



# PRCC, the commonest TFE3 fusion partner in papillary renal carcinoma is associated with pre-mRNA splicing factors

YM Skalsky<sup>\*1</sup>, PM Ajuh<sup>2</sup>, C Parker<sup>1</sup>, AI Lamond<sup>2</sup>, G Goodwin<sup>1</sup> and CS Cooper<sup>1</sup>

<sup>1</sup>Institute of Cancer Research, Molecular Carcinogenesis Section, The Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey, SM2 5NG, UK; <sup>2</sup>Department of Biochemistry, The University of Dundee, MSI/WTB Complex, Dow Street, Dundee, DD1 5EH, UK

**In papillary renal cell carcinomas the TFE3 transcription factor becomes fused to the PSF and NonO pre-mRNA splicing factors and most commonly to a protein of unknown function designated PRCC. In this study we have examined the ability of the resulting PRCC–TFE3 and NonO–TFE3 fusions to activate transcription from the plasminogen activator inhibitor-1 (PAI-1) promoter. The results show that only fusion to PRCC enhanced transcriptional activation, indicating that the ability to enhance the level of transcription from endogenous TFE3 promoters is not a consistent feature of TFE3 fusions. In investigations of the normal function of PRCC we observed that PRCC expressed as a green fluorescent fusion protein colocalizes within the nucleus with Sm pre-mRNA splicing factors. It was also found that endogenous PRCC is coimmunoprecipitated by antibodies that recognize a variety of pre-mRNA splicing factors including SC35, PRL1 and CDC5. Association with the cellular splicing machinery is therefore, a common feature of the proteins that become fused to TFE3 in papillary renal cell carcinomas. *Oncogene* (2001) 20, 178–187.**

**Keywords:** TFE3; PRCC; transcription activation; pre-mRNA splicing

## Introduction

A recurrent translocation involving Xp11.2 is associated with papillary renal cell carcinomas. Several variants of this translocation have been identified including t(X;1)(p11.2;q21), t(X;1)(p11.2;p34) and inv(X)(p11.2;q12). In some cases alteration of Xp11.2 occur as the only cytogenetic abnormality (Meloni *et al.*, 1993; De Jong *et al.*, 1986), suggesting that the genes involved in these translocations play an important role in the process of cellular transformation. Cloning of the breakpoints has revealed that the TFE3 gene at Xp11.2 is fused to either PRCC chromosome 1q21, to PSF at 1p34 or to NonO at Xq12 (Sidhar *et al.*, 1996; Weterman *et al.*, 1996; Clark *et al.*, 1997). Since all cases examined to date were

shown to contain the PRCC–TFE3, PSF–TFE3 or NonO–TFE3 fusions, but not always reciprocal fusions, they are the protein structures that believed to be involved in carcinogenesis.

TFE3 is a transcription factor that was originally identified by its ability to bind the  $\mu$ E3 site in the IgH enhancer (Beckmann *et al.*, 1990). The 576aa protein contains a basic helix–loop–helix domain involved in DNA binding and two transcriptional activation domains: an N-terminal acidic activation domain and a C-terminal proline rich activation domain (Artandi *et al.*, 1995). A leucine zipper region is responsible for dimerization (Roman *et al.*, 1992), homodimerization, heterodimerization with the other family members, such as TFE2, TFEB, TFEC, the Microphthalmia protein (Hemesath *et al.*, 1994), and for binding with PU.1 of the ETS protein family (Tian *et al.*, 1999). A TFE3 target gene, which has been recently identified, is the plasminogen activator inhibitor-1 gene (PAI-1) (Hua *et al.*, 1998), which itself has been linked with oncogenesis/metastasis (Bajou, 1998). TFE3 acts synergistically with SMAD3, which binds to an adjacent site in the PAI-1 promoter in the mediation of TGF $\beta$  induced activation of PAI-1 (Hua *et al.*, 1999).

PSF is a ubiquitously expressed, essential pre-mRNA splicing factor (Patton *et al.*, 1993). The full-length protein has 712 amino acids, which contain an N-terminal sequence rich in proline and glutamines and two consensus RNA binding domains. The PSF–TFE3 fusion protein contains almost all of PSF, fused to the C-terminal portion of TFE3. The second partner, NonO (also called p54<sup>nrB</sup>) is the mammalian equivalent of *Drosophila* visual and courtship song behaviour protein No-on-transient A (Yang *et al.*, 1993). NonO has significant homology with PSF over a 320aa region, encompassing the RNA binding domains (Dong *et al.*, 1993). This striking degree of homology suggested that NonO was also a splicing factor and indeed it has been shown that its depletion from nuclear extract inhibits  $\beta$ -globin splicing *in vitro* (Hallier *et al.*, 1998). NonO may have a role in transcriptional activation (Basu *et al.*, 1997) and was also shown to bind PU.1 an ETS family protein that can alter the splicing activity of NonO *in vitro* (Hallier *et al.*, 1996).

PRCC codes for a protein of 491 amino acids with a proline, glycine and leucine rich N-terminal region. For example, out of its 156 N-terminal amino acids 27%

\*Correspondence: YM Skalsky

Received 21 June 2000; revised 24 October 2000; accepted 24 October 2000

are proline. PRCC is not homologous to any known protein and does not possess motifs indicative of biochemical function. Recent studies have, however, shown that it contains domains capable of transcriptional activation (Weterman *et al.*, 2000). Analysis of six cases of renal cell carcinoma with PRCC–TFE3 fusion genes has demonstrated two distinct breakpoints in PRCC (Sidhar *et al.*, 1996; Weterman *et al.*, 1996). In five of the six cases, the fusion protein contained 156 amino acids of N-terminal PRCC fused to TFE3, while in the sixth case, 393 amino acids of N-terminal PRCC were fused to TFE3.

In this study we have undertaken a search for common features of the TFE3 fusions which could explain their oncogenic activity. Initial investigations indicated that the ability to activate transcription from endogenous TFE3 promoters was not a common factor of TFE3 fusions. Our studies were therefore extended to examine the possibility that PRCC, in common with NonO and PSF, may be a component of the pre-mRNA splicing machinery.

## Results

### *Activation of transcription*

It has recently been proposed, based on studies on the PRCC–TFE3 fusion product that a key feature of TFE3 fusion partners is that they have the ability to enhance levels of transcription directed by TFE3 (Weterman *et al.*, 2000). To assess whether this feature is common to other TFE3 fusion proteins we compared the transcriptional activation properties of PRCC–TFE3 and NonO–TFE3. All proteins were expressed as enhanced green fluorescent protein (EGFP) fusions so that the subcellular location of the protein could be examined in parallel studies. The assay used relied upon the fact that the TFE3 sequences present in PRCC–TFE3 and NonO–TFE3 fusions retain the TFE3 DNA binding domain. To monitor the ability of each protein to activate transcription we utilized a promoter sequence from the plasminogen activator inhibitor type I (PAI-1) gene fused upstream of the luciferase reporter gene (PE2 reporter). The PAI-1 promoter contains adjacent TFE3 binding domain (E-box) and SMAD3 binding sites and therefore this construct can also be used to monitor interactions between TFE3 and SMAD3 in response to signalling through the TGF- $\beta$ /SMAD pathway. All results were normalized for transfection efficiency and level of protein expression, determined by Western blot analysis using anti-GFP and anti-myc antibodies raised against GFP fusions and Myc-tagged SMAD3.

Comparisons of EGFP–TFE3, EGFP–PRCC–TFE3 and EGFP–NonO–TFE3 demonstrated (Figure 1a) that EGFP–TFE3 is a stronger transcriptional activator than EGFP–NonO–TFE3 and both are weaker than EGFP–PRCC–TFE3 which is 2.5–3-fold more active than EGFP–TFE3 and 5–6-fold more

active than EGFP–NonO–TFE3. In fact EGFP–NonO–TFE3 failed to enhance activation when compared to background level measured in cells transfected with PE2 reporter and GFP alone. Similar relative levels of transactivation were observed when assays were performed following TGF- $\beta$  exposure. Cotransfection of SMAD3 further enhanced transcriptional activation both before and after TGF- $\beta$  treatment for EGFP–PRCC–TFE3 and EGFP–TFE3 but interestingly not for EGFP–NonO–TFE3 which again failed to enhance activation when compared to background level measured in cells transfected with PE2 reporter, GFP alone and SMAD3. In these experiments EGFP–PRCC–TFE3 showed twofold greater activity than EGFP–TFE3, and fivefold greater activity than EGFP–NonO–TFE3. All of these activities including the effects of TGF- $\beta$  were dependent on the presence of an intact TFE3 binding site since mutations in the E-box sequence from CACGTG to ACCGAC, which abolishes TFE3 binding, ablated transcriptional activation in all experiments.

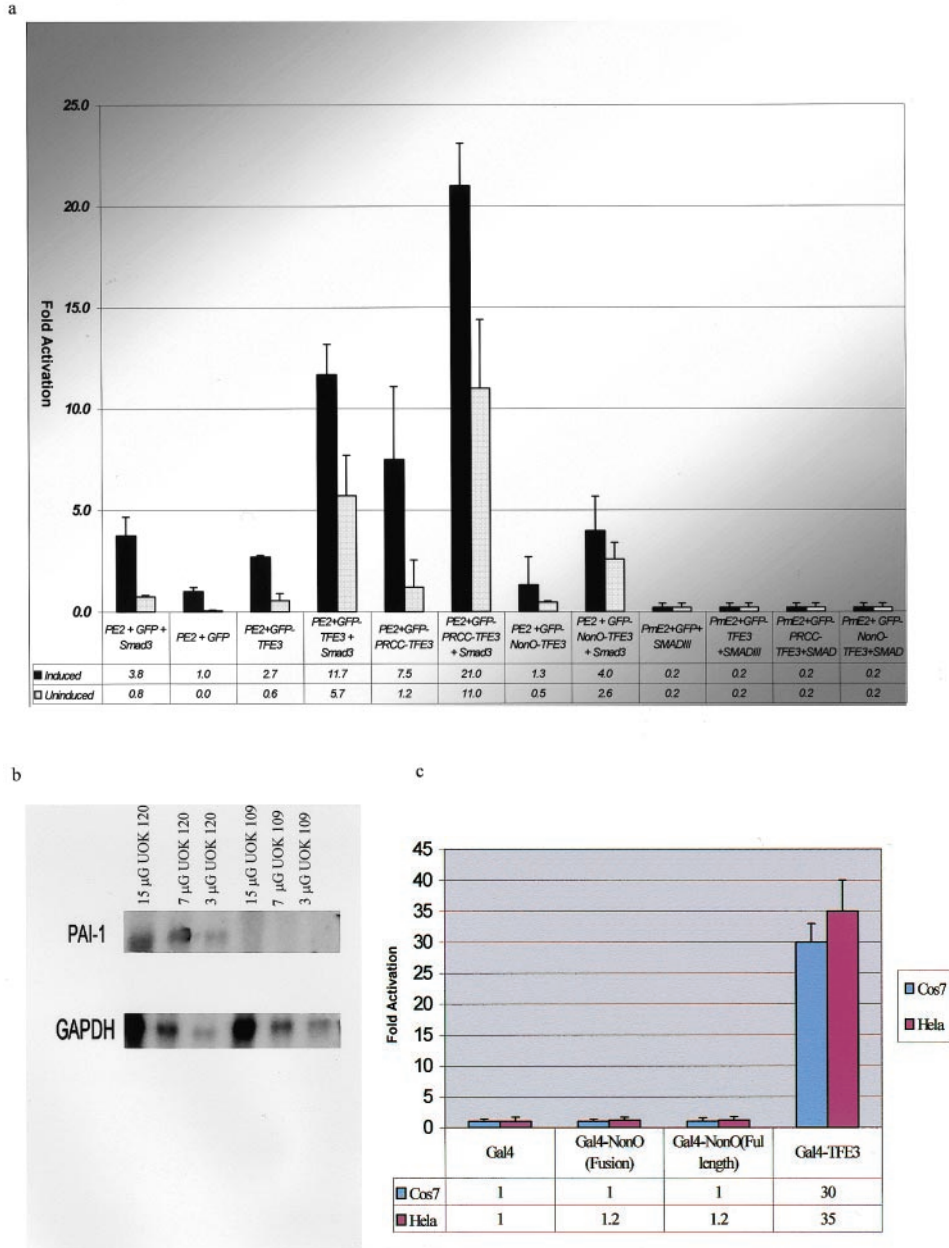
The relative level of PAI-1 expression in papillary renal cell carcinomas cell lines PRCC–TFE3(UOK 120), and NonO–TFE3(UOK 109) was measured by Northern blot analysis (Figure 1b) to confirm the results obtained by the TFE3 specific reporter assay (Figure 1a). Neither UOK120 nor UOK109 harbour normal copy of the TFE3 gene and therefore any TFE3 related changes on the transcription of PAI-1 must be TFE3 fusion related. Our results showed that high levels of PAI-1 transcripts are present in UOK 120, but almost undetectable in UOK 109 as predicted by the transactivation results. Equal levels of RNA loading were demonstrated by the use of a GAPDH probe. These results were further confirmed by quantitative RT–PCR analysis using oligo dT, the random hexamer and PAI-1 specific primer for RT followed by PCR using PAI-1 specific primers (data not shown). To further support the results presented in Figure 1a showing that the fusion NonO–TFE3 failed to enhance activation of the PE2 reporter, a dual luciferase reporter assay was used to verify whether Gal4 fusions of full length or the translocated region of NonO could activate transcription of a reporter plasmid (Figure 1c). In these experiments NonO could not activate transcription neither in HeLa nor in HepG2 cells which is in agreement with the results presented in Figure 1a. The fusion Gal4–TFE3 was used as a positive control showing a marked transcriptional activity.

### *Localization of EGFP–PRCC, EGFP–NonO and EGFP–TFE3*

In the search for features common to the TFE3 fusion partners we compared the location of PRCC and NonO each expressed as an EGFP fusion in Cos7 cells. In all experiments we also compared the localization of the GFP fusion protein with the localization of the endogenous Sm pre-mRNA splicing

protein factors detected using Y12 antibodies. The Y12 antibody specifically detects polypeptides B and D of the Sm proteins which are components of the

U1, U2, U4, U5 and U6 snRNP's snRNAs (Pettersson *et al.*, 1984). Figure 2a and d show the distribution respectively of the EGFP-PRCC and



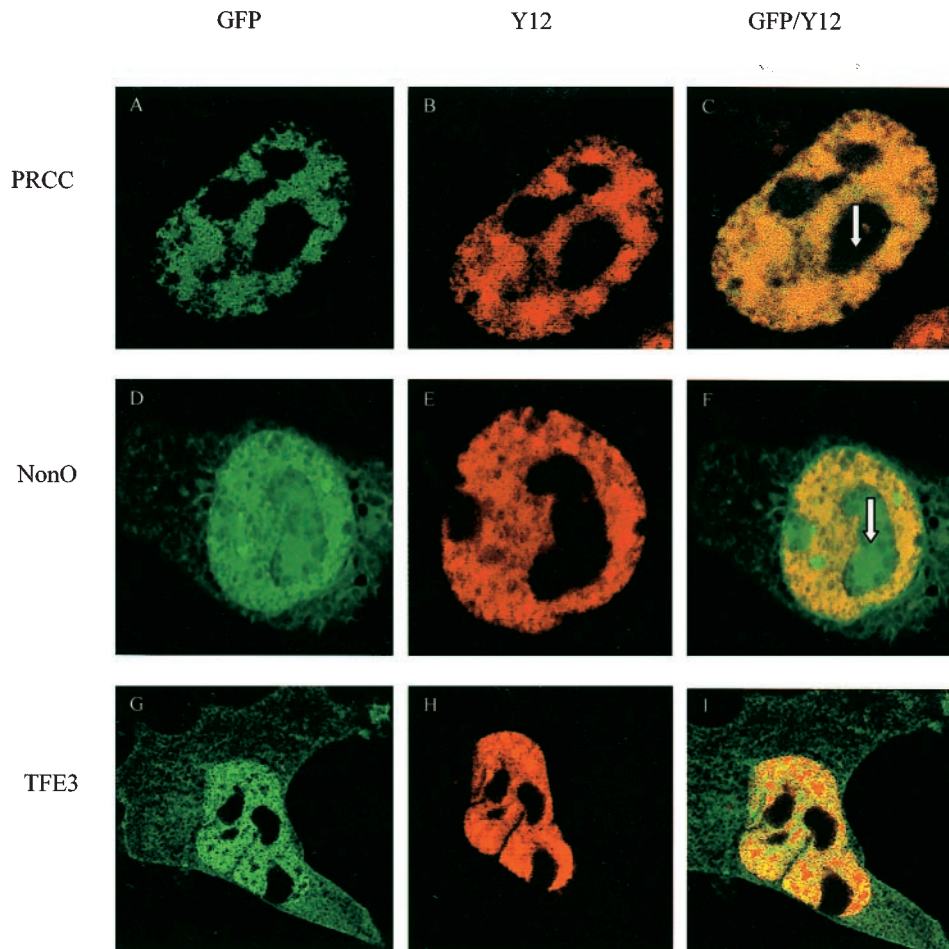
**Figure 1** (a) Transactivation by TFE3, PRCC-TFE3 and NonO-TFE3, in TFE3 specific reporter assay in HepG2 cells. Full length cDNA's of TFE3, PRCC-TFE3 and NonO-TFE3 were cloned in frame with GFP and tested for their ability to (a) activate transcription and (b) synergize with SMAD3 in presence or absence of TGF $\beta$ . The reporters used contained a region of the PAI-1 promoter that has a wild type TFE3 binding site also termed E-box next to a SMAD3 binding site, upstream from the luciferase reporter gene (PE2). Fold activation is an average of seven different experiments and is expressed relative to that of the control GFP vector. Western blot analysis on a fraction of the HepG2 cells pellet using anti-GFP antibody against the GFP fusions and anti-myc antibody against myc-tagged SMAD3 together with cotransfection of renella luciferase were employed to normalize for level of protein expression and transfection efficiency respectively. (b) Northern blot analysis to detect level of PAI-1 expression in cell lines containing the PRCC-TFE3 (UOK120) and NonO-TFE3 (UOK109) fusions. 15, 7 and 3  $\mu$ g's of total RNA were analysed for each cell line. GAPDH bands indicate equal level of RNA loading between cell lines. (c) Transactivation by Gal4 fusions of NonO in HepG2 and HeLa cells. Full-length and translocated cDNA's of NonO were cloned in frame with Gal4 and tested for their ability to activate transcription of a reporter plasmid in a dual luciferase reporter assay. Fold activation is an average of four different experiments and is expressed relative to that of the control Gal4 vector

EGFP-NonO proteins. Both proteins exhibited a nucleoplasmic distribution with some accumulation of protein in speckled structures. However, while EGFP-PRCC was confined to the nucleus, EGFP-NonO was, in addition, found at lower levels throughout the cytoplasm and in the nucleolus. Within the nucleoplasm EGFP-PRCC and EGFP-NonO each colocalized with Y12 antibodies (Figure 2b,c and e,f respectively). These observations are consistent with the known role for NonO in pre-mRNA splicing and provides preliminary evidence that the PRCC protein may also be associated with the cellular splicing machinery.

The EGFP-TFE3 protein exhibited intense fluorescence within the nucleus and was also present at a lower level in the cytoplasm (Figure 2g). Inside the nucleus EGFP-TFE3 was excluded from both the nucleoli and from the structures that stained strongly with the Y12 antibody (Figure 2h,i). EGFP alone did not co-localize with Sm proteins or formed either nuclear structure or specific accumulations (data not shown).

*Treatment with  $\alpha$ -amanatin*

A feature exhibited by splicing factors and snRNPs is that they aggregate in large nuclear speckles when RNA polymerase II transcription is inhibited by treatment with  $\alpha$ -amanatin (Carneo-Fonseca *et al.*, 1992; O’Keefe *et al.*, 1994). When cells expressing EGFP-NonO were treated with  $\alpha$ -amanatin, aggregation of the fluorescent protein into large nuclear speckles is observed (Figure 3d). Interestingly the EGFP-PRCC protein displayed the same behaviour (Figure 3a) and for both EGFP-NonO and EGFP-PRCC colocalization with Y12 antibody was retained following  $\alpha$ -amanatin treatment (Figure 3e,f and b,c respectively). In contrast EGFP-TFE3 did not relocate into large speckles (Figure 3g) following addition of  $\alpha$ -amanatin. Indeed, comparisons with Y12 antibody distribution (Figure 3h) showed that EGFP-TFE3 was completely excluded from the large speckle structures (Figure 3i). These results provide additional support for the view that PRCC may represent a component of the splicing machinery.



**Figure 2** Cellular localization of Sm proteins and GFP-fusions in Cos7 cells. Endogenous Sm proteins were detected by the Y12 antibody (red, centre panel), while PRCC, NonO and TFE3 were detected as EGFP fusions (green, left panel). Combined images of GFP fusions and Sm proteins are shown in the right hand panel. Nucleoli of cells stained with Y12 and transfected with PRCC (c) and NonO (f) are marked with white arrows

GFP on its own shows no response to  $\alpha$ -amanitin treatments or co-localization with Sm protein (data not shown).

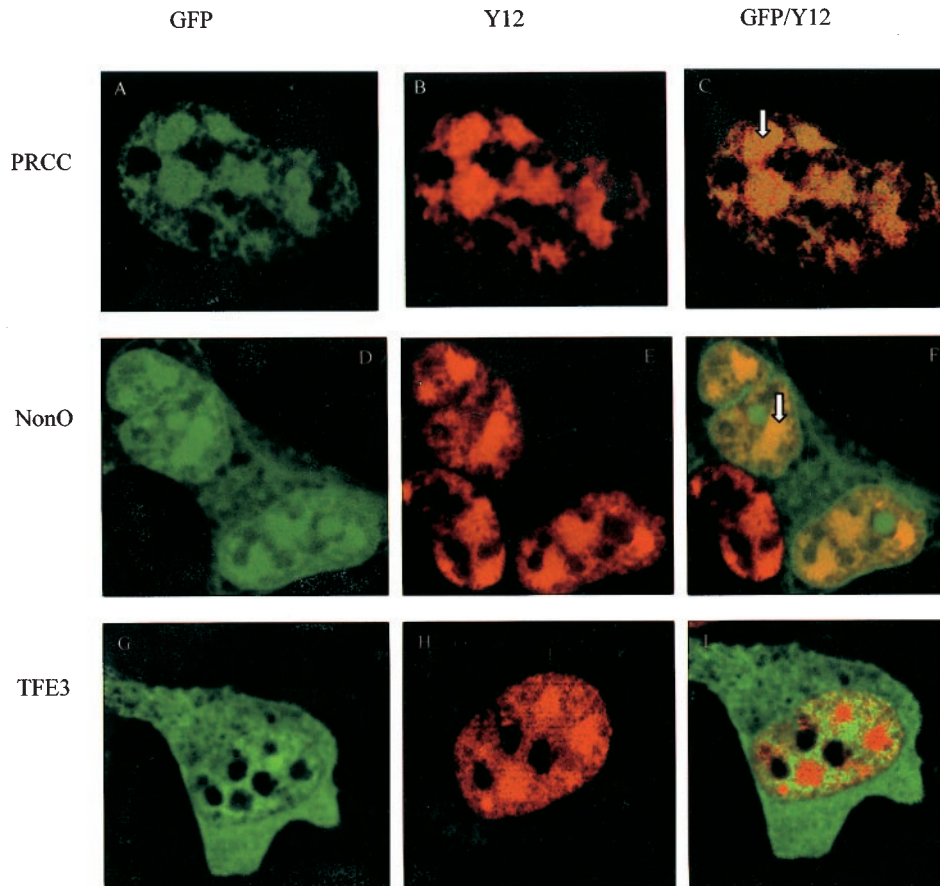
*Association of endogenous PRCC with pre-mRNA splicing proteins*

In order to directly detect endogenous PRCC protein rabbit polyclonal antibodies were raised against GST fusion of the PRCC N-terminal domain. In Western blot analyses of cellular proteins these antibodies detected a 66 kDa PRCC protein in all cell lines examined (Figure 4a). As expected, the antibody also detected abnormal PRCC proteins of 72 and 100 kDa corresponding to PRCC–TFE3 fusion probes respectively in the UOK120 and UOK124 Papillary renal cell lines.

We next used these antibodies to examine directly whether PRCC was associated with splicing machinery. To assay for this association the anti-PRCC antibodies were used to immunoprecipitate PRCC from HeLa cell nuclear extracts and the immunoprecipitated proteins were then subject to Western blot analysis. When the

blots were probed with Y12 antibodies a typical 29 kDa doublet corresponding to the B and D polypeptides of the Sm proteins was observed (Figure 4b). To confirm the coimmunoprecipitation of PRCC and Sm proteins the reciprocal experiment was performed in which proteins were initially immunoprecipitated from HeLa nuclei with the Y12 antibody and Western blots were performed using the anti-PRCC antibody. As expected a 66 kDa band corresponding to the PRCC protein was detected (Figure 4c, lane 1).

In order to further confirm PRCC's interaction with splicing factors, we probed anti-PRCC immunoprecipitates with antibodies to spliceosomal proteins available in our laboratory. The results obtained indicate that PRCC is co-immunoprecipitated by SC-35, hPRL1 and hCDC5 (Figure 4c). The anti-CDC5 antibody detects the human homologue (hCDC5) of the *Schizosaccharomyces pombe* CDC5<sup>+</sup> gene product which is a component of a 40S SnRNP-containing complex and is essential for pre-mRNA splicing (Neubauer *et al.*, 1998; McDonald *et al.*, 1999), anti-PRL1 detects the human homologue (hPRL1) of the *Arabidopsis thaliana* PRL1 gene product which was



**Figure 3** Cellular localization of Sm proteins and GFP-fusions in  $\alpha$ -amanitin treated Cos7 cells. Endogenous Sm proteins were detected by the Y12 antibody (red, centre panel), while PRCC, NonO and TFE3 were detected as EGFP fusions (green, left panel). Combined images of EGFP fusions and Sm proteins are shown in the right hand panel. Some of the cellular regions in which EGFP and Sm proteins colocalized are marked with white arrows

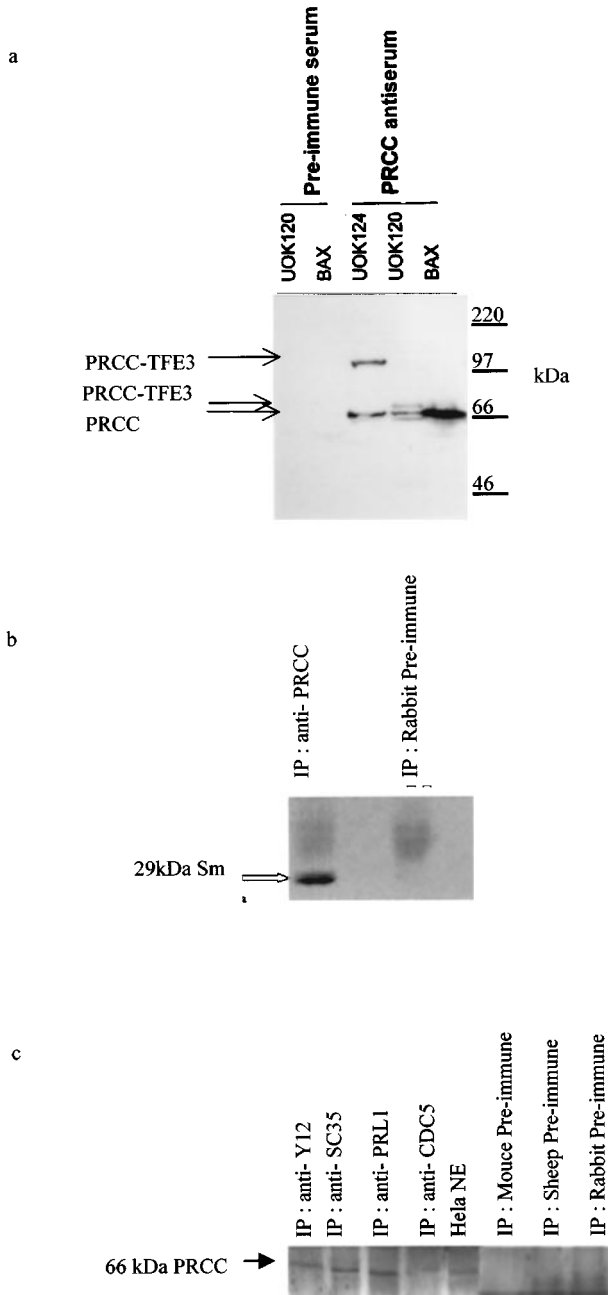
also demonstrated to be a component of the pre-mRNA splicing complex (Neubauer *et al.*, 1998; McDonald *et al.*, 1999). SC35 is member of the SR family of pre-mRNA-splicing factors that play an

important role in the assembly and regulation of the pre-mRNA-splicing complex (Fu, 1995).

*Localization of the EGFP-PRCC-TFE3 and EGFP-NonO-TFE3 fusion proteins*

We have established that EGFP-NonO and EGFP-PRCC both colocalize with Sm pre-mRNA splicing proteins and show that the nuclear distribution of these proteins both before and after  $\alpha$ -amanatin treatment is quite distinct from that of EGFP-TFE3. It was therefore of considerable interest to examine the subcellular localization of the PRCC-TFE3 and NonO-TFE3 chimaeric proteins found in papillary renal carcinoma, again following expression of each protein as a fusion with EGFP (Figure 5a).

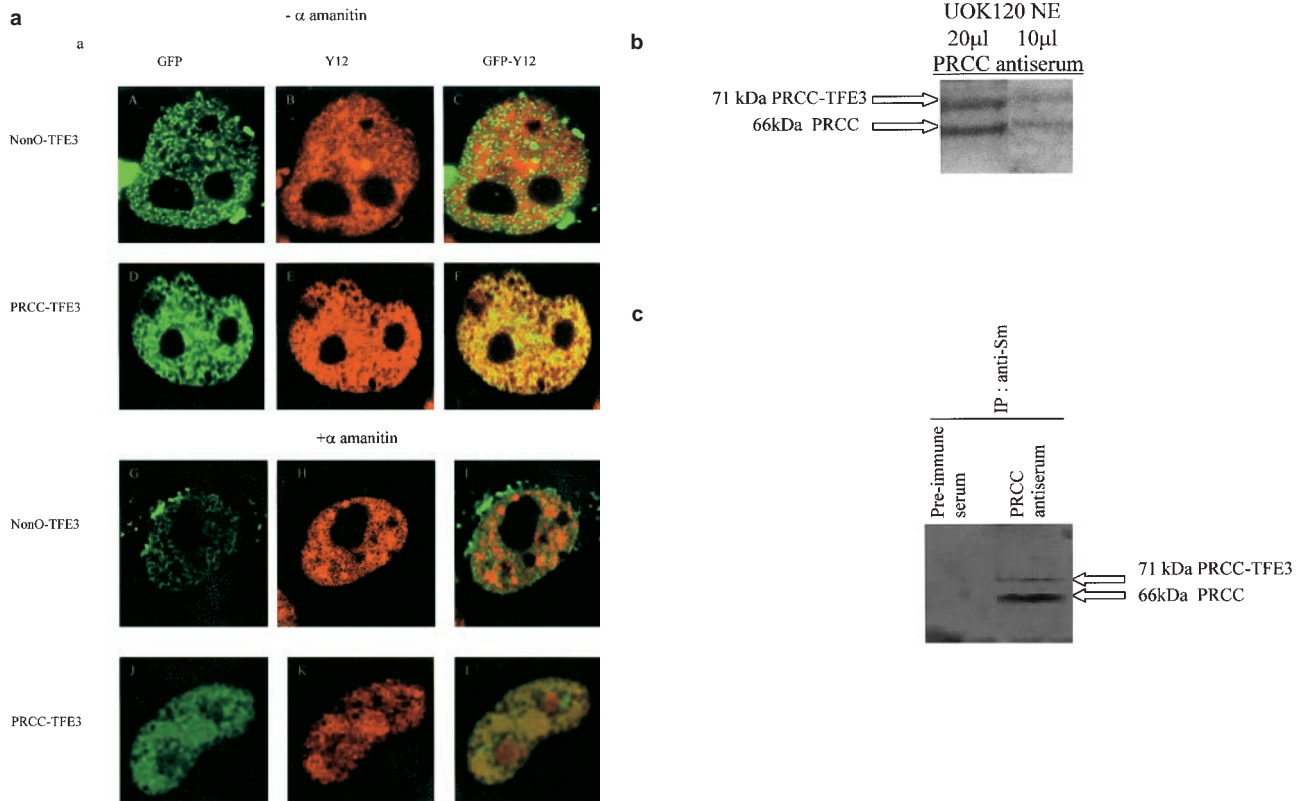
EGFP-NonO-TFE3 formed small aggregates, almost exclusively in the nuclei, although some cytoplasmic staining was also observed in a minority of cells (Figure 5a). This pattern remained unchanged upon treatment of the cells with  $\alpha$ -amanatin (Figure 5g) and both before and after  $\alpha$ -amanatin treatment there was no colocalization with Y12 antibody (Figure 5b,c and h,i respectively). This pattern of localization was unique to EGFP-NonO-TFE3 and quite distinct from that observed for EGFP-TFE3 and for EGFP-NonO. EGFP-PRCC-TFE3 exhibited mostly a nuclear pattern of staining (Figure 5d) but unlike EGFP-NonO-TFE3 some colocalization with the Y12 antibody was still observed (Figure 5f). Following treatment with  $\alpha$ -amanatin in the majority of cells (80%) EGFP-PRCC-TFE3 did not localize with Y12 antibody in large nuclear speckles, although in a minority of cells (20%) it did exhibit partial colocalization (Figure 5l). We conclude from this observation that both EGF-PRCC-TFE3 and EGF-NonO-TFE3 exhibited much less tendency to associate with Sm proteins than their EGF-PRCC and EGF-NonO parent proteins. This conclusion was confirmed by Western blot analysis of UOK120 nuclear extract using anti-PRCC antibodies which shows that PRCC and PRCC-TFE3 proteins are initially present in the nuclear extract at the same level (Figure 5b). Higher level of the fusion PRCC-TFE3 however, is detected by Western blot analysis of Y12-immunoprecipitates from exactly the same nuclear extract (Figure 5c) confirming the loose association between the fusion PRCC-TFE3 and the Sm proteins.



**Figure 4** (a) Western blot analysis using polyclonal rabbit PRCC antiserum and pre-immune serum. Papillary carcinoma lines UOK120, UOK124 and the synovial sarcoma BAX line show the presence of PRCC protein in addition to extra bands corresponding to PRCC-TFE3 fusions in UOK120 and UOK124. No proteins are detected with the pre-immune serum. (b) Immunoprecipitation of Sm proteins (Y12) from HeLa cell nuclear extract using anti-PRCC antibodies. The 29 kDa band corresponding to the Sm protein is indicated by arrow. (c) Immunoprecipitation of PRCC from HeLa cell nuclear extract using anti-splicing factor antibodies Y12 (anti-Sm proteins), anti-SC-35, anti-PRL1 and anti-hCDC5

**Discussion**

In this study we have obtained several lines of evidence to support the view that PRCC is a component of the pre-mRNA splicing machinery. First, PRCC throughout the nucleus is co-localized with the Sm proteins, which are core components of snRNP's. This feature is specific to pre-mRNA splicing factor as evident from the fact that TFE3, which is a transcription factor, does not co-localize with Sm proteins. Secondly, a common feature shared between splicing factors and



**Figure 5** (a) Cellular localization of Sm proteins and GFP-fusions in Cos7 cells before and after  $\alpha$ -amanitin treatment. Endogenous Sm proteins were detected by the Y12 antibody (red, centre panel) while PRCC-TFE3 and NonO-TFE3 were detected as EGFP fusions (green, left panel). Combined images of Sm proteins and EGFP fusions are presented in the right hand panel. (b) Western blot analysis using polyclonal rabbit PRCC antiserum and pre-immune serum. 10 and 15  $\mu$ l of nuclear extract (NE) of papillary carcinoma line UOK120 shows the presence of PRCC protein in addition to extra band at similar level of expression corresponding to the fusion PRCC-TFE3. No proteins are detected with the pre-immune serum. (c) Immunoprecipitation of PRCC and PRCC-TFE3 from UOK120 line nuclear extract using anti-Sm antibodies. The 66 and 71 kDa bands, which correspond to the PRCC and PRCC-TFE3 proteins respectively, are indicated by arrows

snRNPs is to relocate into large speckles when RNA Polymerase II transcription or splicing are inhibited (Carneo-Fonseca *et al.*, 1992; O'Keefe *et al.*, 1994). We have made the use of the transcription inhibitor  $\alpha$ -amanitin to block transcription and demonstrate a precise co-localization of PRCC and NonO with Sm proteins in large speckles. In contrast, TFE3 cellular localization did not change upon inhibition of transcription and was completely excluded from the Sm proteins containing speckles. Lastly, co-immunoprecipitation experiments demonstrate that PRCC is associated with components of the spliceosomal complex such as Sm proteins, SC35, PRL1 and CDC5 in nuclear extract. A fully assembled spliceosome sediments at 50–60S, suggesting a complexity equivalent to that of the ribosomes (Brody and Ableson, 1985; Grabowski *et al.*, 1985). The spliceosome contains five snRNP subunits and each snRNP has one or more snRNAs and a group of common core proteins (Sm proteins) as well as specific proteins. The snRNP's are the main players in the formation of the catalytic core of the spliceosomes and serve as its highly conserved units (Kramer, 1996; Kreivi and Lamond, 1996). Co-localization and co-precipitation

of PRCC with Sm and other spliceosomal proteins suggest that PRCC is a component of this complex.

Weterman *et al.* (2000) found that PRCC-TFE3 is a stronger transactivator than TFE3, which can potentially lead to aberrant expression of TFE3 target genes. Our data, which monitor transcription from the PAI-1 promoter confirm that PRCC-TFE3 is a more potent transcriptional activator than TFE3, but by including NonO-TFE3 into the TFE3 specific reporter assay (Figure 1) we show that this fusion is a weaker transactivator than TFE3 and unable to enhance transcription through synergism with SMAD3. These data are backed by Northern blot analysis for PAI-1 transcription level (a TFE3 target gene) in NonO-TFE3 (UOK109) and PRCC-TFE3 (UOK120) containing cell lines demonstrating that the PAI-1 transcript is almost at an undetectable level in the former line and by the fact that Gal4 fusions of NonO could not activate transcription in a dual luciferase reporter assay. Moreover, NonO-TFE3 cellular distribution into very small speckles compared to the nuclear diffused pattern of wild type TFE3 and PRCC-TFE3 may suggest that only a small proportion of this fusion protein is available to interact with

TFE3 target genes. We conclude therefore that the high level of transcription activation from the endogenous PAI-1 promoter is not a common feature to all TFE3 fusions.

An alternative hypothesis suggested by our observation is that a role of fusion to TFE3 is to remove PRCC, NonO and PSF from the splicing complex. Thus colocalization and biochemical studies with Sm proteins show that the fusions either partially (for PRCC-TFE3) or completely (for NonO-TFE3) excluded from the spliceosome. Interestingly, immunodepletion of NonO or PSF from nuclear extract was shown to inhibit *in vitro* splicing reactions (Hallier *et al.*, 1998 and Paton *et al.*, 1993, respectively). In addition it has been proposed that both PSF and NonO are involved in the control of neuronal differentiation and maturation through splice site specific events occurring at particular developmental stages. These stages are highly correlated with differences in expression level of both NonO (Shinozaki *et al.*, 1999) and PSF (Chanas-Sacre *et al.*, 1999). We envisage that a similar situation might take place in kidney cells where PRCC, NonO and PSF may control the splicing of specific genes. Alteration of the levels ('dosage effect') and therefore the function of these splicing factors may result in inappropriate splicing of specific genes and cellular transformation.

Support for the 'dosage effect' hypothesis is provided by a study of t(6;19)(p21;q13.1) translocation in a patient with Multicystic Renal Dysplasia (MRD). Cloning of the breakpoints has revealed that USF2 on chromosome 19 is fused to CDC5L on chromosome 6 (Groenen *et al.*, 1996, 1998). Interestingly, USF2 is a member of the USF family of transcription factors which like TFE3, bind to an E-box sequence on DNA to activate gene transcription and they both belong to the highly conserved family of basic helix-loop-helix leucine-zipper transcription factors (bHLH-zip). The CDC5L gene product (hCDC5) which is a component of the 40S pre-mRNA splicing complex (Neubauer *et al.*, 1998) and is essential for pre-mRNA splicing (McDonald *et al.*, 1999) was shown in our study to coimmunoprecipitate with PRCC. Unlike PRCC-TFE3 however, which is expressed at a relatively high level, various attempts by Groenen *et al.* (1998) to isolate a transcript of the USF2-CDC5 fusion have been unsuccessful and led the authors to suggest that the translocation may cause MRD through a 'dosage effect' on CDC5 and not USF2. Because mice which are lacking the latter have a very subtle phenotype (Vallet *et al.*, 1997) and display no apparent kidney abnormalities (Groenen pers. comm with M Sawadogo., Houston, TX, 1997).

A further link between splicing and transformation has been obtained in functional studies, which were carried out on the Spi-1/PU.1 transcription factor in Friend erythroleukemia (Hallier *et al.*, 1996). In this disease the insertional mutagenesis of the *Spi-1* gene appears to result in overexpression of the normal Spi-1/PU.1 protein which appears to be related to the emergence of a clonal population of tumorigenic

erythroid cells. This led the authors to assume that its oncogenic potential may stem from 'targeting of inappropriate regulatory elements of some erythroid genes and/or abnormal association with erythroid partners'. Results of the search for such partner revealed that NonO (p54nrb) is associated both *in vivo* and *in vitro* with Spi-1/PU.1. Several experiments, which aimed to assess whether NonO could disturb a normal transcription by Spi-1/PU.1 were conclusively negative. Interestingly however, their discovery that Spi-1/PU.1 could also bind RNA sequences which are similar to those bound by NonO led them to demonstrate that the former inhibit RNA binding by NonO and consequently that addition of Spi-1/PU.1 to *in vitro* splicing reaction inhibited the formation of spliced transcript. This they suggested was likely to happen through disturbance of post-transcriptional gene regulation by Spi-1/PU.1 sequestering of RNA-binding proteins such as NonO.

Another example is provided by the recent discovery that the tumour suppressor protein WT1 can specifically interact with U2AF65, a 3' splice site binding protein suggest that WT1 has a function in pre-mRNA splicing (Davies *et al.*, 1998). It is assumed that 10% of sporadic wilms' tumours, which are kidney malignancies, are associated with WT1 mutations (Little *et al.*, 1999). This raises the possibility that aberrant splicing caused by mutated WT1 is also an important event in this class of renal cancer.

Our future studies will be directly aimed at understanding the exact role that the PRCC protein plays in pre-mRNA splicing. Knowledge of the role of PRCC in pre-mRNA splicing would provide new insight into the effect of the TFE3 fusion on PRCC function and tumorigenesis.

## Materials and methods

### Cell lines

The UOK120, and UOK109 cell lines were derived from primary papillary renal cell carcinoma specimens as described (Anglard *et al.*, 1992). The cell lines were derived respectively from tumours arising in a 30 and 39 years-old males. Molecular and cytogenetic analyses of the cell lines identified the PRCC-TFE3: t(X;1)(p11.2;q21.2) translocation and NonO-TFE3; inv(X)(p11.2;q12) inversion respectively. Another papillary renal cell carcinoma line UOK124 was derived from 21 years-old female and contains the PRCC-TFE3 translocation with 393 amino acids of PRCC N-terminal fused to TFE3 compared to the N-terminal 156aa of PRCC in UOK120. The synovial sarcoma cell line SS255, also known as BAX was obtained and described by Reeves *et al.* (1989). COS7 (Green monkey kidney cell line) and HepG2 (Hepatoblastoma) were obtained from ATCC. All cells were cultured in DMEM with 10% foetal calf serum and incubated in CO<sub>2</sub> (5% v/v) unless otherwise is stated.

### Plasmids description and construction

For transactivation and cellular localization studies full-length cDNAs of PRCC-TFE3, PRCC, NonO, and

NonO-TFE3 were cloned in frame with green fluorescent protein (GFP, Clontech, USA) in pEGFP-C1 vector. NonO-TFE3 and PRCC-TFE3 were cloned in two steps. First, a TFE3 insert was generated by PCR of its C-terminal region using the forward primer 5'-CAATGATGAAATGCTCAGC-3' and the reverse primer 5'-ATATGATCCTCAGGACTCCTCTTCCAT-3' spanning base numbers 1059-1962 (accession number X96717). This insert was digested with *Bam*HI and *Bgl*II, and subcloned into pEGFP-N1 to form an intermediate GFP construct. NonO-TFE3 insert was generated by PCR with the NonO forward primer 5'-ATATAGATCTGCAAAAATGCAGAGTAAT-3' spanning NonO base numbers 141-158 (accession number NM 007363), and the TFE3 reverse primer 5'-CTTCTCTTGCCGTTCCCTTC-3', spanning TFE3 base numbers 1260-1279 (accession number X96717), and a PRCC-TFE3 was generated by PCR using the PRCC forward primer 5'-AT-ATAGATCTTCGCTGGTTGCTTACGCC-3' spanning PRCC base numbers 244-261 (accession no X97124), and the TFE3 reverse primer 5'-CTTTCTTCCGTTCCCTTC-3', spanning TFE3 base numbers 1260-1279 (accession number X96717). These inserts were then digested with *Bgl*II and subcloned into the intermediate pEGFP-C1 construct. Full length PRCC was generated by PCR using the forward primer 5'-ATATAGATCTTCGCTGGTTGCTTACGCC-3' and the reverse primer 5'-ATATAAGCTTCC TAGAATCCATATTTGGC-3' (accession number X97124). Full length NonO was generated by PCR using the forward primer 5'-ATATAGATCTGC AAAAATGCAGAGTAAT-3' and the reverse primer 5'-ATATAAGCTTAT TAGTATCGGCGACG-3' (NM 007363). Both inserts were digested with *Bam*HI and *Bgl*II and subcloned into pEGFP-C1 vector. Full-length cDNA of TFE3 (accession number X96717) was cloned into *Hind*III and *Xba*I restriction sites of pEGFP-C2 vector. All constructs were fully sequenced to confirm correct reading frame and more than one DNA prep of a specific DNA was used in most experiments.

In order to study their transactivation capacity in TFE3 specific reporter assay, of a reporter construct containing a fragment of the PAI-1 promoter with either wild type (PE2) or mutant (PmE2) TFE3 specific binding site (E-box) was cloned upstream a luciferase gene. Both were described (Hua *et al.*, 1998) and kindly provided by Dr Xianxin H (Department of Biology, MIT, USA). SMAD3 plasmid was kindly provided by Dr Caroline Hill (ICRF, UK).

For transactivation studies of NonO, the corresponding full length or the translocated cDNA's regions, were cloned into *Sal*I-*Spe*I restriction sites in frame with Gal4 in PMG147spe vector which contains the MLV (Moloney leukemia virus) promoter upstream the 147 aa of the yeast Gal4 DNA binding domain. cDNA's were generated by PCR using the following primers: Translocated region (accession NM 007363): 5'-ATATGTCGACGCAAAAATGCAGAGTAAT-3' (Forward primer), 5'-ATATACTAGTC-GCAT-CAGGGAAGGTTCC-3' (Reverse primer) and Full length NonO (accession NM 007363) 5'-ATATGTCGACGCAAAAATGCAGAGTAAT-3' (Forward primer), 5'-ATATAC-TAGTAT TAGTATCGGCGACG-3' (Reverse primer).

#### Cellular localization

Cellular localization using enhanced green fluorescent protein (EGFP) and immunolocalization with Y12 antibody were carried out using procedures described by Thaete *et al.* (1999).

#### Transfection and luciferase assays

Transfection into HepG2 cell line, TGF $\beta$  induction, and luciferase reporter assays were carried out as described by Hua *et al.* (1998) except of the following changes: Transfections were carried out into 6-well plates and the amount of transfected DNA was therefore doubled: 0.5  $\mu$ g for Renella luciferase, 2  $\mu$ g for SMAD3, 1  $\mu$ g for either GFP-PRCC, GFP-TFE3 or GFP-NonO, 1  $\mu$ g for GFP-PRCC-TFE3, or GFP NonO-TFE3, 1  $\mu$ g of GFP vector and 2.5  $\mu$ g for PE2 or PmE2 luciferase reporters. Results were normalized for transfection efficiency by the use of co-expressed renella luciferase and for level of protein expression by Western blot analysis using anti-pEGFP against the GFP fusions and anti-myc against myc-tagged SMAD3 antibodies on cell lysate, which was used for measuring the reporter activity. Level of EGFP-fusion expression was also assayed, although to a lesser extent by the confocal microscope.

Transfections of Gal4 fusions into HeLa and HepG2 cells were carried out as described above using the dual luciferase reporter assay (Promega) except for the following changes. No starvation followed by induction took place and amount of transfected DNA was titrated. HeLa cells were transfected by the lipofectamine transfection kit (Life technologies, UK) rather by the Calcium phosphate precipitation method. The observed transcriptional activation by TFE3 was achieved using 0.5  $\mu$ g of Gal4 TFE3 and 2  $\mu$ g of reporter plasmid.

#### Northern blot analysis

A 570 bp of PAI-1 gene probe was prepared by PCR amplification of cDNA which was reverse-transcribed using PAI-1-specific primer: 5'-GGATATGATAA ATATTTAGGT-3'. PCR was carried out with PAI-1 forward 5'-AACG-TGGTTTTCTCACCTA-3' and reverse primer 5'-GTTGG-TCTGAGCCATC ATG-3'. Labelling of probe, preparation of total RNA and Northern blot analysis were carried as described previously by Hodgkinson *et al.* (1993) and according to standard methods (Sambrook *et al.*, 1989).

#### Preparation of PRCC antibodies

Anti serum against the N-terminal portion of PRCC was raised in rabbits immunized with GST-PRCC fusion protein. A PRCC PCR product was generated using the PRCC forward primer 5'-ATATCCCGGGCTCGCTGGTTGCTTACGCCAGCA-3' and the PRCC reverse primer 5'-ATATCCCGGGCTGGATCCCTGAAGGATAG-3'. This PCR product which comprises the first 521 base pairs of PRCC open reading frame was cloned in fusion with GST gene of the pGEX-4T-2 vector (Pharmacia). The GST-PRCC fusion was expressed and isolated from BL21 (DE3) pLyS *E. coli* competent cells (Stratagene) according to the manufacturer's instructions (Pharmacia).

#### Preparation of UOK120 nuclear extract

Nuclear extract preparation was carried out as described by Dignam *et al.* (1983), in collaboration with TCS cell work (UK).

#### Immunoprecipitation and Western blot analysis

Immunoprecipitation and Western blot analysis were carried out according to standard protocols (Sambrook *et al.*, 1989). Antibodies used in immunoprecipitation and Western blotting included mouse monoclonal antibodies to Sm proteins

(anti-Y12 dilution 1:500; Pettersson *et al.*, 1984), rabbit polyclonal antibody to PRCC protein (described above, dilution 1:1000), mouse monoclonal antibody to the SR protein SC-35 (Fu and Maniatis, 1990), Sheep polyclonal affinity purified antibody against hCDC5 which is the human homologue of the *S. pombe* CDC5<sup>+</sup> gene product (Bernstein and Coughlin, 1997; Neubauer *et al.*, 1998) and sheep polyclonal antibody against hPRL1 protein which is the human homologue of the Arabidopsis thaliana PRL1 gene product (Neubauer *et al.*, 1998; McDonald *et al.*, 1999). Ten  $\mu$ l of affinity purified and 20  $\mu$ l of crude sera antibodies was mixed with protein-G sepharose prior to incubation with HeLa or UOK120 cell nuclear extracts.

## References

- Anglard P, Trahan E, Liu S, Latif F, Merino MJ, Lerman MI, Zbar B and Linehan WM. (1992). *Cancer Res.*, **52**, 348–356.
- Artandi SE, Merrell K, Avitahl N, Wong KK and Calame K. (1995). *Nucleic Acids Res.*, **23**, 3865–3871.
- Bajou K. (1998). *Nat. Med.*, **4**, 923–928.
- Basu A, Dong B, Krainer AR and Howe CC. (1997). *Mol. Cell Biol.*, **17**, 677–686.
- Beckmann H, Su LK and Kadesch T. (1990). *Genes Dev.*, **4**, 167–179.
- Bernstein HS and Coughlin SR. (1997). *J. Biol. Chem.*, **272**, 5833–5837.
- Brody E and Ableson J. (1985). *Science*, **24**, 963–967.
- Carneo-Fonseca M, Pepperkok R, Carvalho MT and Lamond AL. (1992). *J. Cell Biol.*, **117**, 1–14.
- Chanas-Sacre G, Mazy-Servais C, Wattiez R, Pirard S, Rogister B, Patton JG, Belachew S, Malgrange B, Moonen G and Leprince P. (1999). *J. Neurosci. Res.*, **1**, 62–73.
- Clark J, Lu YJ, Sidhar SK, Parker C, Gill S, Smedley D, Hamoudi R, Linehan WM, Shipley J and Cooper CS. (1997). *Oncogene*, **15**, 2233–2239.
- Davies RC, Clavio C, Bratt E, Larsson SH, Lamond AI and Hastie ND. (1998). *Genes Dev.*, **12**, 3217–3225.
- De Jong B, Molenaar IM, Leeuw JA, Idenberg VJ and Oosterhuis JW. (1986). *Cancer Genet. Cytogenet.*, **21**, 165–169.
- Dignam JD, Lebovitz RM and Roeder RG. (1983). *Nucleic Acids Res.*, **11**, 1475–1489.
- Dong B, Horowitz DS, Kobayashi R and Krainer AR. (1993). *Nucleic Acids Res.*, **21**, 4085–4092.
- Fu X-D. (1995). *RNA*, **1**, 663–680.
- Fu, X-D and Maniatis T. (1990). *Nature*, **343**, 437–441.
- Grabowski PJ, Seiler SR and Sharp PA. (1985). *Cell*, **42**, 345–353.
- Groenen MA, Garcia E, Debeer P, Devriendt K, Fryns JP and Ven de Van. (1996). *Genomics*, **38**, 141–148.
- Groenen MA, Vanderlinden G, Devriendt K, Fryns JP and Ven de Van. (1998). *Genomics*, **49**, 218–229.
- Hallier M, Lerga A, Barnache S, Tavitian A and Moreau GF. (1998). *J. Biol. Chem.*, **273**, 4838–4842.
- Hallier M, Tavitian A and Moreau GF. (1996). *J. Biol. Chem.*, **271**, 11177–11181.
- Hemesath TJ, Steingrimsson E, McGill G, Hansen MJ, Vaught J, Hodgkinson CA, Arnheiter H, Copeland NG, Jenkins NA and Fisher DE. (1994). *Genes Dev.*, **8**, 2770–2780.
- Hodgkinson CA, Moore KJ, Nakayama A, Steingrimsson E, Copeland NG, Jenkins NA and Arnheiter H. (1993). *Cell*, **74**, 395–404.
- Hua X, Liu X, Ansari D and Lodish H. (1998). *Genes Dev.*, **12**, 3084–3095.
- Hua X, Miller ZA, Wu G, Shi Y and Lodish HF. (1999). *Proc. Natl. Acad. Sci. USA*, **96**, 13130–13135.
- Kramer A. (1996). *Annu. Rev. Biochem.*, **65**, 367–409.
- Kreivi J-P and Lamond AI. (1996). *Curr. Biol.*, **6**, 802–804.
- Little M, Holmes G and Walsh P. (1999). *Bioessays*, **21**, 191–202.
- McDonald WH, Ohi R, Smelkova N, Frenthewey D and Gould KL. (1999). *Mol. Cell Biol.*, **19**, 5352–5362.
- Meloni AM, Dobbs RM, Pontes JE and Sandberg AA. (1993). *Cancer Genet. Cytogenet.*, **65**, 1–6.
- Neubauer G, King A, Rappsilber J, Calvio C, Watson M, Ajuh P, Sleeman J, Lamond A and Mann M. (1998). *Nat. Genet.*, **20**, 46–50.
- O'Keefe R, Mayeda A, Sadowski C, Krainer A and Spector D. (1994). *J. Cell Biol.*, **124**, 249–260.
- Patton JG, Porro EB, Galceran J, Tempst P and Nadal GB. (1993). *Genes Dev.*, **7**, 393–406.
- Pettersson I, Hinterberger M, Mimori T, Gottlieb E and Steitz JA. (1984). *J. Biol. Chem.*, **259**, 5907–5914.
- Reeves BR, Smith S, Fisher C, Warren W, Knight J, Martin C, Chan AM, Guesterson BA, Westbury G and Cooper CS. (1989). *Oncogene*, **4**, 373–378.
- Roman C, Matera AG, Cooper C, Artandi S, Blain S, Ward DC and Calame K. (1992). *Mol. Cell Biol.*, **12**, 817–827.
- Sambrook J, Fritsch EF and Maniatis T. (1989). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Shinozaki A, Arahata K and Tsukahara T. (1999). *Int. J. Biochem. Cell Biol.*, **31**, 1279–1287.
- Sidhar SK, Clark J, Gill S, Hamoudi R, Crew AJ, Gwilliam R, Ross M, Linehan WM, Birdsall S, Shipley J and Cooper CS. (1996). *Hum. Mol. Genet.*, **5**, 1333–1338.
- Thaete C, Brett D, Monaghan P, Whitehouse S, Rennie G, Rayner E, Cooper CS and Goodwin G. (1999). *Hum. Mol. Genet.*, **8**, 585–591.
- Tian G, Erman B, Ishii H, Gangopadhyay SS and Sen R. (1999). *Mol. Cell Biol.*, **1**, 2946–2957.
- Vallet VS, Henrion AA, Bucchini D, Casado M, Raymond-jean M, Kahn A and Vaulont S. (1997). *J. Biol. Chem.*, **272**, 21944–21949.
- Weterman MA, Wilbrink M and van Kessels A. (1996). *Proc. Natl. Acad. Sci. USA*, **93**, 15294–15298.
- Weterman MA, van Groningen JJ, Jansen A and van Kessel AG. (2000). *Oncogene*, **19**, 69–74.
- Yang YS, Hanke JH, Carayannopoulos L, Craft CM, Capra JD and Tucker PW. (1993). *Mol. Cell Biol.*, **13**, 5593–5603.