Disease mechanism for retinitis pigmentosa (RP11) caused by mutations in the splicing factor gene *PRPF31*

Evelyne C. Deery[1], Eranga N. Vithana[2], Richard J. Newbold[1], Victoria A. Gallon[1], Shomi S. Bhattacharya[2], Martin. J. Warren[1], David M. Hunt[2] and Susan E. Wilkie[2],*

[1]School of Biological Sciences, Queen Mary, University of London, London E1 4NS, UK and [2]Division of Molecular Genetics, Institute of Ophthalmology, University College London, 11–43 Bath Street, London EC1V 9EL, UK

This study investigates the functional consequences of two mutations, A194E and A216P, in the splicing factor gene *PRPF31* linked to autosomal dominant retinitis pigmentosa (RP11). Using a yeast complementation assay, we demonstrate that introduction of the human A216P mutation into the yeast orthologue *PRP31p* results in only partial rescue of growth at the restrictive temperature, indicating that splicing function is not fully restored. An *in vivo* assay of splicing function in human cells using a bovine rod opsin splicing template did not detect any defect in splicing efficiency or accuracy attributable to either mutation, suggesting that neither has a dominant negative effect on splicing. However, western analysis and immunofluorescence microscopy of mammalian cells transfected with *PRPF31* revealed that both mutations substantially hinder translocation of the protein into the nucleus. The overall effect may thus be an insufficiency in splicing function, which is revealed only under conditions of elevated splicing demand. With the need to replenish disc proteins on a daily basis, such conditions will exist in rod photoreceptors and this may underlie the disease pathology.

INTRODUCTION

Retinitis pigmentosa (RP) is a genetically heterogeneous disorder characterized by progressive degeneration of the peripheral retina leading to night blindness and loss of visual fields. The disorder is characterized by bone spicule-like pigmentary deposits and a reduced or absent electroretinogram and often progresses to complete blindness in later life. With an incidence of around 1 in 3500, RP can be inherited as an X-linked, autosomal-dominant or autosomal-recessive condition. Eleven known autosomal dominant loci have been reported, including eight identified genes. The three most recently identified genes all encode proteins involved in pre-mRNA splicing, *HPRP3* on chromosome 1p13–q21 (RP18) (1), *PRPC8* on chromosome 17p13.3 (RP13) (2) and *PRPF31* on chromosome 19q13.4 (RP11) (3).

Mutations in *PRPF31* (RP11) have now been identified in British, Japanese and American families, implicating this gene as a major cause of autosomal-dominant RP (3). Affected members from all these families have a type II/regional form of RP and, uniquely, show an 'all or none' form of incomplete penetrance. A detailed clinical description of RP11 families has been reported previously (4,5). Seven mutations have been described including two missense changes (A194E and A216P), deletions leading to frameshifts, a duplication resulting in the insertion of 11 novel amino acids and two splice site mutations (3).

The splicing of pre-mRNA in the nucleus is catalysed by a large ribonucleoprotein complex, the spliceosome. The spliceosome consists of the pre-mRNA substrate and several small nuclear ribonucleoproteins (snRNPs) together with splicing factors not integrated into snRNPs. Each snRNP is a stable complex of an RNA molecule (snRNA) and several proteins, some of which are common to all snRNPs (the Sm proteins) and others specific to a given snRNP (6). Splicing of the vast majority of introns (the so-called 'U2 introns') involves five snRNPs, U1, U2, U4, U5 and U6. Spliceosome assembly involves firstly the binding of U1 and U2 to the pre-mRNA substrate with the formation of a pre-splicing complex (complex A), followed by binding of a trimer of U4/U6 + U5 to form the spliceosome proper (complex B). Two transesterification reactions result in the removal of the intron as a lariat structure and splicing together of the two exons. After splicing, the spliceosome dissociates to the snRNP level (Fig. 1).

The gene *PRPF31* encodes a 61 kDa protein (PRPF31, also referred to as splicing factor 61 K) which has been shown to be
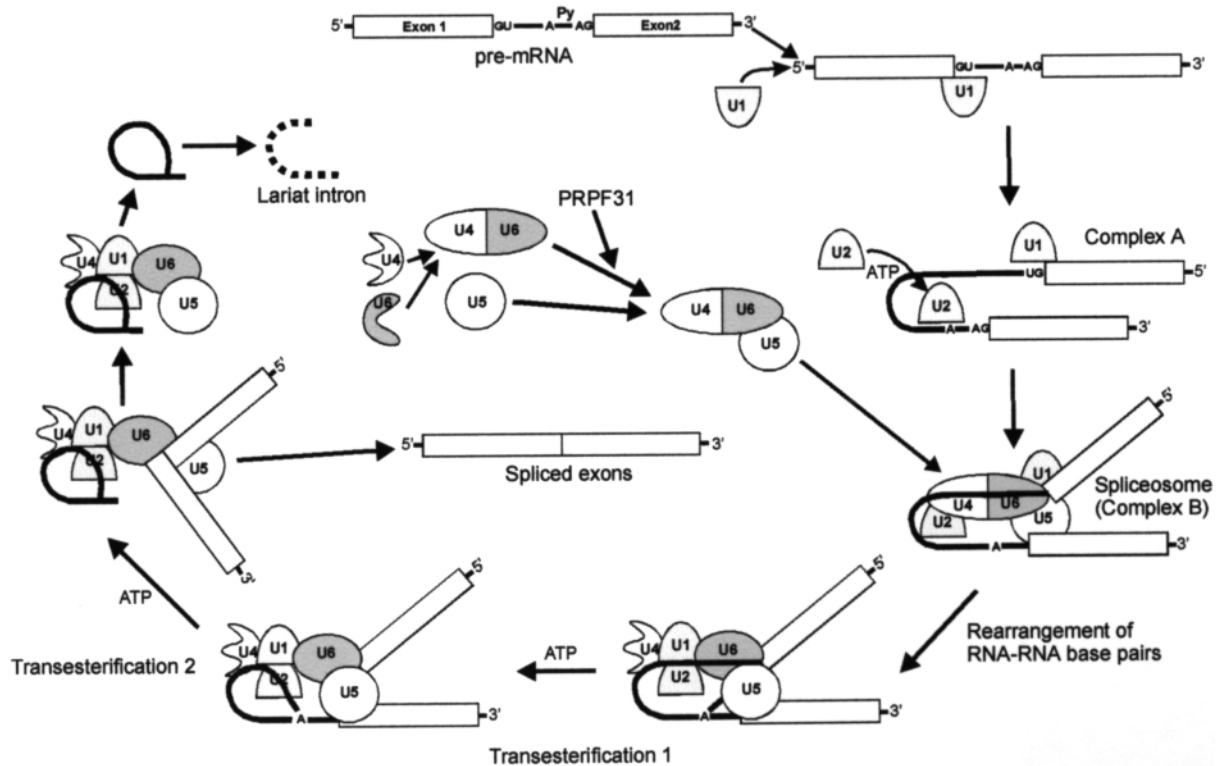
Figure 1. Role of PRPF31 in pre-mRNA splicing. PRPF31 is a splicing factor specific to the U4/U6 snRNP and its primary role is believed to be to recruit and physically tether U5 to U4/U6 to yield the tri-snRNP. Figure adapted from web entry http://lifesci.rutgers.edu./~kiledjian.

integral to the U4/U6 + U5 trimer (7). At each round of pre-mRNA splicing, the U4/U6 + U5 tri-snRNP must be assembled from its components, U4/U6 and U5. PRPF31 is operationally defined as U4/U6-specific as it remains bound to this particle at salt concentrations where the tri-snRNP dissociates. However, as shown by two-hybrid analysis and biochemical assays, it also interacts with a U5-associated 102 kDa protein, PRP6P (8). Thus the primary role of PRPF31 is thought to be to recruit and tether U5 to U4/U6 to yield the tri-snRNP.

The probable function of the RP11 gene in pre-mRNA splicing was deduced on the basis of sequence identity with the *PRP31p* and *PRP31+* genes in the yeasts, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, respectively (9,10). The overall sequence identities between PRPF31 and the two yeast proteins are surprisingly low for factors that are involved in such a fundamental cellular process (20% with *S. cerevisiae* PRP31p, 38% with *S. pombe* PRP31+). However, a Nop domain, which corresponds to a putative snRNA binding domain found in other pre-RNA-processing ribonucleoproteins (Pfam) (11), shows a much higher percentage identity (up to 47%). Whilst yeast PRP31p has been shown to be associated with the U4/U6 + U5 tri-snRNP, functional differences between the mammalian and yeast proteins have been demonstrated. In particular, unlike PRPF31, PRP31p does not appear to be essential for U4/U6 or U4/U6 + U5 assembly but is involved in the recruitment of the tri-snRNP to the pre-spliceosomal complex (12).

In this study we address the consequences of two disease-associated mutations in *PRPF31*, the missense mutations

A194E and A216P, on the functionality and localization of the protein variants. On the basis of yeast mutant complementation data, mammalian cell expression studies and protein localization experiments, we believe that these mutations result in a deficiency in splicing function under conditions of high splicing demand, as pertains in rod photoreceptor cells in the retina and this may underlie the pathology of the disease.

RESULTS

Complementation of PRP31p function in a PRP31p-deficient strain of yeast

Yeast PRP31p was originally identified by screening a panel of EMS-generated temperature-sensitive yeast mutants defective in pre-mRNA splicing (13). One of these strains, JWY2857, was shown to have a defect in the *PRP31p* gene such that it is able to grow at the permissive temperature of 23°C but fails to grow at the restrictive temperature of 37°C. Transformation of this strain with a plasmid expressing wild-type yeast PRP31p fully complements the defective gene (9). In order to examine the functional equivalence of the yeast and human genes, the coding sequence of human *PRPF31* was cloned into the yeast expression vector pMyr and used to transform the PRP31p-deficient strain. Although the transformed strain was able to grow at 23°C, it still failed to grow at the restrictive temperature of 37°C (Fig. 2). Confirmation of expression of the human protein in yeast cells was obtained by using the construct
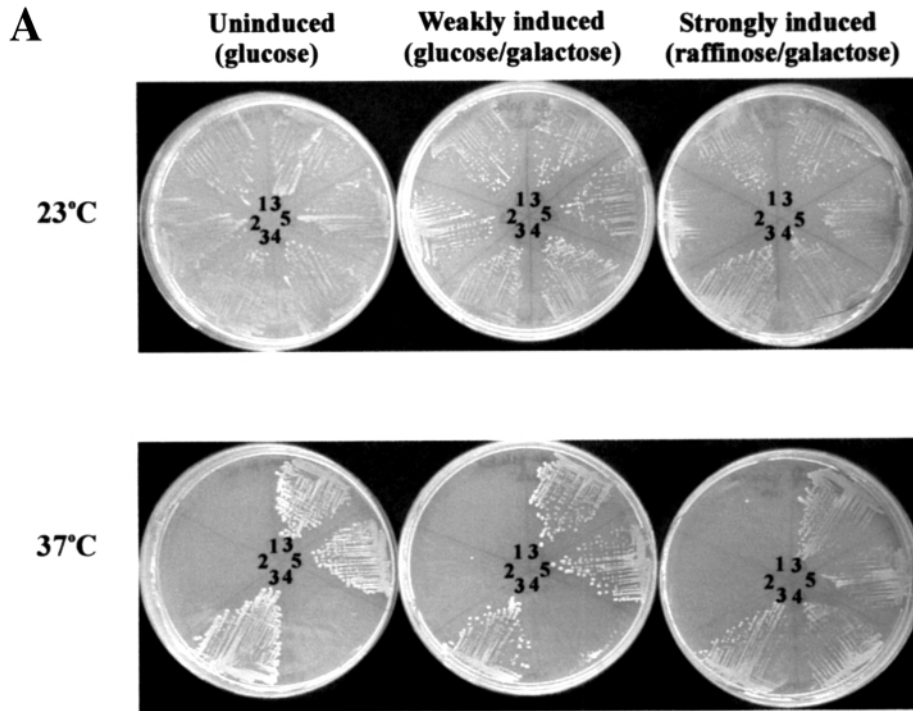
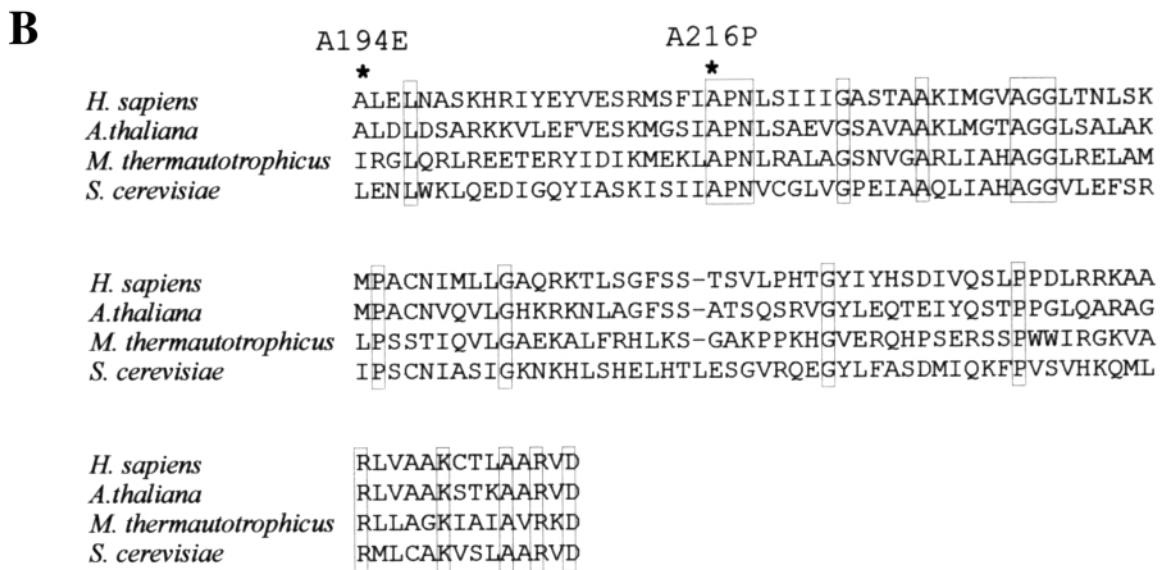| Strain JWY2857 transformed with:-	Growth at 23°C	Growth at 37°C Non-inducing conditions	Growth at 37°C Inducing conditions
1. Empty vector	+++	-	-
2. Human PRPF31	++	-	-
3. Yeast PRP31p (WT)	+++	++++	++++
4. Yeast PRP31p (A226P)	+++	-	++++
5. Yeast PRP31p (L204E)	+++	++++	++++

Figure 2. Complementation of a PRP31p-deficient yeast strain. *S. cerevisiae* strain JWY2857 grows at the permissive temperature of 23°C but fails to grow at the non-permissive temperature of 37°C. Complementation of the deficiency by transformation of the strain with functional PRP31p permits growth at the non-permissive temperature. (**A**) Agar plates showing the effect on growth after transformation with the indicated proteins and, below, table providing key and summarizing results. (**B**) Alignment of protein sequences of the Nop domain of PRPF31 orthologues from mammalian (*Homo sapiens*, accession no. AAK77986), plant (*Arabidopsis thaliana*, accession no. TO2269), yeast (*S. cerevisiae*, accession no. NP_011605) and bacterial (*Methanothermobacter thermautotrophicus*, accession no. NP_276342) species, showing the conservation at position 216 but not at 194.

pMyr-humanPRP[His] that placed a His-tag on to the C-terminus of the protein. Western analysis of cell extracts with the anti-His antibody revealed that human PRPF31 protein was expressed. These results strongly suggest that human PRPF31 cannot complement the function of yeast PRP31p. This failure may be explained by the low degree of sequence similarity between yeast PRP31p and human PRPF31 (20% identity), leading to potential problems in protein–protein interactions, or a lack of common functionality reflected in the rather different roles played by the two proteins in pre-RNA splicing in yeast and humans.

Of the two disease-causing mutations chosen for this study, only A216P occurs at a residue that is widely conserved in PRP31p orthologues across species. Indeed, this residue is conserved from bacteria to plants (Fig. 2B). In *S. cerevisiae*, the equivalent site is A226. In order to investigate whether a corresponding substitution of A226P in yeast PRP31p would have an effect on the ability of the protein to complement the defect in the PRP31p-deficient strain, an A226P substitution was introduced into a yeast PRP31p expression construct by site-directed mutagenesis and used to transform the PRP31p-deficient strain. As shown in Fig. 2A, the strain harbouring the A226P mutation was able to grow at both 23 and 37°C although, at the higher temperature, the strain grew more slowly than when complemented with wild-type yeast PRP31p. This indicates that the A226P mutant does not fully complement the temperature-sensitive mutant when grown at the restrictive temperature of 37°C. Expression from the pMyr vector is leaky, as shown by the growth of the strain under non-inducing conditions (glucose medium) when transformed with wild-type PRP31p. After transformation with the A226P mutant, growth under non-inducing conditions is extremely slow but becomes significant on transfer to inducing conditions (galactose/raffinose medium; Fig. 2A), indicating that the effect of the mutation can be overcome by protein overproduction. The change of the highly invariant alanine at position 226 to the amino acid proline is likely to have an effect on the structure in this highly conserved region of the protein. In yeast, this change would appear to have a significant but relatively mild effect, consistent with a reduced activity of the splicing system.

The A194E mutation was investigated in a similar manner. In the yeast protein, the equivalent position is occupied by a leucine residue (L204) and when this was replaced by glutamate, the resulting construct was able to complement fully the temperature-sensitive strain. Thus the yeast L204E variant would appear to have sufficient activity to support normal growth at 37°C. This can be explained on the basis that the position is not highly conserved (Fig. 2B) and the steric problems that might arise from accommodating a bulky glutamic acid residue into the position normally occupied by an alanine in the human protein would be less pronounced in the yeast protein, which naturally houses a leucine at this position.

Expression of human PRPF31 in a human cell line

No PRPF31-deficient mammalian cells lines are available to enable similar *in vivo* analyses of PRPF31 proteins. However, the effect of a dominant mutation can be monitored in cells transfected with a high-expression mutant construct where the normal endogenous gene product will be diluted by mutant protein. The cDNA sequence for PRPF31 was therefore amplified from human retinal cDNA and cloned into the expression vector pTriEx-1. This vector utilizes the strong chicken β-actin promoter enhanced by the CMV immediate early enhancer (14). The resulting construct was used to transfect the human cell line, HEK 293T.

Protein extracts from transfected cells were examined by SDS–polyacrylamide gel electrophoresis with Coomassie staining. No gross differences in the protein banding patterns indicative of defective splicing were observed between cells transfected with the wild-type construct and those transfected with mutant constructs A194E and A216P. However, a band running at around 61 kDa, presumed to be PRPF31 itself, varied in band intensity (Fig. 3A). This band was more intense in whole cell extracts from cells transfected with either wild-type or mutant PRPF31 constructs in comparison with cells transfected with empty pTriEx-1 vector, indicating that the expression level of the PRPF31 protein was indeed high. The band was also stronger in soluble cytosolic and nuclear fractions from cells transfected with wild-type constructs than from cells transfected with the A194E and A216P mutant constructs.

The pTriEx-1 construct places a C-terminal His.tag onto the expressed protein. Confirmation of the identity of the 61 kDa band present on the Coomassie-stained gel as the PRPF31 protein was obtained therefore by probing a corresponding western blot with an α-His.tag antibody (Fig. 3B). Although the total levels of wild-type and mutant PRPF31 expressed in the cells is similar, lower levels of soluble protein are present in the nucleus and cytosol with the A194E and A216P mutants than with the wild-type. For the wild-type, the combined signal strength from the cytosolic and nuclear fractions is approximately equal to that from the whole cell fractions, indicating that the majority of expressed protein is soluble. However, for the A194E and A216P mutants, it appears that a large fraction of the expressed PRPF31 is insoluble.

In vivo functional analysis in the human cell line

Commensurate with its general function in splicing, PRPF31 is expressed in all human cell types (3), yet the RP11 mutations only appear to be detrimental in the rod photoreceptor cells of the retina. A possible explanation for this apparent inconsistency lies in the observation that the outer segments of rod photoreceptors undergo circadian shedding of disc membranes (15) and the level of mRNA for rod opsin, the most abundant protein of the disc membranes, also displays a daily rhythm (16,17). This potentially places a very high demand on the cell for rod opsin synthesis at these times and even minor defects in mRNA splicing may, over time, compromise cell viability. The corollary of this is that such mutations may only reveal their limitations where the requirements for splicing are particularly high.

In order to further examine the efficiency and accuracy of splicing in HEK 293T cells expressing mutant PRPF31, a construct was made which expressed a synthetic bovine rod opsin transcript containing a single intron, intron 4 (pMT4i4). This splicing construct was driven by the strong adenovirus
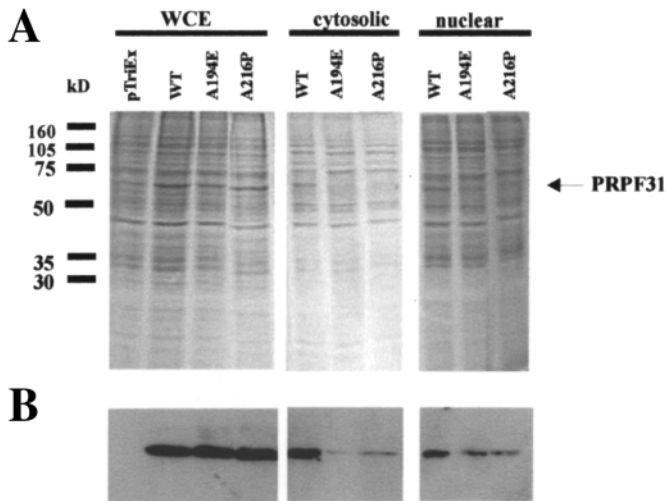
Figure 3. Expression of PRPF31 in HEK 293T cells. (**A**) SDS–PAGE analysis with Coomassie staining of whole cell extracts (WCE) and soluble cytosolic and nuclear proteins from cells transfected with pTriEx-1 expression constructs of wild-type and mutant PRPF31 and with empty pTriEx-1 vector (negative control). (**B**) Corresponding western analysis using α-His.tag primary antibody and HRP-conjugated secondary antibody.

Figure 4. RT–PCR assay on RNA extracts from HEK 293T cells co-transfected with PRPF31 (wild-type and mutants) and splicing template (ST). (**A**) Using exonic primers from either side of intron 4 (E4F/E5R) only spliced transcripts were detected. (**B**) A primer combination of one intronic primer and one exonic primer (In4F/E5R) was thus used to specifically amplify unspliced transcripts. No difference in the number of unspliced transcripts was detected between cells transfected with wild-type or mutant PRPF31. (**C**) Equal loading of total RNA in each assay was demonstrated using primers for the housekeeping gene *PGM1*. As a positive control, cells transfected with splicing template only (ST only) were analysed. Cells treated with plasmid but no transfection reagent gave no bands with the opsin primers (negative control).

major late promoter and would be expected to generate a high copy number of transcripts requiring splicing by the cell. HEK 293T cells were co-transfected with this splicing template and the pTriEx.PRPF31[His] expression construct. Total cellular RNA was then assayed by RT–PCR for the presence of aberrantly spliced or unspliced rod opsin transcripts. To ensure no carry-over of plasmid from the transfection, which could interfere in the RT–PCR, the RNA was treated with DNase1 prior to first strand cDNA synthesis. HEK 293T cells were selected for the assay as they have a high transfection efficiency, which should increase the probability that a high proportion of the cells will be transfected by both plasmids and thereby minimize background from cells transfected by the splicing construct only.

Two primer combinations were used for the RT–PCR, an exonic pair (E4F/E5R) that will amplify both the spliced and unspliced RNA transcripts and an intron–exon pair (In4F/E5R) that will only amplify from unspliced transcripts. In order to obtain an approximate estimate of the relative efficiencies of splicing in the presence of normal and mutant proteins, the number of cycles required for visualization of the PCR product within the exponential phase of amplification was established as 24 for the exonic primers and 28 for the intron–exon primers. For the exonic pair, only a single band indicative of the presence of spliced template was obtained from all co-transfected samples. No reproducible differences in band intensity were detected between cells transfected with wild-type and mutant PRPF31 constructs, in particular no reduction in spliced band intensity was seen with any of the mutants (Fig. 4A). As a positive control, cells transfected with the splicing template construct alone gave a single band of similar intensity, indicating that the endogenous splicing machinery in the cells was adequate to perform the necessary splicing function. RNA from cells treated with plasmid but no transfection reagent gave no RT–PCR band, demonstrating that there was no plasmid carry-over.
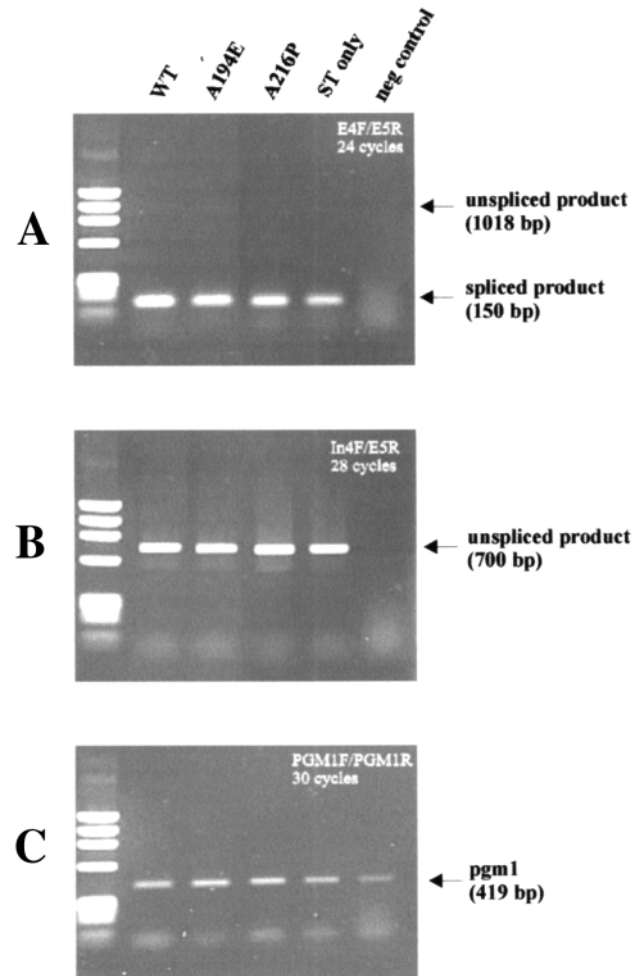
The intron–exon pair of primers was used to examine the possibility that differences may be present in the number of unspliced RNA transcripts present in cells transfected with wild-type and mutant PRPF31 constructs. Results from this PCR amplification indicated that unspliced transcripts were indeed present, although at lower concentrations than the spliced transcripts, as evidenced by the higher cycle number required to visualize the PCR product (Fig. 4B). However, no reproducible differences in band intensity were present between cells transfected with the normal and mutant constructs. PCRs using a primer pair from the housekeeping gene *PGM1* (18) confirmed that equal amounts of total cell RNA had been used in all samples (Fig. 4C). Taken together, these *in vivo* assays fail to detect any difference in splicing function between wild-type PRPF31 and either of the mutants studied, even when the mutant PRPF31 is

expressed at concentrations that would be expected to dilute out the effect of the cell's endogenous protein.

Cellular localization of wild-type and mutant PRPF31 in transfected COS7 cells

The differences in sub-cellular distribution of wild-type and mutant PRPF31 indicated by the western analysis of transfected HEK 293T cells were further investigated by indirect immunofluorescence microscopy using α-His.tag antibodies. HEK 293T cells are unsuitable for microscopic analysis because of their low adhesion to glass cover slips, so a more strongly adherent COS7 cell line was used. Transfected COS7 cells were probed with the α-His.tag primary antibody followed by an FITC-conjugated secondary antibody and nucleic acids in the cells were counter-stained with propidium iodide (PrI). The expressed PRPF31 protein appears as a green signal (Fig. 5, middle panel) whilst PrI gives a red signal and clearly delineates the nuclei and nucleoli (left panel). Co-localization of PRPF31 and nucleic acids can then be seen in merged images as a yellow/orange signal.

Wild-type PRPF31 gives a signal exclusively in the nucleus with the cytoplasm remaining unstained (Fig. 5A) whereas the western analysis indicated the presence of soluble wild-type PRPF31 in the cytosolic fraction. This dichotomy can be accounted for by the dispersal of the protein throughout the relatively large volume of the cytosol reducing the signal to below threshold visibility. Alternatively the cytosolic fraction examined in the western may have been contaminated with nuclear proteins. In Figure 5B, an enlargement of a thin slice through a transfected nucleus is shown, which confirms that the PRPF31 is present within the nucleoplasm but is excluded from the nucleoli. This staining pattern is consistent with that obtained by probing untransfected HeLa cells with an α-PRPF31 antibody (8), although in the present case the increased signal strength due to the high expression level of the recombinant protein would appear to obscure the sub-nucleoplasmic structures seen in the HeLa cells as nuclear speckles and Cajal bodies (8) that are thought to be the sites of snRNP and splicing factor storage (19). Cells transfected with empty pTriEx-1 vector gave no green signal, confirming that the green signal in transfected cells identifies the presence of His-tagged PRPF31 rather than some endogenous histidine-rich protein (result not shown).

Mutants A194E and A216P, however, gave a quite different staining pattern with signal throughout the cytoplasm and rather less intense staining in the nucleus than wild-type (Fig. 5D and F). These results indicate that mutations A194E and A216P result in mis-location of the protein. Some soluble protein does reach the nucleus, as seen from the confocal images and from the western analysis. This suggests that the problem may lie in the efficiency of the targeting mechanism whereby the protein is translocated from its site of synthesis in the cytosol into the nucleus.

Identification of the nuclear localization signal in PRPF31

Targeting of many nuclear proteins has been shown to depend on intrinsic nuclear localization signals (NLS). Three classes of NLS motif have been identified: the classical monopartite NLS, consisting of a single stretch of basic amino acids; the bipartite NLS, which has two basic motifs separated by a spacer of 10–12 amino acids; and a non-basic type of less clearly defined composition (reviewed in 20). For the two basic types, it is known that a simple cluster of basic amino acids is generally not sufficient, yet no clear consensus sequences have been identified (21). The region between residues 351 and 364 of PRPF31 contains a series of basic residues that are highly conserved amongst PRP31 homologues from a wide range of species ([351]RKKRGGRRYRKMKE[364]). This region has previously been identified as a putative NLS (3). A mutant in which this 14 residue sequence was deleted (ΔNLS) was constructed from wild-type PRPF31 by site-directed mutagenesis. This was expressed in COS7 cells and analysed by immunofluorescence microscopy as described above. The protein accumulated in the cytoplasm but was rigorously excluded from the nucleus (Fig. 5E). This result confirms therefore the identification of a classical monopartite NLS in PRPF31.

DISCUSSION

Evaluation of the results presented above should take account of the fact that PRPF31 is an essential gene for splicing in all cells of the body whereas the RP11 mutations exert their pathological effects only in one highly specialized cell type, namely, the rod photoreceptors. The mutations are dominant and thus exert their effect in the presence of wild-type protein. Moreover they show an 'all or none' mode of incomplete penetrance. Thus some individuals carry a mutation with no apparent detrimental effects whereas others with the same mutation show RP from an early age. Taken together, this basic evidence indicates that, for most cells, the level of splicing activity in mutant heterozygotes is adequate for normal cell function and, depending on other as yet undefined factors, may be sufficient in rod photoreceptors. However, if activity falls below a threshold level and this would appear to occur only in the rod photoreceptors of symptomatic patients, then full RP pathology ensues.

We have shown that the A226P mutant of yeast PRP31p, the yeast homologue of human PRPF31, would appear to have a relatively mild effect on splicing since it is able to partially complement the functional deficiency in a temperature-sensitive PRP31p-deficient strain of yeast. At the permissive temperature of 23°C, no difference in the growth rate of the strain is observed when complemented with the A226P mutant or with the wild-type protein. However, at the non-permissive temperature of 37°C, the A226P mutation shows slower growth than the wild-type. As the temperature increases, yeast generally increases its growth rate, with a consequent increased demand on splicing function. Under these conditions therefore, the A226P mutant is able to only partially satisfy the splicing demand and growth becomes limited. Moreover, full complementation of the defect is only seen under inducing conditions, when the mutant PRP31p is expressed at a high level. Hence the incomplete complementation must be due to a reduction in splicing efficiency rather than any toxic effect of the mutation.
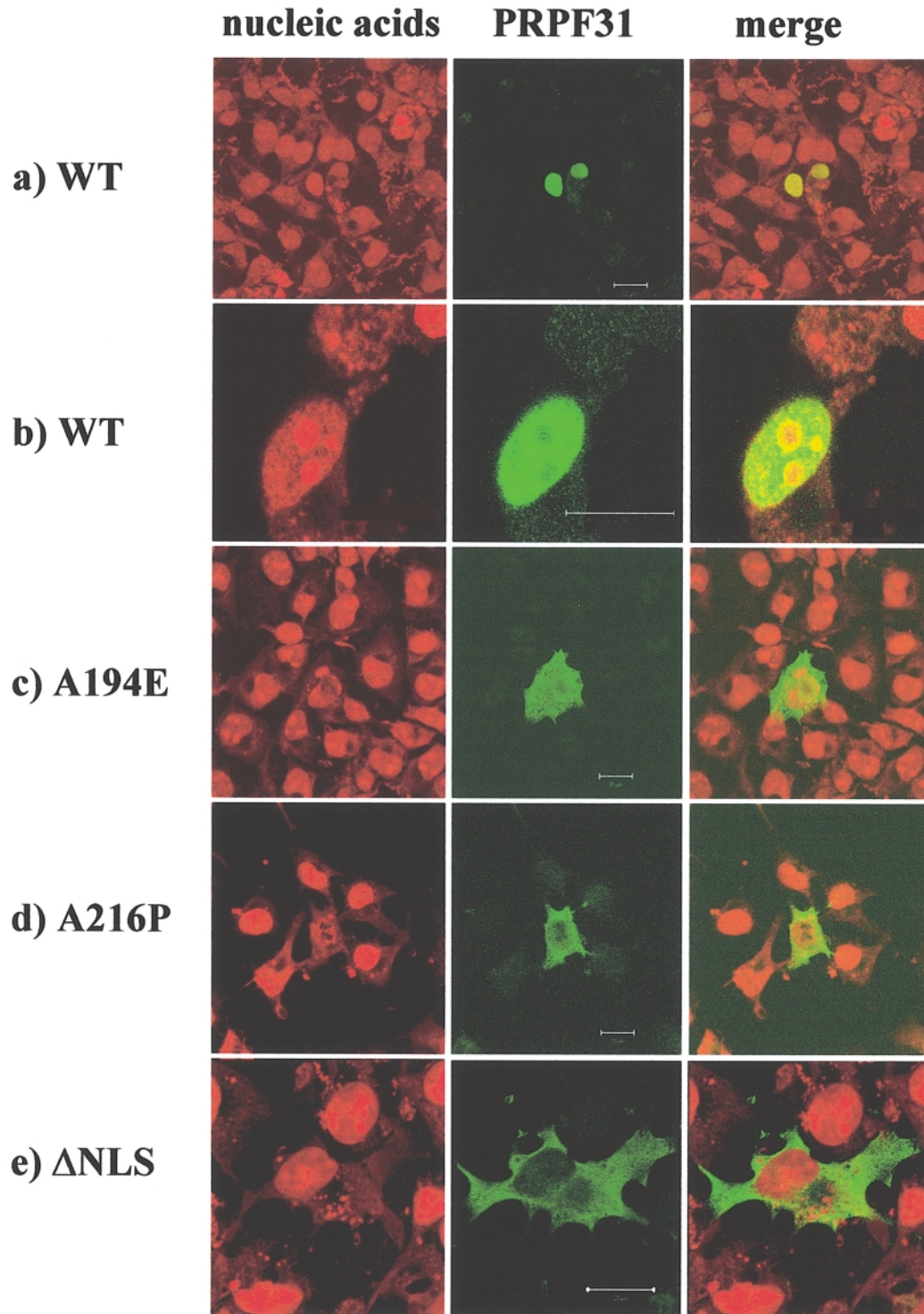
Figure 5. Confocal immunofluorescent images showing cellular localization of His-tagged wild-type and mutant PRPF31 proteins expressed in COS7 cells. Transfected cells were probed with α-His.tag primary antibody and with FITC-conjugated secondary antibody and stained with propidium iodide. Size bars 20 μm.

In the yeast experiment the A226P mutation is effectively present in a homozygous condition. Thus we can deduce that the mutant protein has functional activity, although reduced in comparison to wild-type. It is tempting to extrapolate from this that a corresponding A216P mutation in PRPF31 would also reduce but not eliminate splicing activity. However, the limited sequence similarity between yeast PRP31p and human PRPF31 and known functional differences between the two proteins referred to above mean that such extrapolation should be treated with caution. The differences between the two proteins

are highlighted by the failure of the human protein to complement the PRP31-deficient yeast strain and by the observation that the introduction of the human mutation at position 194 into the yeast protein at this non-conserved position fully complements the temperature-sensitive mutation of PRP31p.

An *in vivo* functional assay of PRPF31 activity that utilized a splicing template of bovine rod opsin requiring the removal of intron 4 was used to assess the effect of high levels of expression of the mutant protein in human HEK 293T cells against a background of endogenous normal PRPF31 activity. The assay was designed to detect the presence of aberrant splicing or failure to splice by the mutant PRPF31. Neither was detected, which is perhaps not surprising since such an effect would almost certainly be lethal in all cells of the body. RP11 disease is dominant, so the mutant protein in patients must take its effect in the presence of normal gene product. Failure to demonstrate any effect on splicing suggests that there is no dominant negative effect on the wild-type protein and that the pathology more likely arises from other causes.

Studies on the sub-cellular localization of PRPF31 protein in transfected mammalian cells show that the A194E and A216P mutations both impede but do not entirely prevent the translocation of protein from its site of synthesis in the cytoplasm to the nucleus. This conclusion is supported by the western analysis, which showed that less mutant than wild-type protein was present in the nucleus and that a higher proportion of the expressed mutant protein was insoluble. This suggests that the mutant proteins may be more susceptible to mis-folding and produce insoluble aggregates in the cytosol. Alternatively the mutants may fold normally but retention at an abnormally high concentration in the cytosol may cause aggregation and loss of solubility. It must be remembered, however, that in the transfected cells the PRPF31 proteins were expressed at abnormally high levels. Whether protein insolubility is a problem when expressed at physiological levels is unknown.

Taken together, these results offer two possible explanations for the rod photoreceptor degeneration in RP11 patients with the A194E or A216P mutations. A net reduction in concentration of functional PRPF31 protein in the nucleus, owing to its retention in the cytoplasm and/or loss of function, for example through inactivation of the Nop domain, could lead to an insufficiency of splicing function. The consequences of this may be a reduction in the concentrations of certain key proteins in the cell. Reduced levels of rhodopsin in the disorganized rod outer segments of mice heterozygous for the targeted deletion of the rod opsin gene (22) have been shown to be detrimental; photoreceptor degeneration in RP11 patients may arise therefore from the long-term effects of reduced levels of rhodopsin in rod outer segment membranes. Our results with the yeast complementation system suggest that such an insufficiency of splicing function might be overcome by protein overproduction. This would make RP11 an exciting proposition for gene therapy.

An alternative explanation for the photoreceptor degeneration is that damage may result from the accumulation over time of insoluble aggregates. Recent studies on protein deposition diseases such as Alzheimer's disease have shown that amorphous aggregates, even if derived from otherwise normally functional protein, can be highly toxic to cells (23,24). However, in the case of PRPF31, the absence of any

apparent detrimental effect in other cells of the body where the protein is expressed makes this a less likely explanation.

The presence in PRPF31 of a classical monopartite NLS indicates that importation into the nucleus is likely to be via the importin/Ran.GTP pathway (20). Work is now in progress to investigate the effect of mutations A194E and A216P on the binding of PRPF31 to importin.

MATERIALS AND METHODS

Oligonucleotides used as primers were purchased from Sigma-Genosys. The nucleotide sequences of all primers used in this study may be obtained on application to the authors. Except where otherwise stated, all chemicals and reagents were purchased from Sigma-Aldrich.

Plasmid construction

Constructs for expression of human PRPF31 in mammalian cell lines. The complete cDNA sequence of *PRPF31* was amplified from human retinal cDNA (human retina quick-clone cDNA from Clontech) using HotStarTaq DNA polymerase (Qiagen) and primers 5′PRP31 (*Bsp*HI-tagged) and 3′PRP31-His (*Xho*I-tagged) such that the C-terminus of PRPF31 was cloned in-frame with the eight histidines of pTriEx-1 (Novagen). This was achieved by first cloning into pGEM.Teasy (Promega) and subsequently transferring the *Bsp*HI–*Xho*I fragment into the expression vector pTriEx-1 via the *Nco*I and *Xho*I sites to generate the plasmid pTriEx.PRPF31^{His}. Sequencing of this plasmid insert revealed a synonymous difference (G506A) compared with the GenBank submitted sequence, AL050369 and a coding difference (G765A) resulting in a V244M substitution. The V244M difference was also found in the sequence amplified from normal human genomic DNA in this laboratory, so it was assumed that G765A is a database error.

Constructs for expression of human PRPF31 and S. cerevisiae PRP31p in yeast. In order to study the complementation of the yeast temperature-sensitive *PRP31* mutant strain (JWY2857, gift of Dr John L. Woolford), the human and yeast *PRP31* coding sequences were cloned into pMyr (Stratagene). The pMyr plasmid was digested with *Xho*I and partially digested with *Spe*I (such that only the site at position 16 was cut). The human *PRPF31* cDNA was reamplified with the primers 5′PRP31H (*Spe*I-tagged) and the 3′PRP31H-stop (*Xho*I-tagged) such that the PRPF31 protein was expressed up to its natural stop codon to generate the pMyr-human PRPF31. A second construct was made by amplification from pTriEx.PRPF31^{His} to include the polyhistidine tail, achieved using the 5′PRP31H (*Spe*I-tagged) and the CAGDOWN primer (Novagen) with *Proof start* DNA polymerase (Qiagen). This fragment was cloned into the pMyr vector cut with *Xho*I, rendered blunt-ended and *Spe*I (cut only once at position 16). This His-tagged construct, pMyr-human PRPF31^{His}, allowed the expression of the human PRPF31 protein in yeast to be monitored by western blotting.

The yeast *PRP31p* was amplified from *S. cerevisiae* genomic DNA (Sigma) with the primers 5′PRP31Y (*Spe*I tagged) and

3′PRP31Y (*Xho*I tagged), and cloned into pMyr to give the construct pMyr–yeast PRP31. Two changes were found in the yeast sequence, I60V and V440A.

Construct for expression of splicing template.   The sequence of bovine rod intron 4 was PCR amplified from bovine genomic DNA (Sigma) using *Pfu* turbo polymerase (Stratagene) and the primers E4F (*Apa*I-tagged) and E5R (*Sal*I-tagged). After cloning into pGEM.Teasy (Promega), the DNA sequence was obtained by sequencing in both directions. The sequence was then excised from the pGEM.Teasy vector as an *Apa*I–*Sal*I fragment and used to replace the *Apa*I–*Sal*I fragment in pMT4 (25). pMT4 is a eukaryotic expression construct containing a synthetic bovine rhodopsin cDNA under the control of the adenovirus major late promoter. The resulting construct was named pMT4i4.

Generation of site-directed mutants

Single-point mutations were introduced into human *PRPF31* in the pGEM-Teasy vector using the GeneEditor site directed mutagenesis kit (Promega) using primers A194E and A216P. A deletion mutant in which the nuclear localization sequence between nucleotides 1051 and 1092 was removed (ΔNLS) was obtained from the same template using a QuickChange site-directed mutagenesis kit (Stratagene) and primers ΔNLS-forward and ΔNLS-reverse. After verification by DNA sequencing, the *PRPF31* fragments were transferred to pTriEx-1 as before.

The *Aat*II/*Eco*RI fragment of the yeast *PRP31* was cloned into pGEM-Teasy and used as a template for the L204E and A226P mutagenesis using the GeneEditor site-directed mutagenesis kit (Promega). After ensuring that the correct base changes had been made, the mutated fragments were reintroduced into the pMyr–yeast *PRP31* plasmid and the entire gene was sequenced.

Complementation of the temperature-sensitive, PRP31p-deficient yeast strain

Yeast transformations were performed following the lithium acetate procedure (26). The temperature-sensitive JWY2857 strain was grown in YPD medium (Sigma) at 23°C. After transformation with the pMyr derived constructs, the transformants were grown at 23°C on agar containing yeast nitrogen base without amino acids (Difco) plus yeast synthetic drop-out medium supplement (without uracil) and 2% glucose. The pMyr vector contains the *GAL1* promoter inducible by galactose. The strains were subsequently streaked in duplicate on the same medium with different carbon sources, either 2% glucose, or 0.25% glucose and 2% galactose, or 1% raffinose and 2% galactose, and incubated at 23 or 37°C.

Expression of wild-type and mutant human His-tagged PRPF31 in HEK 293T cells

Constructs were transiently transfected into HEK 293T cells (ECACC) in 90 mm diameter tissue culture dishes using Lipofectamine (Gibco BRL). To ensure a similar transfection efficiency for all samples, transfection with wild-type or mutant constructs was performed at the same time using aliquots of the same cells. Cells were harvested by scraping 48 h later, washed twice with PBS and pelleted gently by centrifugation. The cell pellet from one dish of cells was resuspended in 1 ml buffer A (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT, 0.5 mM PMSF, 0.25 mM benzamidine) and incubated on ice for 10 min. The cells were re-pelleted by spinning for 10 min in a microfuge at 2000 rpm at 4°C and then resuspended in 300 µl buffer A. For each preparation a 100 µl aliquot was removed for analysis of whole cell proteins. NaCl was added to this aliquot to a concentration of 1 M to lyse all membranes and release the proteins. To the remaining 200 µl of cell suspension detergent Nonidet P-40 was added to a concentration of 0.2% and the mixture was incubated on ice for 10 min or until all cell membranes were lysed, as checked by microscope. The suspension was spun at 3500 rpm for 10 min at 4°C and the supernatant containing soluble cytosolic proteins was removed. The pellet containing nuclei and insoluble components was resuspended in 200 µl buffer B (5 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF, 0.25 mM benzamidine) and NaCl was added to a concentration of 1 M. After incubation for 45 min at 4°C, nuclear lysis occurred as demonstrated by a marked increase in viscosity. The viscosity was reduced by shearing the DNA by passage 10 times through a 25 gauge needle. The suspension was then spun in a microfuge at 13 000 rpm for 40 min at 4°C and the supernatant containing soluble nuclear proteins was preserved. The protein concentration of all extracts was determined using a BioRad protein assay kit with BSA as a standard.

Western analysis of protein extracts

Aliquots containing equal amounts of total proteins were electrophoresed in 12% SDS–polyacrylamide gels in duplicate. One gel was stained with Coomassie blue stain. The proteins on the other were transferred by electroblotting to Zetaprobe membrane (BioRad) using methanol-free transfer buffer (50 mM Tris pH 9.1, 390 mM glycine, 0.04% SDS). Blots were blocked with 5% (w/v) milk powder proteins in PBS and then probed with an α-His.tag monoclonal antibody (Novagen; 100 ng/ml). After washing the blots were probed with a horseradish peroxidase-conjugated α-mouse secondary antibody (Jackson Immunoresearch Laboratories; 160 ng/ml). Immunoreactive protein was detected using enhanced chemiluminescence (National Diagnostics).

In vivo splicing assay in HEK 293T cells

HEK 293T cells in 3.5 cm dishes were co-transfected with equal amounts of pMT4i4 and wild-type or mutant pTriEx.PRPF31[His] using Lipofectamine (Gibco BRL). Forty-eight hours later, cells were washed twice with PBS and used to prepare total cell RNA using TRIzol reagent (Gibco BRL). Cells were lysed directly in the dish with 1 ml TRIzol containing 250 µg/ml glycogen per dish of cells. After incubation with chloroform and phase separation, isopropanol was added to the aqueous phase to precipitate the RNA. The RNA was pelleted by centrifugation in a microfuge at 12 000 rpm for 10 min at 4°C, rinsed in 70% aqueous ethanol

and re-dissolved in 20 μl cDNA synthesis buffer (10 mM Tris pH 8.3, 50 mM KCl, 2.5 mM MgCl$_2$). Contaminating DNA carried over from the cell transfection was then removed by treatment with RNase-free DNase1 (Roche Molecular Biochemicals) at 37°C for 15 min and the DNase1 was inactivated at 75°C for 10 min. First-strand cDNA synthesis was then performed using AMV reverse transcriptase and oligo-p(dT)$_{15}$ primer (both Roche Molecular Biochemicals) by incubation first at room temperature for 10 min, then at 42°C for 1 h, followed by enzyme inactivation at 99°C for 5 min. This cDNA was then assayed for the presence of spliced, aberrantly spliced and unspliced transcripts derived from the splicing template construct (pMT4i4) using a PCR-based approach. The primer combination E4F/E5R was used to probe for the presence of spliced, aberrantly spliced and unspliced transcripts, whilst the combination In4F/E5R was used to probe for the presence of unspliced transcripts only. As an internal control for total RNA in each assay, PCRs were also conducted using primers for the housekeeping gene *PGM1* (PGM1F and PGM1R) (18). Thermocycling was conducted for 24 cycles (E4F/E5R), 28 cycles (In4F/E5R) and 30 cycles (PGM1F/PGM1R), respectively, to ensure amplification was still in the exponential (not plateau) phase, as verified by varying cycle number. PCR products were analysed on ethidium bromide-stained agarose gels.

Immunofluorescence analysis of transfected COS7 cells by confocal microscopy

COS7 cells were plated on to glass coverslips and transfected with wild type or mutant pTriEx.PRPF31^{His} constructs using the method specified above. After 48 h, the cells were washed with PBS, fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X-100 and blocked with 5% fetal calf serum (First Link). They were then probed with α-His.tag monoclonal antibody (Novagen; 100 ng/ml). After repeated washing with PBS, the cells were re-blocked with 5% goat serum (Sigma) and then incubated with a fluorescein isothiocyanate-conjugated α-mouse secondary antibody (Jackson ImmunoResearch Laboratories; 30 μg/ml). After washing the cells were stained for nucleic acids with propidium iodide (PrI; Sigma) and mounted in DAKO fluorescent mounting medium. Cells were viewed using a Zeiss LSM510 confocal fluorescent microscope.

REFERENCES

1. Chakarova, C.F., Hims, M.M., Bolz, H., Abu-Safieh, L., Patel, R.J., Papaioannou, M.G., Inglehearn, C.F., Keen, T.J., Willis, C., Moore, A.T. *et al.* (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.
2. McKie, A.B., McHale, J.C., Keen, T.J., Tarttelin, E.E., Goliath, R., van Lith-Verhoeven, J.J.C., Greenberg, J., Ramesar, R.S., Hoyng, C.B., Cremers, F.P.M. *et al.* (2001) Mutations in the pre-mRNA splicing factor gene *PRPC8* in autosomal dominant retinitis pigmentosa (RP13). *Hum. Mol. Genet.*, **10**, 1555–1562.
3. Vithana, E.N., Abu-Safieh, L., Allen, M.J., Carey, A., Papaioannou, M., Chakarova, C., Al-Maghtheh, M., Ebenezer, N.D., Willis, C., Moore, A.T. *et al.* (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.
4. Moore, A.T., Fitzke, F., Jay, M., Arden, G.B., Inglehearn, C.F., Keen, T.J., Bhattacharya, S.S. and Bird, A.C. (1993) Autosomal dominant retinitis pigmentosa with apparent incomplete penetrance: a clinical, electrophysiological, psychophysical and molecular genetic study. *Br. J. Ophthal.*, **77**, 473–479.
5. Evans, K., Al-Maghtheh, M., Fitzke, F.W., Moore, A.T., Jay, M., Inglehearn, C.F., Arden, G.B. and Bird, A.C. (1995) Bimodal expressivity in dominant retinitis pigmentosa genetically linked to chromosome 19q. *Br. J. Ophthal.*, **79**, 841–846.
6. Will, C.L. and Luhrmann, R. (2001) Spliceosomal UsnRNP biogenesis, structure and function. *Curr. Opin. Cell Biol.*, **13**, 290–301.
7. Teigelkamp, S., Achsel, T., Mundt, C., Gothel, S.F., Cronshagen, U., Lane, W.S., Marahiel, M. and Luhrmann, R. (1998) The 20 kD protein of human [U4/U6.U5] tri-snRNPs is a novel cyclophilin that forms a complex with the U4/U6-specific 60 kD and 90 kD proteins. *RNA*, **4**, 127–141.
8. Makarova, O.V., Makarov, E.M., Liu, S., Vornlocher, H-P. and Luhrmann, R. (2002) Protein 61 K, 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.
9. Weidenhammer E.M., Singh, M., Ruiz-Noriega, M., and Woolford, J.L. Jr. (1996) The *PRP31* gene encodes a novel protein required for pre-mRNA splicing in *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **24**, 1164–1170.
10. Bishop D.T., McDonald, W.H., Gould, K.L. and Forsburg, S.L. (2000) Isolation of an essential *Schizosaccharomyces pombe* gene, *PRP31+*, that links splicing and meiosis. *Nucleic Acids Res.*, **28** 2214–2220.
11. Bateman, A., Birney, E., Durbin, R., Eddy, S.R., Howe, K.L. and Sonnhammer, E.L. (2000) The Pfam protein families database. *Nucleic Acids Res.*, **28**, 263–266.
12. Weidenhammer, E.M., Ruiz-Noriega, M. and Woolford, J.L. Jr. (1997) Prp31p promotes the association of the U4/U6.U5 tri-snRNP with prespliceosomes to form spliceosomes in Saccharomyces cerevisiae. *Mol. Cell Biol.*, **17**, 3580–3588.
13. Maddock, J.R., Roy, J. and Woolford, J.L. (1996) Six novel genes necessary for pre-mRNA splicing in *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **24**, 1037–1044.
14. Miyazaki, J., Takaki, S., Araki, K., Tashiro, F., Tominaga, A., Takatsu, K. and Yamamura, K. (1989) Expression vector system based on the chicken beta-actin promoter directs efficient production of interleukin-5. *Gene*, **79**, 269–277.
15. Young, R.W.J. (1967) The renewal of photoreceptor cell outer segments. *J. Cell Biol.*, **33**, 61–72.
16. Korenbrot, J.I. and Fernald, R.D. (1989) Circadian rhythm and light regulate opsin mRNA in rod photoreceptors. *Nature*, **337**, 454–457.
17. Von Schantz, M., Lucas, R.J. and Foster, R.G. (1999) Circadian oscillation of photopigment transcript levels in the mouse retina. *Brain Res. Mol. Brain Res.*, **72**, 108–114.
18. Whitehouse, D.B., Putt, W., Lovegrove, J.U., Morrison, K., Hollyoake, M., Fox, M.F., Hopkinson, D.A. and Edwards, Y.H. (1992) Phosphoglucomutase 1: complete human and rabbit mRNA sequences and direct mapping of this highly polymorphic marker on human chromosome 1. *Proc. Natl Acad. Sci. USA*, **89**, 411–415.
19. Sleeman, J.E. and Lamond, A.I. (1999) Nuclear organisation of pre-mRNA splicing factors. *Curr. Opin. Cell Biol.*, **11**, 372–377.
20. Jans, D.A., Xiao, C-Y. and Lam, M.H.C. (2000) Nuclear targeting signal recognition: a key control point in nuclear transport? *BioEssays*, **22**, 532–544.

21. Chelsky, D., Ralph, R. and Jonak, G. (1989) Sequence requirements for synthetic peptide-mediated translocation to the nucleus. *Mol. Cell. Biol.*, **9**, 2487–2492.

22. Humphries, M. M., Rancourt, D., Farrar, G. J., Kenna, P., Hazel, M., Bush, R. A., Sieving, P. A., Sheils, D. M., McNally, N., Creighton, P. *et al.* (1997) Retinopathy induced in mice by targeted disruption of the rhodopsin gene. *Nat. Genet.*, **15**, 216–219.

23. Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C.M. and Stefani, M. (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*, **416**, 507–511.

24. Walsh, D.M., Klyubin, I., Fadeeva, J.V., Cullen, W.K., Anwyl, R., Wolfe, M.S., Rowan, M.J. and Selkoe, D.J. (2002) Naturally secreted oligomers of amyloid β protein potently inhibit hippocampal long-term potentiation *in vivo*. *Nature*, **416**, 535–538.

25. Franke, R.R., Sakmar, T.P., Oprian, D.D. and Khorana, H.G. (1988) A single amino acid substitution in rhodopsin (K248→L) prevents activation of transducin. *J. Biol. Chem.*, **263**, 2119–2122.

26. Ito H., Fukuda Y., Murata K. and Kimura A. (1983). Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.*, **153**, 163–168.