, comprising seven zinc finger domains, and it is expressed in all normal adult tissues; it is involved in the transcriptional activation of snRNA and snRNA-type genes transcribed by RNA polymerase II and III (Myslinski *et al.*, 1998).

The next step was the confirmation of the interactions with a second technique. Ideally, all the interactants identified through a technique (such as the yeast two-hybrid) should be confirmed with a second one; however this is not feasible due to time and resources available. Hence we narrowed down according to the results obtained: in yeast, the interaction between PRPF31 and human species of the three interactants favoured MAGI3 and TRAK2, whereas ZNF143 appeared to be a weak interacting partner (section 3.8). As ZNF143 is a nuclear protein, it is more likely to interact with a splicing factor; moreover, the blue colour developed in the β-galactosidase assay by the interaction of ZNF143 with PRPF31 was comparable to the one developed by PRPF6 with PRPF31, symptomatic that perhaps it should have been included for further validation together with MAGI3 and TRAK2. However ZNF143 was not pursued for the following reason: at the time of this choice, the auto-activation test for the isolated prey species had not been done, and in parallel to it the immunoprecipitation experiments were set up. Interactants were then chosen according to which were less likely to bind

DNA spuriously (*i.e.* without the presence of a bait), thus excluding ZNF143. However, after the auto-activation test, it appeared that ZNF143 was not a false positive, but a true interacting partner, and hence it should have been included in the downstream work. Furthermore, a negative auto-activation test does not exclude a candidate from being a false positive: even though MAGI3 and TRAK2 have failed to auto-activate the reporter genes and hence may represent good interacting candidates, it may be that these proteins are to some extent "sticky", and will attach to any bait, consequently triggering the expression of the reporter genes and falling into the false positive category.

Only few potential interacting partners were isolated from the libraries. Several factors could have influenced this outcome:

1. Knowledge and good practice of the methods, necessity to implement the system in our lab;
2. Time-frame allowed;
3. Richness of the libraries;
4. Expression of the bait;
5. Stringency of the system;
6. Type of protein screened for.

Therefore, many factors may influence the process of a yeast two-hybrid screening. On one hand, some factors may relate to how the screening was conducted (points 1 and 2). The fact that the facilities and knowledge were not in our hands, and that we had to implement the system in our lab had played a part. Thus it may be that the methods were involuntarily not performed as appropriate: as an example, a low mating efficiency may be caused by a too vigorous shaking after combination of the yeast mating types, thus inhibiting the mating.

Also, the time-frame allowed for the exchange to learn the technique maybe should have been extended to allow a less tight schedule for the initial, crucial steps, as they constitute the core of a screening (*i.e.* the mating process and the initial selection of positive clones).

On the other hand, the remaining factors relate to the nature of the parts constituting the chosen method. Again, a low mating efficiency can be due to a lower viability of the library cells (as in this case, as they were aliquoted and store frozen at -80°C).

Another factor is that the libraries may not be fully comprehensive, and may not be representative of all the genes: this may hinder them from being singled out (it could be the case of PRPF6 since it was not detected in any screen). The amplification of the libraries (section 3.9) showed that PRPF6 was present in both libraries; however this is debatable as the amplification of the aliquots may have favoured some plasmids against others. Regardless, the number of independent clones screened was more than a million per bait and per library, which is the minimum number recommended in order for a screening to be valid (using less than a million of clones reduces the chance of detecting genuine interactions).

In addition, the stringency of the system was high, which brings the advantage of keeping false positives to a minimum, but on the other hand has the disadvantage of missing interacting partners, especially weak ones. Therefore the screen could be repeated, using less stringent conditions (no selection against the adenine reporter gene), but with the disadvantage of more false positives. In addition, it was later discovered that it is not necessary for all the reporter genes to be expressed to declare a clone is a genuine interacting partner: this is due to the accessibility of the UAS to the recombinant proteins complex, therefore even if the clone elicits the expression of two out of three reporter genes, it can be a true positive and therefore can be considered for further studies. This way, it may appear that many interacting partners can be considered, but care must be taken in deciding which could be true positives and which could be false, and the choice is upon the characteristics of the interactants. On these considerations, the following interactants did not activate the *LacZ* gene after co-transformation but could have been followed-up as they appeared to be in frame with the activation domain: PNMA1 (paraneoplastic Ma antigen 1), USP47 (ubiquitin specific peptidase 47), SV2B (synaptic vesicle glycoprotein 2B). Other clones could be considered too, knowing that yeast has the capacity of skipping termination codons and translating the downstream sequence (yeast tolerates translational frameshift, skipping stop codons and continues the translation).

Another key factor is the type of bait used. We chose to produce four different PRPF31 baits to increase the chances of detecting partners. Not all the baits were expressed at a high level, only BD-F1 was highly expressed (bait encompassing the NOSIC domain of

PRPF31). The crucial experiment for the bait characterisation is the auto-activation, which excluded any spurious activation of the reporter genes (section 3.3.1), therefore allowing for the yeast two-hybrid screenings to take place. Concerning the type of protein, it is difficult to predict whether or not a protein will have plenty of interactions: first because of the aforementioned factors which influence a screening, and secondly because of the nature of the protein. As an example, the PRPF31 NOP domain is involved in contacting RNA and proteins, hence perhaps many PRPF31 interactions occur when the two entities - RNA and protein - are together, such as for protein 15.5K and RNA: PRPF31 binds to both at the same time, but fails to bind them separately (Liu *et al.*, 2007).

In conclusion, we suggest that not many interacting partners were isolated due to the use of high stringency system: in fact, the control interaction of PRPF6 and PRPF31 was remarkably weak (section 3.4), suggesting that the high stringency may have contributed negatively. Makarova and co-workers (2002) reported a tight interaction between PRPF6 and PRPF31; although the authors have used the same methodology, the characteristics of the methods are not comparable: a strong interaction may be a weak one in a different system, and/or vice versa.

Chapter 4 focused on confirming the interactions found through the yeast two-hybrid system, using the co-immunoprecipitation technique. As previously mentioned, the results concerned only MAGI3 and TRAK2, as these two proteins showed the strongest interaction when using human full length sequences to test direct interactions in the yeast two-hybrid system.

We chose to express the proteins with N-terminal tags (6xHis for PRPF31 and V5 for the interactants), to render the experiments more practical, and also to mimic the yeast two-hybrid conditions, as the baits and preys had an N-terminal GAL4 domain. The tagged proteins were highly expressed in mammalian cells (HEK293), thus facilitating the downstream immunoprecipitation work, as the more protein is produced, the greater the likelihood of a good performance of the experiment (*i.e.* detecting a band on a western blot even though there may be a loss of protein due to the experimental procedures).

Both MAGI3 and TRAK2 (as V5-proteins) were separately proven to be part of the 6xHis-PRPF31 protein complex. The authenticity of these results was proven by the fact that PRPF6 was also found complexed with PRPF31 using the same conditions (PRPF6 is a known interacting partner of PRPF31; Makarova *et al.*, 2002).

As PRPF31 binds MAGI3 and TRAK2, it is not excluded that the two proteins are part of the same protein complex. It would be interesting to provide evidence for this, which can be achieved through co-immunoprecipitation experiments, expressing the proteins and use the same procedure described in Chapter 4.

The immunoprecipitation was also performed in the reverse way (coupling the anti-V5 antibody to the support), showing a cross-reaction between the anti-V5 antibody and the 6xHis tag. In fact, the western blot showed that 6xHis-PRPF31 was present on immune pellets of cells expressing V5 tag and 6xHis-PRPF31 (6xHis-PRPF31 was detected by the anti-V5 antibody, confirming that the antibody binds 6xHis-PRPF31 in both western blot and IP procedures; compare with Table 4.2 in Chapter 4). Besides, a band corresponding to the 6xHis tag was not observed, in line with the fact that the tag is very short and hence not detectable on the western blot, but not excluding its presence in the immune pellet. However, no wild type PRPF31 was observed in the immune pellet of cells expressing only the V5 tag, suggesting that PRPF31 did not originate the cross-reaction, and that any spurious interaction occurs when the 6xHis tag is present. On the other hand, this procedure allowed the detection of wild type PRPF31 when the V5-protein was expressed, thus confirming the results of the previous co-IP experiments.

For future work and/or improvements to this system, we would suggest continuing using the V5-tag and the anti-V5 antibody, for their high specificity and robustness; instead we would substitute the 6xHis tag with a different one (HA tag for example), which would give less background and be more easily detectable in immunocytochemistry.

It is to say that the experiments discussed above may have to be supported by co-immunoprecipitation on the endogenous complex, to support the potential biological significance of these interactions: so far, it has been proved that there is association only when overexpressing the proteins with a tag. Hence, the results must be interpreted with caution.

Finally, another approach to investigate the components of the PRPF31 complex would be to utilise mass spectrometry. This method could also provide more insights about the partners in retina if the complex is purified from fresh retina preparations.

8.2 Localisation Experiments

The results from the co-localisation in epithelial cells were described in Chapter 5, in support of MAGI3 and TRAK2 as novel interacting partners of PRPF31.

Unexpectedly, the 6xHis-PRPF31 protein localised not only to the nucleus but also strongly throughout the cytoplasm; it also formed large aggregates in the nucleus which resembled the nucleoli. On the other hand, wild type endogenous PRPF31 was found mainly in the nucleoplasm, particularly in speckles and appeared to be absent from the nucleoli; moreover, it was present in the cytoplasm to a much lower level and also found at cell-cell contacts, supporting in part the pattern of localisation indentified with the 6xHis-PRPF31 protein.

Tagged PRPF31 and MAGI3 appeared to localise together to the nucleus in structures reminiscent of the nucleoli, in forms of aggregates in which PRPF31 lined the outside and MAGI3 the inside. Any cytoplasmic labelling indicative of co-localisation was quite weak, therefore suggesting that in presence of MAGI3, PRPF31 was "compartmentalised" in the nucleus, especially the "nucleoli". The localisation of wild type MAGI3 and PRPF31 in epithelial cells revealed that the proteins partially overlapped at intercellular junctions, indicating that the endogenous proteins may be recruited to a common site (wild type PRPF31 was shown to co-localise with E-cadherin to the adherens junctions). The localisation of endogenous PRPF31 at the junctions supported what was previously shown with the over-expressed PRPF31, that it labelled the cytoplasm (although without clear distinction of subcellular structures). Adamsky and colleagues (2003) reported MAGI3 protein at adherens junctions of aged astrocytes, and also at tight junctions of epithelial cells. Hence it may be that there is a cross-talk between PRPF31 and MAGI3, meaning that either protein shuttles to/from the other protein's location, triggering a downstream effect which is yet to be elucidated. These data also suggest a participation of PRPF31 in MAGI3 pathways at the junctions: it is possible that PRPF31 exerts a modulation of MAGI3 activity, thus altering cell

morphology, adhesion or migration. On the other hand, the "nucleolar" localisation of both proteins is suggestive of other pathways, maybe related to PRPF31 RNA-binding activity. The latter hypothesis is supported by the co-localisation of the recombinant proteins, and also by the co-localisation of the wild type proteins upon hyperosmotic stress treatment of the cells.

Tagged PRPF31 and TRAK2 localised both to the same sites of PRPF31 and MAGI3, the "nucleoli", again in form of aggregates were PRPF31 was detected on the outer side and TRAK2 on the inner side. The two proteins were also observed to localise to the cytoplasm, and the signal was very prominent on cell extensions (on average, half of the cells presented these spike-like structures). It is possible that the over-expression of TRAK2 influences the localisation of PRPF31, hence the latter is transported to these peripheral cell structures; however, this does not exclude that normally a fraction of PRPF31 already resides in the periphery of the cell, but is not detected. In the SK-N-SH cell line (human neuroblastoma), wild type PRPF31 and wild type TRAK2 signals seemed to overlap at the cytoplasmic region close to the nucleus. Regarding cell-cell contact sites, both signals were very faint and could not be thoroughly analysed to understand whether there was any co-localisation; no "spike-like" structures were observed hence suggesting that the over-expression of such protein in epithelial cells induces a change in morphology. These findings suggest a role for TRAK2 in sorting PRPF31 from the nucleus to the cytoplasm, hence explaining TRAK2 presence around the nucleus. However, the fact that both proteins co-localised to the nucleoli, as seen for PRPF31 and MAGI3, is suspicious that maybe the over-expression of the splicing factor triggered a segregation of TRAK2 in the nucleus.

Chapter 6 comprises several observations regarding PRPF31 localisation in epithelial cells. We chose to continue using MDCK cells as they were extensively used for localisation studies of over-expressed proteins (Chapter 4), and showed a peculiar labelling for PRPF31 at the junctions, therefore we investigated this aspect further as PRPF31 was not reported to localise at intercellular junctions. In wild type epithelial cells, PRPF31 localised to the nucleoplasm in speckles, with a very faint to no presence in the nucleoli as the double staining with SC-35 revealed. PRPF31 also weakly labelled the cytoplasm and more interestingly, it was found at cell-cell contacts. We became

more interested in the latter localisation, as PRPF31 was presumed to be nuclear due to its involvement in splicing; however, there are several lines of evidence documenting that splicing can occur in the cytoplasm too (as discussed later). Hence, it was observed that PRPF31 localised to the adherens junctions, as confirmed by co-staining with E-cadherin (a marker for adherens junctions). Moreover, the application of stress treatments did not alter this localisation. Interestingly, giving the cells a hyperosmotic stress enhanced the localisation of PRPF31 to the junctions, and favoured the localisation of the splicing factor to the "nucleoli", where it co-localised with MAGI3. We also investigated the occurrence of PRPF31 at the junctions according to the cell stage, and, as cells in their physiological environment are in G0/G1 phase, we investigated arresting the cells in such phase. The signal was still observed, although its pattern was not as clear as the wild type: perhaps the inhibitors were toxic to the cell and contributed in changing its morphology, or maybe PRPF31 is not as present as in the untreated cells. The latter hypothesis still indicates that in normal proliferative states PRPF31 may be found at intercellular junctions.

Taken together, these results showed that in epithelial cells PRPF31 has a dual localisation (nuclear and junctional) which does not appear to be dramatically affected by the cell stage; it also seems that junctional PRPF31 is more prominent during stress events. These observations may suggest that either splicing extensively occurs at the junctions, or PRPF31 has a moonlighting function, being involved in other cell pathways other than splicing. Besides, the veracity of the observations regarding V5-MAGI3 and 6xHis-PRPF31 "nucleolar" co-localisation is strengthened by the fact that the corresponding wild type proteins still co-localise to the nucleolus in hyperosmotic conditions, therefore suggesting a role of the nucleolus in sequestering these proteins as a complex. However, how this segregation is triggered and what are the downstream effects remain to be elucidated.

The role of the nucleolus is interesting to consider: wild type PRPF31 was not documented to be in the nucleolus (Makarova *et al.*, 2002), thus suggesting to reserve further judgement when considering these observations, especially since no nucleolar marker was concomitantly used. However, PRPF31 has RNA binding activity via its

NOP domain: NOP56/58 proteins are involved in small nucleolar RNP biogenesis (Aittaleb *et al.*, 2003). PRPF31 is known to contact a ribonucleoprotein platform formed by U4 RNA and protein 15.5K; interestingly, protein 15.5K is involved in the assembly of three RNP complexes: the U4/U6 snRNP (splicing), the box C/D snoRNP and the box B/C motif of U3 (involved in processing of rRNA and ribosome respectively). However, Liu and colleagues (2007) showed that PRPF31 contacts only U4 RNA and not other types of RNAs despite the similarity of its NOP domain with Nop56 and Nop58 nucleolar proteins. Nonetheless, when PRPF31 was over-expressed in epithelial cells, it formed aggregates which resembled the nucleoli, hence suggesting that in normal conditions the splicing factor is naturally found in the nucleolus, but not at a detectable level. Moreover, in particular stress conditions PRPF31 was found in the nucleolus. Studies regarding the nucleolar proteome have not pin-pointed PRPF31 though (nucleolar proteome database http://www.lamondlab.com/NOPdb/; Leung *et al.*, 2003). However, the protein SART1 involved in the tri-snRNP assembly (Makarova *et al.*, 2001) was found to be present in the nucleolus (Hinsby *et al.*, 2006); in addition, the protein Prp43p, involved in the removal of U2, U5 and U6 snRNP from the post-spliceosomal complex was purified not only with spliceosomal components but also with preribosomal complexes (Combs *et al.*, 2006). Therefore, the hypothesis of PRPF31 in the nucleolus is strengthened by the fact that the splicing factor not only possesses domains commonly found in nucleolar proteins (NOSIC and NOP domains), but also by the evidence that other spliceosomal proteins are also found in the nucleolus or associated with ribosomal components.

The multifunctional basis of the nucleolus was recently illustrated: it is a subnuclear compartment principally involved in the ribosome-subunit biogenesis, but also involved in coordinating and regulating cell-cycle control events and stress responses by sequestering proteins (Boisvert *et al.*, 2007). It is plausible that cell treatment with a transfecting reagent, and subsequently the overwhelming production of proteins has played a role, with the visible outcome of the protein being sequestered in the nucleolus, instead of being degraded. However, cells subjected to mock transfection appeared not to be affected; also, it seems conceivable that as 6xHis-PRPF31 concentrates in the nucleoli, the V5-MAGI3 and V5-TRAK2 will also do if there is genuine interaction.

Regardless, the fact that each recombinant protein was found together with 6xHis-PRPF31 in the nucleolus is suggestive of a direct interaction between the proteins.

The immunolabelling for PRPF31 in murine retina reinforced the observations from the cell work as it attests that PRPF31 has different localisations, perhaps also in human retina. In fact, in mouse retina PRPF31 was shown to be weakly present in the outer nuclear layer, meaning that splicing in the photoreceptors may not take place in the nucleus (at least at a level detectable by the method), but either takes place in the inner segments or at the synaptic termini; or else, it may be that in photoreceptors, PRPF31 has a different function altogether, for which it does not require to be in the nuclei. However PRPF31 was observed in between the nuclei of the photoreceptor cells, in a streaks-like fashion. The same localisation was shared with PRPF3 and PRPF8, other splicing factors that are part of the tri-snRNP with PRPF31, and mutations in which also cause RP: they were not present in the nuclei but visible as streaks through the outer nuclear layer. On the other hand, another splicing factor, PRPF16, was found to localise to the photoreceptor nuclei, and also very prominently to the inner nuclear layer. These observations suggest that perhaps in the retina the splicing factor PRPF31 (and maybe PRPF3 and PRPF8) are involved in a separate function and/or which occurs in a different subcellular compartment other than the nucleus. This hypothesis is reinforced by the fact that the localisation of PRPF31 in SK-N-SH cells was cytoplasmic as well as nuclear, hence suggesting that PRPF31 may have a different behaviour in some neuronal cells than epithelial cells.

Interestingly, a recent publication (Solovei *et al.*, 2009) pointed out that the chromatin pattern in mouse rod nuclei is inverted according to the canonical pattern (euchromatin at the centre of the nucleus, heterochromatin at the periphery and interspersed in the nucleoplasm). This inverted nuclear architecture is common to all the nocturnal mammals analysed by the authors, whereas the conventional arrangement is observed in pig and horse; cow showed an intermediate pattern. Given these considerations, it is plausible to reason that splicing occurs at the periphery of the (mouse) rod nuclei, hence supporting the observations about PRPF31 (and PRPF3, PRPF8) presence in streaks in

between the photoreceptor nuclei. Nonetheless, this does not explain why the splicing factor(s) was (were) found at the inner segment and at the outer plexiform layer.

So far, there is mounting evidence suggestive of a cytoplasmic localisation of PRPF31, thus indicating either a new function, or occurrence of splicing also outside the nucleus. Both options suit the type of interactants found. Interestingly, there are investigations which attest for cytoplasmic splicing. Firstly it was suggested by Denis and co-workers (2005), who put forward this hypothesis by showing that human platelets contain spliceosome factors and they are able to process resident pre-mRNAs in response to external cues. Later it was also shown that neuronal dendrites are capable of splicing reporter pre-mRNA, but the authors failed to provide a functional significance of this dendritic splicing (Glanzer *et al.*, 2005). More recent scientific data provided insights into a cytoplasmic segregation of the minor spliceosome (see paragraph 1.5.1 of the Introduction; König *et al.*, 2007), and PRPF31 is known to belong to the protein machinery of the minor tri-snRNP (Will & Luhrmann 2005). Conversely, the findings from König and co-workers are opposed by those of Frilander's group (Pessa *et al.*, 2008): they state that the minor spliceosome components are localised predominantly to the nucleus.

Considering the spatial organisation of the retina, it could be realistic that splicing occurs outside the nucleus in photoreceptor cells. It is known that mRNAs are trafficked to dendrites and that their translation is regulated according to neuronal activity; such mRNAs are transported in form or large granules, which contain mRNAs, RNA-binding proteins, ribosomes and translational factors, and the transport is microtubule dependent (reviewed by Martin & Zukin 2006). However, how unspliced or partially spliced transcripts can evade surveillance and be exported outside the photoreceptor nucleus remains elusive, if at all possible.

Assuming such pre-mRNA translocation happens, it may be that PRPF31 is part of the pre-mRNA-containing granule which escapes the nucleus, hence explaining why PRPF31 is found to be cytoplasmic, especially in retina (inner segment and outer plexiform layer). That could explain why a protein like TRAK2 interacts with PRPF31: it may contact the complex via PRPF31 and carry it along microtubules through the cell.

This may be plausible since TRAK2 has been found to be involved in the axonal transport of cargoes (Brickley *et al.*, 2005): interestingly, in the retina PRPF31 is found at the synaptic termini of photoreceptor cells, where also TRAK2 was observed (unfortunately, no co-localisation could be performed, see Chapter 6).

It is known that the minor spliceosome has a lower processing rate than the major counterpart (Pessa *et al.*, 2006), and that minor introns are not as common as the major ones - representing only 0.15-0.34% of all introns. Minor splicing seems to escape mitotic downregulation, and it may be linked to cell proliferation (König *et al.*, 2007). Considering that the minor splicing has a lower rate of processing the minor introns, and also that PRPF31 takes part in both spliceosomes, it may be that a decreased rate of processing due to lack of PRPF31 affects the viability of minor-intron-containing transcripts crucial for photoreceptor survival. Noteworthy, PrBP/δ contains a minor intron (as predicted by database found at http://genome.imim.es/datasets/u12/; Alioto, 2007). PrBP/δ was formerly known as PDE6D and was believed to be a subunit of rod PDE6 (the phosphodiesterase which catalyses cGMP hydrolysis in the phototransduction cascade), but later it was found to play a role in intracellular trafficking of newly synthesised prenylated proteins, such as PDE6 and GRK1 (rhodopsin kinase, shared between rods and cones). However PrBP/δ is expressed in other tissues which do not express PDE6 (Norton *et al.*, 2005); nonetheless, it was reported that PrBP/δ is present in retina at higher levels than other tissues. Knockout of the *Pde6d* gene in mouse resulted in an adult mouse of significantly reduced body size but of normal viability, development and fertility, but with a severely compromised rod and cone photoreceptor physiology (Karan *et al.*, 2008). GRK1, rod PDE6 and cone PDE6α' showed a defective transport to the outer segment due to absence of PrBP/δ; the *Pde6d*$^{-/-}$ rod outer segments lack GRK1 and show a minor mistargeting of PDE6, showing a delay in dark-adaptation; also cGMP may rise to toxic levels, an event linked to various retinal degeneration (*rd* mouse as an example; Karan *et al.*, 2008).
Given these circumstances, it is possible that a limited quantity of PrBP/δ affects PDE6 trafficking to the outer segments, leading to impaired phototransduction first (as there is no stoichiometric quantity of the cascade components) and high levels of cGMP, thus

leading to degeneration. The same mechanism could be evoked also by the other splicing factors causing RP, as they all take part in the minor spliceosome. It is known that rhodopsin, which constitute the 90% of the proteins in the rod outer segments (Palczewski, 2006), is produced in only two hours (von Schantz *et al.*, 1999), hence it is likely that its splicing is also happening in within two hours and thus it must be efficient and any impairment slowing it may lead to death of the rod cells in the long run, possibly by apoptosis. However, it is possible that instead splicing happens within 24 hours, and partially-spliced transcripts are kept on site (such as the inner segments) and ready to be ultimately spliced (explaining the splicing factors' presence to the inner segments) and then translated in those two hours.

Regarding MAGI3, we postulated before a role in the nucleolus together with PRPF31 in particular conditions, such as stress events; we also postulated a role of PRPF31 in signalling at the junctions, modulating the role of MAGI3 in cell-cell adhesion and morphology. However, it is hard to speculate what is happening in the retina as MAGI3 seemed to co-localise with PRPF31 only in few cells within the inner nuclear layer, and to the area corresponding to the outer plexiform layer. Retina cells are not epithelial cells but neurons, thus they will not form typical tight junctions, however they still may form other types of junctions: as MAGI3 is known to localise to adherens junctions in astrocytes (Adamsky *et al.*, 2003), it is possible that the neuronal cells of the retina form some type of junctions/connections in which MAGI3 is localised, explaining the round pattern observed (see magnification on Figure 6.16 of Chapter 6).
On the other hand, interestingly MAGI3 has two WW domains. WW domains are ~40 residues long and characterised by two strictly spaced tryptophan residues, and they fold as a triple-stranded β-sheet, they recognise and interact with proline-rich regions mediating intracellular processes (Macias *et al.*, 2002). WW domains of CA150 and FBP11 (spliceosomal proteins) were found to be associated with the early spliceosome, particularly with the U2 snRNP and with less affinity to the other snRNPs (Lin *et al.*, 2004); the authors postulated that WW domain-containing proteins have the ability to nucleate the assembly of splicing factors at the 3' splice site of the intron facilitating splicing or exon selection. It may be possible that MAGI3 exerts a similar role, by

interacting with a splicing factor on U2 snRNP via its WW domains, and contacting PRPF31 which is part of the tri-snRNP, facilitating the transition of the spliceosome complex A (prespliceosome) to complex B (precatalytic spliceosome). This could perhaps happen at the junctions of epithelial cells. It is known of Symplekin, a protein involved in 3'-mRNA processing and polyadenylation, having a dual localisation at the tight junctions and at the nucleus in epithelial cells (Keon *et al.*, 1996; Hofmann *et al.*, 2002). Symplekin interacts with ZONAB (Kavanagh *et al.*, 2006), which is a transcription factor which also localises to both tight junctions and the nucleus, hence Symplekin seems to regulate gene expression not only through its polyadenylation activity but also through the transcription factor ZONAB. It is possible to make a parallel hypothesis with PRPF31, as it is found at adherens junctions and in the nucleus simultaneously, hence it may similarly be involved in regulating gene expression which escapes nuclear regulation to facilitate more prompt expression according to local signals, in concert with MAGI3.

TRAK2 is probably involved in the transport of mitochondria via kinesin to the synaptic termini (Brickley *et al.*, 2005; Smith *et al.*, 2006): nerve terminals have a high metabolic requirement for ATP. This could imply that there is a mechanism by which the cells are able to couple the splicing of transcripts to the levels of ATP or vice versa, explaining the interaction between PRPF31 and TRAK2; the reason for this may be that splicing requires ATP, hence when it is at its highest rate of splicing, PRPF31 is the link between ATP supply via TRAK2, which would bring mitochondria on site.
The TRAK2 homologue, TRAK1, was found associated *in vivo* with RNA Polymerase II and OGT (an enzyme responsible for the *O*-GlcNAc –acetylglucosamine– post-translational modification), leading to speculation that TRAK1 may target OGT to transcriptional complexes (Iyer *et al.*, 2003). RNA Polymerase II was shown to be extensively *O*-GlcNAc-modified, hence suggesting an interaction with TRAK1-OGT complex, which has a nuclear localisation. TRAK2 does not seem to localise to the nucleus, however it does interact with OGT, hence we hypothesised that TRAK2, being present at the border of the photoreceptor nuclei (see Figure 6.17, Chapter 6) may contact RNA Polymerase II in transcriptionally active granules via OGT (as mouse

photoreceptor nuclei have an inverted chromatin pattern, see before), and link them to PRPF31 (splicing can occur co-transcriptionally; Neugebauer, 2002). The reason for this is not known, although it could be a mechanism of transcription/splicing regulation, as it may be that TRAK2 targets OGT to transcription factors and spliceosomal proteins for them to be glycosylated.

Iyer and colleagues (2003) also reported that TRAK2 homodimerises via its coiled-coil domain and they reasoned that it is possible that TRAK2 and TRAK1 may interact with each other (association proven *in vitro*, but authors did not provide information whether this interaction occurs physiologically) hence there may be a cross-talk, explaining the nuclear (nucleolar) presence of TRAK2 upon over-expression.

The C-terminal domain (CTD) of RNA Polymerase II can be glycosylated; this modification was proposed to regulate transcription (Comer & Hart, 2001). Therefore it may be that there is a cross-talk between PRPF31, TRAK2 and the CTD to regulate transcription according to splicing. This idea of a linker between the polymerase and splicing is also reinforced by the fact that the WW domain-containing protein CA150 mentioned before interacts with CTD (Lin *et al.*, 2004). Similarly, MAGI3 can be viewed as a regulator of transcription and splicing in epithelial cells, perhaps not in mouse photoreceptor nuclei.

8.3 Isoforms and Mutation Screening

Chapter 7 presented the results from the isoform analysis of MAGI3 and TRAK2, and the mutation screening regarding *MAGI3*.

As we were interested in alternatively-spliced retina transcripts which result in new MAGI3 and TRAK2 proteins, we pursued the amplification of the transcripts in retina. The analysis of the results from the amplification of the transcripts coding regions suggested that neither MAGI3 nor TRAK2 may have retina-specific isoforms (which could therefore perhaps relate to RP11). Nevertheless, MAGI3 was found to be present in two different species, 1481 and 1125 amino acids long, meaning that MAGI3 in human retina has an isoform of ~170 kDa, and a second one of ~130 kDa. Database searches did not identify any relevant domain in these C-terminal regions. TRAK2 was found to have only one transcript, generating a protein of 914 amino acids (~105 kDa).

Interestingly, both genes generated 5' UTR which differed from the sequence deposited in Ensembl. EST database searches focussed on these regions did support a longer 5' UTR for *MAGI3*, whereas for *TRAK2* no such longer 5' UTR was retrieved. These results suggested that perhaps the transcripts undergo different types of regulation in the retina. Hence it would be interesting to investigate their 5' UTRs in various tissues for comparison, however, the research was not pursued further as we were interested in the protein product of these transcripts.

The approach aimed to encompass the types of isoform which shared the annealing regions of the primers used: if a shorter/longer isoform was present, it would have been detected because the fragments were overlapping by two exons, facilitating the detection of missing exons or extra-exons. However, if the new isoform did not present the region for the primers to anneal, then no amplification would have been observed. Therefore the presence of a retinal isoform is not entirely excluded, and it could be identified with a different method, rapid amplification of cDNA ends (RACE) for example; however, that method was not chosen due to the level of characterisation of the transcripts in the Ensembl database. Other ways to detect isoforms are northern blotting and western blotting, done on human retina RNA and proteins respectively. The northern blot approach will rely on the detection of RNA, hence a different migration pattern in the gel would be suggestive of different splicing products (provided that the probe is annealing to all the species); on the other hand, western blot is quite similar to the northern blot, as it will allow the detection of protein bands migrating at different molecular weights (also in this case the epitope recognised by the antibody should be present in all the protein species). However, the latter case is not feasible for MAGI3 as it already displays isoforms and it would be difficult to detect any isoform of a similar molecular weight.

Alternatively, a human EST database search for the entire transcript sequence would be a useful tool to discover any new transcribed region(s), however it has the disadvantage of being time-consuming, and is preferred to corroborate experimental data.

Regarding the mutation screening, 95 patients affected by autosomal dominant RP (with no mutations in most common genes) were subjected to sequence analysis for the *MAGI3* gene, with the aim to discover mutations that could perhaps lead to retinal

degeneration. We chose to screen the *MAGI3* gene by virtue of MAGI3 nuclear/nucleolar localisation with PRPF31 and its WW domains, hence it may be favoured to have a role in splicing. Four single nucleotide changes were detected, three were non-synonymous, meaning that the translated sequence bore one amino acid change. The presence of each non-synonymous change was investigated in 105 controls, two of which showed one change. For the remaining two changes, the DNA of affected family members for each patient was sequence analysed, resulting in non-segregation of the mutation. Therefore it is possible to infer that the observed changes are rare single nucleotide polymorphisms.

Conversely, this screening should not be regarded as conclusive and exhaustive, as it has not entirely excluded *MAGI3* from being a gene implicated in retinitis pigmentosa: the gene may still cause RP, but assuming its frequency is very low, it may be that 95 patients are not sufficient to detect it. Furthermore, any large deletions/insertions or rearrangements will be missed by direct sequencing method. Also, it may be possible that the change(s) lie(s) in intronic regions, which were not sequenced. *TRAK2* was not screened due to time constraints, but also (as demonstrated by screening *MAGI3*) the chances of finding a mutation in this manner ("candidate gene" approach) are relatively slim, as there is no mapping data from a family to refer back to and therefore no indication about the mode of inheritance of the potential mutation. Subsequently it is difficult to ascertain which type of patients to screen. Moreover, using a "candidate gene" approach to screening does not imply that mutations in the particular candidate gene cause retinal degeneration.

8.4 Final Considerations and Future Perspectives

Upon the results observed and discussed, it is hard to postulate an explanation, linking mutations in the *PRPF31* gene and retinitis pigmentosa, different from haploinsufficiency: first because of the mounting evidence in literature favouring the haploinsufficiency theory; secondly because this project unveiled unexpected insights into PRPF31. The results presented herein are novel and relate to a more general background of PRPF31 function rather than specifically to the retina and/or retinitis pigmentosa. In fact, before postulating any theory other than that of haploinsufficiency

to explain RP, more information must be gathered on the role of PRPF31 in the cell, as the results presented herein suggest PRPF31 may play a larger role.

From this work several hypotheses can be put forward:

1. The process of splicing can undergo several modes of regulation. One of these may be initiated by stress, and it may target the spliceosome, or some of its components, to the nucleolus. Both MAGI3 and TRAK2 may influence this process, since they were found co-localising to the nucleoli with PRPF31; it does not exclude that PRPF31 may take part in processing of rRNA;
2. Implication of MAGI3 in splicing, suggested by its WW domains. It may be that MAGI3 plays a role in linking the U2 snRNP to the tri-snRNP contacting PRPF31; this may happen in both the nucleus and the junctions according to regulation events;
3. Similar to the previous hypothesis, TRAK2 may be implicated in regulating splicing by coupling splicing to transcription: if TRAK2 is able to control the C-terminal domain (CTD) of the RNA Polymerase II through OGT, it may also contact PRPF31 regulating the splicing process of the nascent transcript. This hypothesis would suit more the (mouse) photoreceptor nuclei architecture rather than cells displaying a standard chromatin arrangement.
4. In the case of cytoplasmic splicing, it is possible that the mitochondrial presence at particular sites is connected to splicing rates through a pathway implying an interaction between two key players: PRPF31 for splicing and TRAK2 for trafficking of mitochondria;
5. Again for cytoplasmic splicing, it is possible that TRAK2 contacts RNA granules through PRPF31 and sort them to different cellular destinations via kinesin and microtubules;
6. Regarding RP, it may be that transcripts crucial for rod survival contain minor-introns and their splicing is slowed down by the absence of half the normal PRPF31 protein product, triggering cell death and apoptosis in the long run; this could be as the case of the PrBP/δ protein as less protein would lead to mistargeting of PDE6 and absence of GRK1 in the outer segments, impairing phototransduction;

The hypothesis of PRPF31 having a moonlighting function is fascinating, however considering the domains the protein bears it is likely to be a function in connection with RNA processing or regulation of pathways through RNA. Another interesting alternative would be to further investigate the relation with ZNF143: the hypothesis that can be put forward (in case of genuine interaction with ZNF143) is that PRPF31 is able to alter the transcription of U6 through ZNF143 in case U6 snRNA is needed: ZNF143 is known to target several snRNA genes for transcription (Myslinski *et al.*, 1998) and activates U6 transcription (Yuan *et al.*, 2007).

Due to the evidence gathered in the last two years about the mechanism of RP11, it is likely that the disease is caused by an insufficient amount of splicing factor, meaning haploinsufficiency. This work focuses more on PRPF31 as a key component of splicing, and does not fulfil the questions raised by RP11. Therefore the future work that we propose is directed towards answering questions regarding the biology of PRPF31 and splicing, and the correlation between the splicing factors and the new interactors, MAGI3 and TRAK2.

It is of great interest to investigate the role of PRPF31 and MAGI3 and TRAK2 to confirm any splicing regulation due to these interacting partners or to confirm if/how PRPF31 plays a role in their pathways. Live cell imaging could be used to validate whether the interactions are transient and where in the cell they happen. Moreover we would propose to knockdown one of the interacting partners in cells and follow up the effects on PRPF31, and vice versa (perhaps with a partial knockdown of PRPF31, as the *Prpf31*$^{-/-}$ animal model was not viable, suggesting that a complete lack of the splicing factor is embryonic lethal; Bujakowska *et al.*, 2009). We would suggest performing experiments involving stress and to detect the effects on PRPF31 and perhaps on splicing; also upon these treatments, it would be of interest to investigate the nucleolar interaction of MAGI3 and PRPF31. Experiments involving stress, perhaps using a stably transfected cell line for a partial knock down of PRPF31, could mimic the conditions in retina (a highly stressed tissue), which may give better indications linking to retinitis pigmentosa.

Nonetheless, the validation of the ZNF143 interaction with a different method is reccommended, and, if positive, pursue co-localisation in cells and retina, as done for MAGI3 and TRAK2 with a comprehensive follow up with other experiments as mentioned above. Besides, a further analysis of the other isolated yeast two-hybrid clones is advised, in case any interesting protein (perhaps retina-specific) escaped our analysis, due to the high stringency.

Moreover, to investigate the complex of protein surrounding PRPF31 we would suggest co-immunoprecipitation experiments: first, to check whether or not PRPF31 and E-cadherin complex together (using MDCK cells); secondly, to analyse the protein complex from cytosolic and nuclear fractions using mass spectrometry (this would be of interest to understand whether or not the spliceosomal complex is present in cytosol and with which components); third, to perform mass spectrometry of PRPF31 protein complex, isolated from the retina.

Also we would suggest comparing the splicing rates of retina (perhaps isolating each layer) and another tissue which is known to have a high metabolism (pancreas or liver as an example); also comparing the immunohistochemistry of retina with another tissue, to confirm whether or not the presence of PRPF31 outside the nuclei is specific to the retina (and therefore suggesting a link to RP) or whether PRPF31 when investigated in the context of a tissue displays a different localisation than in cultured cells.

Finally, it would be worthwhile to investigate the presence of other splicing factors, in both retina and cells, to observe their subcellular localisation. This should give more information on splicing: why three components of a spliceosome particle are causing RP and are all localising in an unexpected way (PAP1 and BRR2 were not tested), whereas PRPF16, not involved in the tri-snRNP assembly, shows a localisation more in line with the expectations? It may be that either the tri-snRNP can be assembled outside the nucleus, supporting cytoplasmic splicing of transcripts which require immediate translation “on-site”. However, it is rather controversial that a partially spliced transcript reaches a subcellular region with less difficulties of a completely spliced transcript, which would be shorter hence easier to translocate. It may be that a pre-mRNA transcript folds and is then sorted to the extranuclear site. It appears that in retina the splicing machinery is translocated outside the nucleus, favouring the hypothesis of

splicing happening in the minute cytoplasmic area bordering the nucleus, as the SC-35 labelling suggested (Solovei *et al.*, 2009). Regardless, more experimental data must be gathered before the patho-physiology of PRPF31 mutations can be determined.

Appendix

A.1 Oligonucleotides for Expression Profiles

Table A.1. List of the oligonucleotides used for the expression profiles on *MAGI3*, *TRAK2* and *ZNF143*.

Name	Sequence	Size bp	Mg	T_a
MAGI3 F17	CATCATGGTCCACCATCAGGAA	264/844	2mM	64
MAGI3 R18	ATTGTGCCGACTGCCTACAACT			
TRAK2 F11	GAGATTGAGGGGACTATGCGTA	247/544	2mM	64
TRAK2 R12	TCCCCTGGTTCAGGAGTGATTT			
ZNF143 F4	CTCATAGATGGCCAGGTCATTC	157/682	2mM	64
ZNF143 R5	TGGAGGTGTGGTGAATAAATGC			

The "size" column lists both the cDNA and the genomic DNA sizes of the fragments amplified by the corresponding primer pair.

A.2 Oligonucleotides for Library Amplification.

Table A.2. List of the oligonucleotides used for the amplification of *PRPF6* fragments from the human retina and the bovine retina yeast two-hybrid libraries

Name	Sequence	Size bp	Mg	T_a
PRP6_F7	GTGTCTGACTCCGTGAGTGGACAG	415	2mM	66
PRP6_R9	CTTTCCGAAGAACCCGCTTCTTTG			
PRP6_F13	TGTGTAGCCCACAATGCCCTGGAG	582	2mM	59
PRP6_R17	TGTGGAGTGGGGACACTTCTTCAA			
PRP6_F17	TTGAAGAAGTGTCCCCACTCCACA	622	2mM	68
PRP6_R21	TCAGAAGGTGTTCTTGATGCGGCC			

A.3 Oligonucleotides for *MAGI3* Splice Variants

Table A.3. List of the oligonucleotides used for the amplification of *MAGI3* in retina.

Name	Sequence	Mg	T_a
5' UTR			
A	GAGGAGGAGGAGGTTGCGGCTGG	1mM	60
B	CCTTTCCCGCCGCTCGGGCGCTGAGCC	1mM	60
C	AGTCCGGTCAGCCCCGGGGAGCGCAGC	1mM	60
D	GCCGCCCAGGGCCCCCGGGCTGAGACG	1mM	60
E	CACTCCTGCACCTTGCTGAGCCAGTGC	1mM	60
Coding Sequence			
1_F	ATGTCGAAGACGCTGAAGAAGAAGAAGC	1mM	60
4_R	GCCTCTCTCTGTTTTCTGCGTTTCCACTGC		
2_F	AGACTGTGAAACCAGGCAAAGTCATTAATA	1mM	60
8_R	GTGATTTTTCCATATCTGGTTTTGAAGACC		
6_F	GATCTACTTCATTGACCACAATACCAAGAC	1mM	60
10_R	TTGGTGAAGGAGGACCTCCTCTTAAGATAA		
8_F	GACATACTATGTTGATCACCTTAACCAGAA	1mM	60
12_R	CTTTGGTGTTTGGTGGTTTATCTTCATATA		
10_F	GGAAAATTGCACCAGGCGATGTTATTGTAG	1mM	60
14_R	CCTCGGGTTGTTTTTCTCCATAGAAGATCT		
12_F	AAAACTGCCAAAATGAAAACAGATAAAAAG	1mM	60
17_R	CCTTCTAGGGCACTTCTGTCCCCAGGTATG		
15_F	AGATCTTCTATGGAGAAAAACAACCCGAGG	1mM	60
21_R	CACGTAGGAGCACAGACCGGAAGGAGCC		
20_F	GATGGCAGAATTCATGTTGGTGACCAGATT	1mM	60
21_R	GAGAAGAGACTGATGCTTTATCTCAGGC		
3' UTR			
F	CCTACACAAGGAATCACACATACTCGAGC	1mM	60
G	CACGTAGGAGCACAGACCGGAAGGAGCC	1mM	60
H	CCCGAGCAACATTAAGGCTTTCAGGGC	1mM	60
I	GAGAAGAGACTGATGCTTTATCTCAGGC	1mM	60

Sizes of various generated fragments are discussed in chapter 5.

A.4 Oligonucleotides for *TRAK2* Splice Variants

Table A.4. List of the oligonucleotides used for the amplification of *TRAK2* in retina.

Name	Sequence	Mg	T_a
5'UTR			
A	CGCTGGGAACGCATTCGCCCCACCCGCGC	1mM	60
B	CCGAGGTCGCCGAGTGATGATGTTGTG	1mM	60
C	CCTTAGTAATTCCTCCCTTAACTCCAGGC	1mM	60
D	GTCATGACTGTCCAAAGTATGATAATCAC	1mM	60
E	CTCTGATGAGTCAAATGACATCCGTAGC	1mM	60
F	CAGTGATGCTCTCCGAGTCTCTGTGATTGC	1mM	60
M	CGGACGACAGAGGATGCCGAACCACTCC	1mM	60
Coding Sequence			
1_F	CCGAGGTCGCCGAGTGATGATGTTGTG	1mM	60
8_R	GTTTGTGCTG AAGGTCTACA ATCTGTGAC		
6_F	GTTAATCAGCTGCAGCATGAGCTATGC	1mM	60
12_R	AACCTCCTCT GAGCTGCTCC CCTGGTTC		
10_F	CTGCACGAGTTACAAGACAGGAATATGG	1mM	60
15_R	CTGGGGGTAA CCTGAGTGAT GTCAGAGG		
13_F	GAGCTCCCAGAAGATGGGCCAACCAGGAC	1mM	60
16_R	CTGCCTCCTGGCATCTTCCATTCTCTTGG		
3'UTR			
G	TTGCAACCGCCAACCCAGGAAAGTGCCTG	1mM	60
H	GGTTCCCTTCATTATCCTGTGGAAGTAGCG	1mM	60
I	CAAGCTACTCATGTTTATTAACTGATCTAC	1mM	60
L	GAATGGGCCATGAGTGTTCAGAGGGAG	1mM	60

Sizes of various generated fragments are discussed in chapter 5.

A.5 Oligonucleotides for *MAGI3* Screening

Table A.5. List of the oligonucleotides used for the amplification of *MAGI3* exons for the mutation screening of adRP patients.

name	Sequence (5'…3')	Size	Mg	T_a	DMSO
MAGI3_F1	GAGGAGGAGGAGGTTGCGGCTGG	967	2mM	64	10%
MAGI3_R1	GGACAATACG CGGTGGGGAG GTG				
MAGI3_F2	CTGGTTTAAGAGTATTCAGTTACTAGC	309	2mM	56	-
MAGI3_R2	TTCATGTGAAATACTCCATCAAACTATC				
MAGI3_F3	GTGAATTTTACCTTAATAAAGTTGG	508	2mM	56	-
MAGI3_R3	GTCTGGCTGA TTTTTAGTAT ACTG				
MAGI3_F4	CTCAAATGATTCACCCATCTCGAC	583	2mM	56	-
MAGI3_R4	CTCACTTTAG CTTATAGGAT GGTC				
MAGI3_F5	GACTAGCATCAGAGGTGCTATACAGAG	639	2mM	56	-
MAGI3_R5	GATCCTCCTG CTCTTATAAA ACCAGAAG				
MAGI3_F6	ATCTGAGGAATACAAATGGAGCCTGTTG	364	2mM	56	-
MAGI3_R6	ATCAGTTGAT TATGATATCC ACCCACTG				
MAGI3_F7	CATCTTCTGTTCTGGGCTTCTGATTGC	602	2mM	56	-
MAGI3_R7	CCTTGATAAT TCATCTAGTC TGTTGCCC				
MAGI3_F7a	TGGGTGACAGAATGAGTGCCTGTCTCC	487	2mM	68	-
MAGI3_R7a	GTGCAGTGGC TCACGCCTGG AATCTC				
MAGI3_F8	GCTTCACAGCTTGGCATCTTCTGCA	426	2mM	64	-
MAGI3_R8	ACAGTAACAA CTAGACACTC TTTAGTGTC				
MAGI3_F9	GACTTCATCTTCCTGGTATAGTACTTTGC	757	2mM	56	-
MAGI3_R9	CAGATAAATG CTTGAGGAGA TGTATACCAC				
MAGI3_F10	CTATCTTTTCAGCAGTTAATAGAGGTTTCC	882	2mM	60	-
MAGI3_R10	TCAGACAGAA CCTTAAGGAA AATGCACAC				
MAGI3_F11	CAAAGACACCAAAGTTCAATGTCACACAGG	580	2mM	60	-
MAGI3_R11	TGGTACAATA GGCCTTTAGT AATCCTCCAG				
MAGI3_F12	GTAAATCTGCTGTGTCTTTGATCCTGCC	382	2mM	60	-
MAGI3_R12	GCTTCTGAAA GATGGGAAGG TCCTAGTC				
MAGI3_F13	GTTGAATTGTGTGTGTGTGTTACACTC	483	2mM	64	-
MAGI3_R13	ACTGTCTAAT TTGCAGCTGA CATATTC				
MAGI3_F14	TGCCATTTTGGAGAATGGGCACTTTCTC	531	2mM	64	-
MAGI3_R14	ATCCTTTTGT GTTTGCATAG CCTTCCAC				
MAGI3_F15	ACTGCTGTTGCCTACATCTGCACTCAG	558	2mM	64	-
MAGI3_R15	AGATAAACTT GAAGCCTGGA AGGTTCGG				
MAGI3_F16	GTTTGGATGCGATGATGACGTGTGGTAC	475	2mM	64	-
MAGI3_R16	GAAGAATAAG ACTCCAGCGT GCATACAC				

MAGI3_F17	GTATAGTTCCTGGAGCTCTGTCTTCTAG	587	2mM	64	-
MAGI3_R17	GGACGGTATT TGAGGAACAG GCAAGAG				
MAGI3_F18	TGTCAGAGGGAAAATGGCAAGGTCGAG	607	2mM	60	-
MAGI3_R18	GTCAATGTGC AATGCAACAA TCAGAGAAC				
MAGI3_F19	GTTCTCTGATTGTTGCATTGCACATTGACC	440	2mM	60	-
MAGI3_R19	CGAAGTTTCT AGTCTAGAGC CTGAACC				
MAGI3_F20	CACATGGCAGTACTGGGTACACGTTAGG	491	2mM	60	-
MAGI3_R20	CATTCCATTC CAACGAAAGT CCTGCCTAAG				
MAGI3_F21	ACACCCACGCAGTACTTGCCTCATC	882	2mM	64	-
MAGI3_R21	AGGAACAATA CGTTTTGAAC TTTGACAGC				
MAGI3_F21a	ACTCAGGCTTTTTAAGCAGTTACGAC	1286	2mM	64	-
MAGI3_R21a	AATGCTGCTG TAAGAAAAGA TTACAAC				

Bibliography

Abd El-Aziz MM, Barragan I, O'Driscoll CA, Goodstadt L, Prigmore E, Borrego S, Mena M, Pieras JI, El-Ashry MF, Safieh LA, Shah A, Cheetham ME, Carter NP, Chakarova C, Ponting CF, Bhattacharya SS, Antinolo G. 2008. *EYS*, encoding an ortholog of *Drosophila* spacemaker, is mutated in autosomal recessive retinitis pigmentosa. *Nat Genet* 40:1285-1287.

Abu-Safieh L, Vithana EN, Mantel I, Holder GE, Pelosini L, Bird AC, Bhattacharya SS. 2006. A large deletion in the adRP gene PRPF31: evidence that haploinsufficiency is the cause of disease. *Mol Vis 12*:384-388.

Adamsky K, Arnold K, Sabanay H, Peles E. 2003. Junctional protein MAGI-3 interacts with receptor tyrosine phosphatase beta (RPTP beta) and tyrosine-phosphorylated proteins. *J Cell Sci 116*:1279-1289.

Aittaleb M, Rashid R, Chen Q, Palmer JR, Daniels CJ, Li H. 2003. Structure and function of archaeal box C/D sRNP core proteins. *Nat Struct Biol 10*:256-263.

Alioto TS. 2007. U12DB: a database of orthologous U12-type spliceosomal introns. *Nucleic Acids Res 35*:D110-115.

al-Maghtheh M, Inglehearn CF, Keen TJ, Evans K, Moore AT, Jay M, Bird AC, Bhattacharya SS. 1994. Identification of a sixth locus for autosomal dominant retinitis pigmentosa on chromosome 19. *Hum Mol Genet 3*:351-354.

Beales PL, Elcioglu N, Woolf AS, Parker D, Flinter FA. 1999. New criteria for improved diagnosis of Bardet-Biedl syndrome: results of a population survey. *J Med Genet* 36:437–46.

Beck M, Brickley K, Wilkinson HL, Sharma S, Smith M, Chazot PL, Pollard S, Stephenson FA. 2002. Identification, molecular cloning, and characterisation of a novel GABAA receptor-associated protein, GRIF-1. *J Biol Chem 277*:30079-30090.

Bellare P, Small EC, Huang X, Wohlschlegel JA, Staley JP, Sontheimer EJ. 2008. A role for ubiquitin in the spliceosome assembly pathway. *Nat Struct Mol Biol 15*:444-451.

Black DL. 2003. Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem 72*:291-336.

Boisvert FM, van Koningsbruggen S, Navascues J, Lamond AI. 2007. The multifunctional nucleolus. *Nat Rev Mol Cell Biol 8*:574-585.

Bornancin F, Mechtcheriakova D, Stora S, Graf C, Wlachos A, Devay P, Urtz N, Baumruker T, Billich A. 2005. Characterization of a ceramide kinase-like protein. *Biochim Biophys Acta* 1687:31–43.

Boughman JA, Vernon M, Shaver KA. 1983. Usher syndrome: definition and estimate of prevalence from two high-risk populations. *J Chronic Dis* 36: 595–603.

Bowne SJ, Sullivan LS, Blanton SH, Cepko CL, Blackshaw S, Birch DG, Hughbanks-Wheaton D, Heckenlively JR, Daiger SP. 2002. Mutations in the inosine monophosphate dehydrogenase 1 gene (IMPDH1) cause the RP10 form of autosomal dominant retinitis pigmentosa. *Hum Mol Genet 11*:559-568.

Boylan JP, Wright AF. 2000. Identification of a novel protein interacting with RPGR. *Hum Mol Genet 9*:2085-2093.

Brickley K, Smith MJ, Beck M, Stephenson FA. 2005. GRIF-1 and OIP106, members of a novel gene family of coiled-coil domain proteins: association in vivo and in vitro with kinesin. *J Biol Chem 280*:14723-14732.

Bron AJ, Tripathi RC, Tripathi BJ. 1997. Wolff's Anatomy of the Eye. 8th ed. Chapman & Hall Medical, London.

Bujakowska KM, Maubaret C, Chakarova CF, Tanimoto N, Beck SC, Fahl E, Humphries MM, Kenna P, Makarov E, Makarova O, Paquet-Durand F, Ekstrom P, van Veen T, Leveillard T, Humphries P, Seeliger M, Bhattacharya SS. 2009. Study of gene targeted mouse models of splicing factor gene Prpf31 implicated in human autosomal dominant retinitis pigmentosa (RP). *Invest Ophthalmol Vis Sci*.

Chakarova CF, Hims MM, Bolz H, Abu-Safieh L, Patel RJ, Papaioannou MG, Inglehearn CF, Keen TJ, Willis C, Moore AT, Rosenberg T, Webster AR, Bird AC, Gal A, Hunt D, Vithana EN, Bhattacharya SS. 2002. Mutations in HPRP3, a third member of pre-mRNA splicing factor genes, implicated in autosomal dominant retinitis pigmentosa. *Hum Mol Genet 11*:87-92.

Chakarova CF, Papaioannou MG, Khanna H, Lopez I, Waseem N, Shah A, Theis T, Friedman J, Maubaret C, Bujakowska K, Veraitch B, Abd El-Aziz MM, Prescott de Q, Parapuram SK, Bickmore WA, Munro PM, Gal A, Hamel CP, Marigo V, Ponting CP, Wissinger B, Zrenner E, Matter K, Swaroop A, Koenekoop RK, Bhattacharya SS. 2007. Mutations in TOPORS cause autosomal dominant retinitis pigmentosa with perivascular retinal pigment epithelium atrophy. *Am J Hum Genet 81*:1098-1103.

Chapple JP, Grayson C, Hardcastle AJ, Bailey TA, Matter K, Adamson P, Graham CH, Willison KR, Cheetham ME. 2003. Organization on the plasma membrane of the retinitis pigmentosa protein RP2: investigation of association with detergent-resistant membranes and polarized sorting. *Biochem J* 372:427–433.

Chen P, Hao W, Rife L, Wang XP, Shen D, Chen J, Ogden T, Van Boemel GB, Wu L, Yang M, Fong HK. 2001. A photic visual cycle of rhodopsin regeneration is dependent on Rgr. *Nat Genet* 28:256–260.

Clarke G, Goldberg AF, Vidgen D, Collins L, Ploder L, Schwartz L, Molday LL, Rossant J, Szel A, Molday RS, Birch DG, McInnes RR. 2000. Rom-1 is required for rod photoreceptor viability and the regulation of disk morphogenesis. *Nat Genet* 25:67–73

Combs DJ, Nagel RJ, Ares M, Jr., Stevens SW. 2006. Prp43p is a DEAH-box spliceosome disassembly factor essential for ribosome biogenesis. *Mol Cell Biol 26*:523-534.

Comer FI, Hart GW. 2001. Reciprocity between O-GlcNAc and O-phosphate on the carboxyl terminal domain of RNA polymerase II. *Biochemistry 40*:7845-7852.

Comitato A, Spampanato C, Chakarova C, Sanges D, Bhattacharya SS, Marigo V. 2007. Mutations in splicing factor PRPF3, causing retinal degeneration, form detrimental aggregates in photoreceptor cells. *Hum Mol Genet 16*:1699-1707.

Daiger SP, Bowne SJ, Sullivan LS. 2007. Perspective on Genes and Mutations Causing Retinitis Pigmentosa. *Arch Ophthalmol*. 125(2): 151–158.

Davies WL, Carvalho LS, Hunt DM. 2007. SPLICE: a technique for generating in vitro spliced coding sequences from genomic DNA. *Biotechniques 43*:785-789.

Deery EC, Vithana EN, Newbold RJ, Gallon VA, Bhattacharya SS, Warren MJ, Hunt DM, Wilkie SE. 2002. Disease mechanism for retinitis pigmentosa (RP11) caused by mutations in the splicing factor gene PRPF31. *Hum Mol Genet 11*:3209-3219.

Denis MM, Tolley ND, Bunting M, Schwertz H, Jiang H, Lindemann S, Yost CC, Rubner FJ, Albertine KH, Swoboda KJ, Fratto CM, Tolley E, Kraiss LW, McIntyre TM, Zimmerman GA, Weyrich AS. 2005. Escaping the nuclear confines: signal-dependent pre-mRNA splicing in anucleate platelets. *Cell 122*:379-391.

Dhallan RS, Macke JP, Eddy RL, Shows TB, Reed RR, Yau KW, Nathans J. 1992. Human rod photoreceptor cGMP-gated channel: amino acid sequence, gene structure, and functional expression. *J Neurosci* 12:3248–3256.

Dryja TP. 2000. Molecular genetics of Oguchi disease, fundus albipunctatus, and other forms of stationary night blindness: LVII Edward Jackson Memorial Lecture. *Am J Ophthalmol* 130:547–563.

Dryja TP, Adams SM, Grimsby JL, McGee TL, Hong DH, Li T, Andreasson S, Berson EL. 2001. Null RPGRIP1 alleles in patients with Leber congenital amaurosis. *Am J Hum Genet 68*:1295-1298.

Eudy JD, Weston MD, Yao S, Hoover DM, Rehm HL, Ma-Edmonds M, Yan D, Ahmad I, Cheng JJ, Ayuso C, Cremers C, Davenport S, Moller C, Talmadge CB, Beisel KW, Tamayo M, Morton CC, Swaroop A, Kimberling WJ, Sumegi J. 1998.Mutation of a gene encoding a protein with extracellular matrix motifs in Usher syndrome type IIa. *Science* 280: 1753-1757.

Fields S, Song O. 1989. A novel genetic system to detect protein-protein interactions. *Nature 340*:245-246.

Francois J. 1968. Leber's congenital tapetoretinal degeneration. *Int Ophthalmol Clin* 8:929–947.

Franklin JL, Yoshiura K, Dempsey PJ, Bogatcheva G, Jeyakumar L, Meise KS, Pearsall RS, Threadgill D, Coffey RJ. 2005. Identification of MAGI-3 as a transforming growth factor-alpha tail binding protein. *Exp Cell Res 303*:457-470.

Frio TR, McGee TL, Wade NM, Iseli C, Beckmann JS, Berson EL, Rivolta C. 2009. A single-base substitution within an intronic repetitive element causes dominant retinitis pigmentosa with reduced penetrance. *Hum Mutat 30*:1340-1347.

Funke L, Dakoji S, Bredt DS. 2005. Membrane-associated guanylate kinases regulate adhesion and plasticity at cell junctions. *Annu Rev Biochem 74*:219-245.

Gamundi MJ, Hernan I, Muntanyola M, Maseras M, Lopez-Romero P, Alvarez R, Dopazo A, Borrego S, Carballo M. 2008. Transcriptional expression of cis-acting and trans-acting splicing mutations cause autosomal dominant retinitis pigmentosa. *Hum Mutat 29*:869-878.

Gerard MA, Krol A, Carbon P. 2007. Transcription factor hStaf/ZNF143 is required for expression of the human TFAM gene. *Gene 401*:145-153.

Glanzer J, Miyashiro KY, Sul JY, Barrett L, Belt B, Haydon P, Eberwine J. 2005. RNA splicing capability of live neuronal dendrites. *Proc Natl Acad Sci U S A 102*:16859-16864.

Gonzalez-Santos JM, Cao H, Duan RC, Hu J. 2008. Mutation in the splicing factor Hprp3p linked to retinitis pigmentosa impairs interactions within the U4/U6 snRNP complex. *Hum Mol Genet 17*:225-239.

Gorska-Andrzejak J, Stowers RS, Borycz J, Kostyleva R, Schwarz TL, Meinertzhagen IA. 2003. Mitochondria are redistributed in Drosophila photoreceptors lacking milton, a kinesin-associated protein. *J Comp Neurol 463*:372-388.

Graham DK, Dawson TL, Mullaney DL, Snodgrass HR, Earp HS. 1994. Cloning and mRNA expression analysis of a novel human protooncogene, c-mer. *Cell Growth Differ* 5(6):647-57.

Grainger RJ, Beggs JD. 2005. Prp8 protein: at the heart of the spliceosome. *Rna 11*:533-557.

Green WR, Enger C. 1993. Age-related macular degeneration histopathologic studies. The 1992 Lorenz E. Zimmerman Lecture. *Ophthalmology* 100:1519–1535.

Gregory CY, Evans K, Wijesuriya SD, Kermani S, Jay MR, Plant C, Cox N, Bird AC, Bhattacharya SS. 1996. The gene responsible for autosomal dominant Doyne's honeycomb retinal dystrophy (DHRD) maps to chromosome 2p16. *Hum Molec Genet* 5:1055-1059.

Grishin A, Li H, Levitan ES, Zaks-Makhina E. 2006. Identification of gamma-aminobutyric acid receptor-interacting factor 1 (TRAK2) as a trafficking factor for the K+ channel Kir2.1. *J Biol Chem 281*:30104-30111.

Grossman CE, Qian Y, Banki K, Perl A. 2004. ZNF143 mediates basal and tissue-specific expression of human transaldolase. *J Biol Chem 279*:12190-12205.

Hamel C. 2006. Retinitis pigmentosa. *Orphanet J Rare Dis* 11;1:40.

Hargrave PA. 2001. Rhodopsin structure, function, and topography: the Friedenwald lecture. *Invest Ophthalmol Vis Sci* 42: 3–9.

Hartong DT, Berson EL, Dryja TP. 2006. Retinitis pigmentosa. *Lancet 368*:1795-1809.

He J, Bellini M, Inuzuka H, Xu J, Xiong Y, Yang X, Castleberry AM, Hall RA. 2006. Proteomic analysis of beta1-adrenergic receptor interactions with PDZ scaffold proteins. *J Biol Chem 281*:2820-2827.

Hinsby AM, Kiemer L, Karlberg EO, Lage K, Fausboll A, Juncker AS, Andersen JS, Mann M, Brunak S. 2006. A wiring of the human nucleolus. *Mol Cell 22*:285-295.

Hofmann I, Schnolzer M, Kaufmann I, Franke WW. 2002. Symplekin, a constitutive protein of karyo- and cytoplasmic particles involved in mRNA biogenesis in Xenopus laevis oocytes. *Mol Biol Cell 13*:1665-1676.

Hong DH, Pawlyk B, Sokolov M, Strissel KJ, Yang J, Tulloch B, Wright AF, Arshavsky VY, Li T. 2003. RPGR isoforms in photoreceptor connecting cilia and the transitional zone of motile cilia. *Invest Ophthalmol Vis Sci* 44:2413–2421.

House AE, Lynch KW. 2008. Regulation of alternative splicing: more than just the ABCs. *J Biol Chem 283*:1217-1221.

Iyer SP, Akimoto Y, Hart GW. 2003. Identification and cloning of a novel family of coiled-coil domain proteins that interact with O-GlcNAc transferase. *J Biol Chem 278*:5399-5409.

James P, Halladay J, Craig EA. 1996. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics 144*:1425-1436.

Karan S, Zhang H, Li S, Frederick JM, Baehr W. 2008. A model for transport of membrane-associated phototransduction polypeptides in rod and cone photoreceptor inner segments. *Vision Res 48*:442-452.

Kavanagh E, Buchert M, Tsapara A, Choquet A, Balda MS, Hollande F, Matter K. 2006. Functional interaction between the ZO-1-interacting transcription factor ZONAB/DbpA and the RNA processing factor symplekin. *J Cell Sci 119*:5098-5105.

Keen TJ, Hims MM, McKie AB, Moore AT, Doran RM, Mackey DA, Mansfield DC, Mueller RF, Bhattacharya SS, Bird AC, Markham AF, Inglehearn CF. 2002. Mutations in a protein target of the Pim-1 kinase associated with the RP9 form of autosomal dominant retinitis pigmentosa. *Eur J Hum Genet 10*:245-249.

Keon BH, Schafer S, Kuhn C, Grund C, Franke WW. 1996. Symplekin, a novel type of tight junction plaque protein. *J Cell Biol 134*:1003-1018.

Kirk E, Chin LS, Li L. 2006. GRIF1 binds Hrs and is a new regulator of endosomal trafficking. *J Cell Sci* 119:4689-701.

Kirschner R, Rosenberg T, Schultz-Heienbrok R, Lenzner S, Feil S, Roepman R, Cremers FPM, Ropers H-H, Berger W. 1999. RPGR transcription studies in mouse and human tissues reveal a retina-specific isoform that is disrupted in a patient with X-linked retinitis pigmentosa. *Hum Molec Genet* 8:1571-1578.

Kobayashi M, Takezawa S, Hara K, Yu RT, Umesono Y, Agata K, Taniwaki M, Yasuda K, Umesono K. 1999. Identification of a photoreceptor cell-specific nuclear receptor. *Proc Nat Acad Sci* 96:4814-4819.

Konig H, Matter N, Bader R, Thiele W, Muller F. 2007. Splicing segregation: the minor spliceosome acts outside the nucleus and controls cell proliferation. *Cell 131*:718-729.

Korschen HG, Beyermann M, Muller F, Heck M, Vantler M, Koch KW, Kellner R, Wolfrum U, Bode C, Hofmann KP, Kaupp UB. 1999. Interaction of glutamic-acid-rich proteins with the cGMP signalling pathway in rod photoreceptors. *Nature* 400:761–766

Koutalos Y, Nakatani K, Yau KW. 1995. The cGMP-phosphodiesterase and its contribution to sensitivity regulation in retinal rods. *J Gen Physiol* 106:891–921

Letteboer SJ, Roepman R. 2008. Versatile screening for binary protein-protein interactions by yeast two-hybrid mating. *Methods Mol Biol 484*:145-159.

Leung AK, Andersen JS, Mann M, Lamond AI. 2003. Bioinformatic analysis of the nucleolus. *Biochem J 376*:553-569.

Li N, Mei H, Macdonald IM, Jiao X, Hejtmancik F. 2009. Mutations in ASCC3L1 on chromosome 2q11.2 are associated with autosomal dominant retinitis pigmentosa in a Chinese Family. *Invest Ophthalmol Vis Sci*.

Li Q, Lee JA, Black DL. 2007. Neuronal regulation of alternative pre-mRNA splicing. *Nat Rev Neurosci 8*:819-831.

Li ZY, Possin DE, Milam AH. 1995. Histopathology of bone spicule pigmentation in retinitis pigmentosa, *Ophthalmology* 102:805–816.

Lim KP, Yip SP, Cheung SC, Leung KW, Lam ST, To CH. 2009. Novel PRPF31 and PRPH2 mutations and co-occurrence of PRPF31 and RHO mutations in Chinese patients with retinitis pigmentosa. *Arch Ophthalmol* 127:784-90.

Lin KT, Lu RM, Tarn WY. 2004. The WW domain-containing proteins interact with the early spliceosome and participate in pre-mRNA splicing in vivo. *Mol Cell Biol 24*:9176-9185.

Liu Q, Zuo J, Pierce EA. 2004. The retinitis pigmentosa 1 protein is a photoreceptor microtubule-associated protein. *J Neurosci* 24:6427–6436.

Liu S, Li P, Dybkov O, Nottrott S, Hartmuth K, Luhrmann R, Carlomagno T, Wahl MC. 2007. Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4 snRNP. *Science 316*:115-120.

Liu S, Rauhut R, Vornlocher HP, Luhrmann R. 2006. The network of protein-protein interactions within the human U4/U6.U5 tri-snRNP. *Rna 12*:1418-1430.

Macias MJ, Wiesner S, Sudol M. 2002. WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett 513*:30-37.

Maita H, Kitaura H, Ariga H, Iguchi-Ariga SM. 2005. Association of PAP-1 and Prp3p, the products of causative genes of dominant retinitis pigmentosa, in the tri-snRNP complex. *Exp Cell Res 302*:61-68.

Maita H, Kitaura H, Keen TJ, Inglehearn CF, Ariga H, Iguchi-Ariga SM. 2004. PAP-1, the mutated gene underlying the RP9 form of dominant retinitis pigmentosa, is a splicing factor. *Exp Cell Res* 300:283–296.

Makarova OV, Makarov EM, Liu S, Vornlocher HP, Luhrmann R. 2002. Protein 61K, encoded by a gene (PRPF31) linked to autosomal dominant retinitis pigmentosa, is required for U4/U6*U5 tri-snRNP formation and pre-mRNA splicing. *Embo J 21*:1148-1157.

Makarova OV, Makarov EM, Luhrmann R. 2001. The 65 and 110 kDa SR-related proteins of the U4/U6.U5 tri-snRNP are essential for the assembly of mature spliceosomes. *Embo J 20*:2553-2563.

Martin KC, Zukin RS. 2006. RNA trafficking and local protein synthesis in dendrites: an overview. *J Neurosci 26*:7131-7134.

Mathew R, Hartmuth K, Mohlmann S, Urlaub H, Ficner R, Luhrmann R. 2008. Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome. *Nat Struct Mol Biol 15*:435-443.

McKie AB, McHale JC, Keen TJ, Tarttelin EE, Goliath R, van Lith-Verhoeven JJ, Greenberg J, Ramesar RS, Hoyng CB, Cremers FP, Mackey DA, Bhattacharya SS, Bird AC, Markham AF, Inglehearn CF. 2001. Mutations in the pre-mRNA splicing factor gene PRPC8 in autosomal dominant retinitis pigmentosa (RP13). *Hum Mol Genet 10*:1555-1562.

Mordes D, Yuan L, Xu L, Kawada M, Molday RS, Wu JY. 2007. Identification of photoreceptor genes affected by PRPF31 mutations associated with autosomal dominant retinitis pigmentosa. *Neurobiol Dis 26*:291-300.

Muniz A, Villazana-Espinoza ET, Hatch AL, Trevino SG, Allen DM, Tsin AT. 2007. A novel cone visual cycle in the cone-dominated retina. *Exp Eye Res 85*:175-184.

Myslinski E, Gerard MA, Krol A, Carbon P. 2007. Transcription of the human cell cycle regulated BUB1B gene requires hStaf/ZNF143. *Nucleic Acids Res 35*:3453-3464.

Myslinski E, Krol A, Carbon P. 1998. ZNF76 and ZNF143 are two human homologs of the transcriptional activator Staf. *J Biol Chem 273*:21998-22006.

Neugebauer KM. 2002. On the importance of being co-transcriptional. *J Cell Sci 115*:3865-3871.

Norton AW, Hosier S, Terew JM, Li N, Dhingra A, Vardi N, Baehr W, Cote RH. 2005. Evaluation of the 17-kDa prenyl-binding protein as a regulatory protein for phototransduction in retinal photoreceptors. *J Biol Chem 280*:1248-1256.

Nottrott S, Urlaub H, Luhrmann R. 2002. Hierarchical, clustered protein interactions with U4/U6 snRNA: a biochemical role for U4/U6 proteins. *Embo J 21*:5527-5538.

Palczewski K. 2006. G protein-coupled receptor rhodopsin. *Annu Rev Biochem 75*:743-767.

Palczewski K, McDowell JH, Jakes S, Ingebritsen TS, Hargrave PA. 1989. Regulation of rhodopsin dephosphorylation by arrestin. *J Biol Chem* 264:15770–15773

Patel AA, McCarthy M, Steitz JA. 2002. The splicing of U12-type introns can be a rate-limiting step in gene expression. *Embo J 21*:3804-3815.

Pellikka M, Tanentzapf G, Pinto M, Smith C, McGlade CJ, Ready DF, Tepass U. 2002. Crumbs, the Drosophila homologue of human CRB1/RP12, is essential for photoreceptor morphogenesis. *Nature* 416:143–149

Pessa HK, Ruokolainen A, Frilander MJ. 2006. The abundance of the spliceosomal snRNPs is not limiting the splicing of U12-type introns. *Rna 12*:1883-1892.

Pessa HK, Will CL, Meng X, Schneider C, Watkins NJ, Perala N, Nymark M, Turunen JJ, Luhrmann R, Frilander MJ. 2008. Minor spliceosome components are predominantly localized in the nucleus. *Proc Natl Acad Sci U S A 105*:8655-8660.

Pugh EN, Lamb TD. 2000. Phototransduction in vertebrate rods and cones: molecular mechanism of amplification, recovery and light adaptation. Chapter 5, pp 183-255. Handbook of biological physics, volume 3, molecular mechanisms of visual transduction. Edited by Stavenga DG, de Grip WJ, Pugh EN. Elsevier Science

Reidel B, Goldmann T, Giessl A, Wolfrum U. 2008. The translocation of signaling molecules in dark adapting mammalian rod photoreceptor cells is dependent on the cytoskeleton. *Cell Motil Cytoskeleton 65*:785-800.

Reiners J, Van Wijk E, Marker T, Zimmermann U, Jurgens K, te Brinke H, Overlack N, Roepman R, Knipper M, Kremer H, Wolfrum U. 2005. Scaffold protein harmonin (USH1C) provides molecular links between Usher syndrome type 1 and type 2. *Hum Mol Genet* 14:3933–3943.

Rice DS, Huang W, Jones HA, Hansen G, Ye GL, Xu N, Wilson EA, Trouhton K, Vaddi K, Newton RC, Zambrowicz BP, Sands AT. 2004. Severe retinal degeneration associated with disruption of semaphorin 4A. *Invest Ophthalmol Vis Sci* 45:2767–2777.

Rio Frio T, Civic N, Ransijn A, Beckmann JS, Rivolta C. 2008a. Two trans-acting eQTLs modulate the penetrance of PRPF31 mutations. *Hum Mol Genet 17*:3154-3165.

Rio Frio T, Wade NM, Ransijn A, Berson EL, Beckmann JS, Rivolta C. 2008b. Premature termination codons in PRPF31 cause retinitis pigmentosa via haploinsufficiency due to nonsense-mediated mRNA decay. *J Clin Invest 118*:1519-1531.

Rivolta C, McGee TL, Rio Frio T, Jensen RV, Berson EL, Dryja TP. 2006. Variation in retinitis pigmentosa-11 (PRPF31 or RP11) gene expression between symptomatic and asymptomatic patients with dominant RP11 mutations. *Hum Mutat 27*:644-653.

Rivolta C, Sharon D, DeAngelis MM, Dryja TP. 2002. Retinitis pigmentosa and allied diseases: numerous diseases, genes, and inheritance patterns. *Hum Mol Genet 11*:1219-1227.

Roepman R, Bernoud-Hubac N, Schick DE, Maugeri A, Berger W, Ropers HH, Cremers FP, Ferreira PA. 2000a. The retinitis pigmentosa GTPase regulator (RPGR) interacts with novel transport-like proteins in the outer segments of rod photoreceptors. *Hum Mol Genet 9*:2095-2105.

Roepman R, Schick D, Ferreira PA. 2000b. Isolation of retinal proteins that interact with retinitis pigmentosa GTPase regulator by interaction trap screen in yeast. *Methods Enzymol 316*:688-704.

Rong J, Li SH, Li XJ. 2007. Regulation of intracellular HAP1 trafficking. *J Neurosci Res 85*:3025-3029.

Saari JC, Nawrot M, Kennedy BN, Garwin GG, Hurley JB, Huang J, Possin DE, Crabb JW. 2001. Visual cycle impairment in cellular retinaldehyde binding protein (CRALBP) knockout mice results in delayed dark adaptation. *Neuron* 29:739–748.

Saishin Y, Ishikawa R, Ugawa S, Guo W, Ueda T, Morimura H, Kohama K, Shimizu H, Tano Y, Shimada S. 2000. Retinal fascin: functional nature, subcellular distribution, and chromosomal localization. *Invest Ophthalmol Vis Sci* 41:2087–2095.

Saunier S, Salomon R, Antignac C. 2005. Nephronophthisis. *Curr Opin Genet Dev* 15:324-331.

Schwahn U, Lenzner S, Dong J, Feil S, Hinzmann B, van Duijnhoven G, Kirschner R, Hemberger M, Bergen AAB, Rosenberg T, Pinckers AJLG, Fundele R, Rosenthal A, Cremers FPM, Ropers H-H, Berger W. 1998. Positional cloning of the gene for X-linked retinitis pigmentosa 2. *Nature Genet.* 19: 327-332.

Sheibani N, Sorenson CM, Frazier WA. 2000. Differential modulation of cadherin-mediated cell-cell adhesion by platelet endothelial cell adhesion molecule-1 isoforms through activation of extracellular regulated kinases. *Mol Biol Cell 11*:2793-2802.

Smith MJ, Pozo K, Brickley K, Stephenson FA. 2006. Mapping the GRIF-1 binding domain of the kinesin, KIF5C, substantiates a role for GRIF-1 as an adaptor protein in the anterograde trafficking of cargoes. *J Biol Chem 281*:27216-27228.

Solovei I, Kreysing M, Lanctot C, Kosem S, Peichl L, Cremer T, Guck J, Joffe B. 2009. Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. *Cell 137*:356-368.

Stanek D, Neugebauer KM. 2006. The Cajal body: a meeting place for spliceosomal snRNPs in the nuclear maze. *Chromosoma 115*:343-354.

Stowers RS, Megeath LJ, Gorska-Andrzejak J, Meinertzhagen IA, Schwarz TL. 2002. Axonal transport of mitochondria to synapses depends on milton, a novel Drosophila protein. *Neuron 36*:1063-1077.

Swaroop A, Xu J, Pawar H, Jackson A, Skolnick C, Agarwal N. 1992. A conserved retina-specific gene encodes a basic motif/leucine zipper domain. *Proc Nat Acad Sci* 89:266-270.

Tanackovic G, Rivolta C. 2009. PRPF31 alternative splicing and expression in human retina. *Ophthalmic Genet 30*:76-83.

Thomas M, Laura R, Hepner K, Guccione E, Sawyers C, Lasky L, Banks L. 2002. Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation. *Oncogene 21*:5088-5096.

Tieder M, Levy M, Gubler MC, Gagnadoux MF, Broyer M. 1982. Renal abnormalities in the Bardet-Biedl syndrome. *Int J Pediatr Nephrol* 3:199–203.

To K, Adamian M, Jakobiec FA, Berson EL. 1998 Histopathologic and immunohistochemical study of dominant cone degeneration. *Am J Ophthalmol* 126:140–142.

Travis GH, Golczak M, Moise AR, Palczewski K. 2007. Diseases caused by defects in the visual cycle: retinoids as potential therapeutic agents. *Annu Rev Pharmacol Toxicol 47*:469-512.

Travis GH, Sutcliffe JG, Bok D. 1991. The retinal degeneration slow (rds) gene product is a photoreceptor disc membrane-associated glycoprotein. *Neuron* 6:61–70.

Umen JG, Guthrie C. 1995. Prp16p, Slu7p, and Prp8p interact with the 3' splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. *RNA* 1:584–597.

van der Spuy J, Chapple JP, Clark BJ, Luthert PJ, Sethi CS, Cheetham ME. 2002. The Leber congenital amaurosis gene product AIPL1 is localized exclusively in rod photoreceptors of the adult human retina. *Hum Mol Genet 11*:823-831.

Vithana EN, Abu-Safieh L, Allen MJ, Carey A, Papaioannou M, Chakarova C, Al-Maghtheh M, Ebenezer ND, Willis C, Moore AT, Bird AC, Hunt DM, Bhattacharya SS. 2001. A human homolog of yeast pre-mRNA splicing gene, PRP31, underlies autosomal dominant retinitis pigmentosa on chromosome 19q13.4 (RP11). *Mol Cell 8*:375-381.

Vithana EN, Abu-Safieh L, Pelosini L, Winchester E, Hornan D, Bird AC, Hunt DM, Bustin SA, Bhattacharya SS. 2003. Expression of PRPF31 mRNA in patients with autosomal dominant retinitis pigmentosa: a molecular clue for incomplete penetrance? *Invest Ophthalmol Vis Sci 44*:4204-4209.

Vollrath D, Feng W, Duncan JL, Yasumura D, D'Cruz PM, Chappelow A, Matthes MT, Kay MA, LaVail MM. 2001. Correction of the retinal dystrophy phenotype of the RCS rat by viral gene transfer of Mertk. *Proc Natl Acad Sci USA* 98:12584–12589.

von Schantz M, Lucas RJ, Foster RG. 1999. Circadian oscillation of photopigment transcript levels in the mouse retina. *Brain Res Mol Brain Res 72*:108-114.

Wahl MC, Will CL, Luhrmann R. 2009. The spliceosome: design principles of a dynamic RNP machine. *Cell 136*:701-718.

Wakasugi T, Izumi H, Uchiumi T, Suzuki H, Arao T, Nishio K, Kohno K. 2007. ZNF143 interacts with p73 and is involved in cisplatin resistance through the transcriptional regulation of DNA repair genes. *Oncogene 26*:5194-5203.

Wang A, Forman-Kay J, Luo Y, Luo M, Chow YH, Plumb J, Friesen JD, Tsui LC, Heng HH, Woolford JL, Hu J. 1997. Identification and characterization of human

genes encoding Hprp3p and Hprp4p, interacting components of the spliceosome. *Hum Mol Genet* 6:2117–2126.

Wang DY, Chan WM, Tam PO, Baum L, Lam DS, Chong KK, Fan BJ, Pang CP. 2005. Gene mutations in retinitis pigmentosa and their clinical implications. *Clin Chim Acta 351*:5-16.

Wassle H. 2004. Parallel processing in the mammalian retina. *Nat Rev Neurosci 5*:747-757.

Weidenhammer EM, Ruiz-Noriega M, Woolford JL, Jr. 1997. Prp31p promotes the association of the U4/U6 x U5 tri-snRNP with prespliceosomes to form spliceosomes in Saccharomyces cerevisiae. *Mol Cell Biol 17*:3580-3588.

Weng J, Mata NL, Azarian SM, Tzekov RT, Birch DG, Travis GH. 1999. Insights into the function of Rim protein in photoreceptors and etiology of Stargardt's disease from the phenotype in abcr knockout mice. *Cell* 98:13–23.

White K, Marquardt A, Weber BHF. 2000. VMD2 mutations in vitelliform macular dystrophy (Best disease) and other maculopathies. *Hum Mutat* 15:301-308.

Will CL, Luhrmann R. 2005. Splicing of a rare class of introns by the U12-dependent spliceosome. *Biol Chem 386*:713-724.

Wright GJ, Leslie JD, Ariza-McNaughton L, Lewis J. 2004. Delta proteins and MAGI proteins: an interaction of Notch ligands with intracellular scaffolding molecules and its significance for zebrafish development. *Development 131*:5659-5669.

Wu Y, Dowbenko D, Spencer S, Laura R, Lee J, Gu Q, Lasky LA. 2000. Interaction of the tumor suppressor PTEN/MMAC with a PDZ domain of MAGI3, a novel membrane-associated guanylate kinase. *J Biol Chem 275*:21477-21485.

Xi Q, Pauer GJ, Marmorstein AD, Crabb JW, Hagstrom SA. 2005. Tubby-like protein 1 (TULP1) interacts with F-actin in photoreceptor cells. *Invest Ophthalmol Vis Sci* 46:4754–4761.

Xue L, Gollapalli DR, Maiti P, Jahng WJ, Rando RR. 2004. A palmitoylation switch mechanism in the regulation of the visual cycle. *Cell* 117:761–771.

Yang Z, Alvarez BV, Chakarova C, Jiang L, Karan G, Frederick JM, Zhao Y, Sauve Y, Li X, Zrenner E, Wissinger B, Hollander AI, Katz B, Baehr W, Cremers FP, Casey JR, Bahattacharya SS, Zhang K. 2005. Mutant carbonic anhydrase 4 impairs pH regulation and causes retinal photoreceptor degeneration. *Hum Mol Genet* 14:255–265.

Yuan CC, Zhao X, Florens L, Swanson SK, Washburn MP, Hernandez N. 2007. CHD8 associates with human Staf and contributes to efficient U6 RNA polymerase III transcription. *Mol Cell Biol 27*:8729-8738.

Zhang H, Wang D, Sun H, Hall RA, Yun CC. 2007. MAGI-3 regulates LPA-induced activation of Erk and RhoA. *Cell Signal 19*:261-268.

Zhou Z, Licklider LJ, Gygi SP, Reed R. 2002. Comprehensive proteomic analysis of the human spliceosome. *Nature* 419:182–185.