2011

Serine/arginine-rich Protein 30c Activates Human Papillomavirus Type 16 L1 mRNA Expression Via a Bimodal Mechanism

Monika Somberg
Uppsala University, Sweden

Xiaoze Li
Uppsala University, Sweden

Cecilia Johansson
Uppsala University, Sweden

See next page for additional authors

Recommended Citation

Somberg, M., Li, X., & Johansson, C. (2011). Serine/arginine-rich protein 30c activates human papillomavirus type 16 L1 mRNA expression via a bimodal mechanism Free. *Journal of General Virology*, vol. 92, no. 10. doi:10.1099/vir.0.033183-0

Journal of General Virology (2011), **92**, 2411–2421 DOI 10.1099/vir.0.033183-0

Serine/arginine-rich protein 30c activates human papillomavirus type 16 L1 mRNA expression via a bimodal mechanism

Monika Somberg,[1] Xiaoze Li,[1,2] Cecilia Johansson,[1] Beatrice Orru,[3] Roger Chang,[1] Margaret Rush,[1] Joanna Fay,[3]† Fergus Ryan[3] and Stefan Schwartz[1,2,3]

Correspondence
Stefan Schwartz
Stefan.Schwartz@med.lu.se

[1]Department of Medical Biochemistry and Microbiology, Uppsala University, Biomedical Center, Husargatan 3, Box 582, 751 23 Uppsala, Sweden

[2]Department of Laboratory Medicine, Section of Medical Microbiology, Lund University, 221 84 Lund, Sweden

[3]Dublin Institute of Technology, Kevin Street, Dublin 8, Ireland

Two splice sites on the human papillomavirus type 16 (HPV-16) genome are used exclusively by the late capsid protein L1 mRNAs: SD3632 and SA5639. These splice sites are suppressed in mitotic cells. This study showed that serine/arginine-rich protein 30c (SRp30c), also named SFRS9, activated both SD3632 and SA5639 and induced production of L1 mRNA. Activation of HPV-16 L1 mRNA splicing by SRp30c required an intact arginine/serine-repeat (RS) domain of SRp30c. In addition to this effect, SRp30c could enhance L1 mRNA production indirectly by inhibiting the early 3′-splice site SA3358, which competed with the late 3′-splice site SA5639. SRp30c bound directly to sequences downstream of SA3358, suggesting that SRp30c inhibited the enhancer at SA3358 and caused a redirection of splicing to the late 3′-splice site SA5639. This inhibitory effect of SRp30c was independent of its RS domain. These results suggest that SRp30c can activate HPV-16 L1 mRNA expression via a bimodal mechanism: directly by stimulating splicing to late splice sites and indirectly by inhibiting competing early splice sites.

Received 17 April 2011
Accepted 17 June 2011

INTRODUCTION

Human papillomavirus type 16 (HPV-16) is the most common high-risk HPV type in the human population (Howley & Lowy, 2001; zur Hausen, 2002). Although the vast majority of HPV-16 infections are cleared by the immune system of the host, chronic infections with HPV-16 cause more than 400 000 cases of cervical cancer globally each year (Longworth & Laimins, 2004; zur Hausen, 2002). HPV-16 persistence probably requires a successful delay in the production of the immunogenic capsid proteins L1 and L2 at early stages in the virus life cycle to avoid host immune responses. The L1 and L2 capsid proteins of HPV-16 are normally expressed in terminally differentiated cells in the upper layers of the mucosal epithelium (Doorbar, 2005; Longworth & Laimins, 2004). In chronic HPV-16 infections that deteriorate to anogenital cancer, L1 and L2 genes are efficiently shut down. Cervical cancer cells do not express L1 or L2 protein, suggesting that inhibition of late gene expression is a prerequisite for progression to cervical cancer.

HPV late gene expression is regulated at least in part at the level of RNA processing, and several reports have described regulatory elements on late HPV mRNAs (Graham, 2008; Schwartz, 2008; Zheng & Baker, 2006). These HPV RNA elements control RNA processing such as viral mRNA stability, splicing and translation. The late 3′-untranslated region (3′UTR) has been identified as the location of RNA regulatory elements in many HPV types, including HPV-16 (Graham, 2008; Zhao *et al.*, 2007). RNA regulatory elements have also been identified in the L1 and L2 protein-encoding regions of HPV-16 (Collier *et al.*, 1998, 2002; Oberg *et al.*, 2003). Reports also suggest that the control of late gene expression extends to regulatory elements in the early region of HPV-16, and polyadenylation elements at the early poly(A) signal pAE may influence late gene expression (Zhao *et al.*, 2005). The identification of host cellular factors binding to these regulatory sequences would help to resolve the mechanism of HPV-16 late gene suppression in undifferentiated cells, including cancer cells. The following factors have been shown to bind to the HPV-16 and/or HPV-31 late 3′UTR: HuR, U2 small

†Present address: National Children's Research Centre, Our Lady's Children's Hospital, Crumlin, Dublin 12, Ireland.

A supplementary table is available with the online version of this paper.

nuclear ribonucleoprotein auxiliary factor 65 kDa subunit ($U2AF^{65}$), cleavage stimulation factor-64 (CstF-64), alternative splicing factor/splicing factor 2 (ASF/SF2) and CUG-binding protein (Cumming *et al.*, 2002; Goraczniak & Gunderson, 2008; Koffa *et al.*, 2000; McPhillips *et al.*, 2004), whilst heterogeneous nuclear ribonucleoprotein (hnRNP) C1/C2, HuR and poly(A)-binding protein (PABP) bind to the late 3′UTR of HPV-1 (Sokolowski & Schwartz, 2001; Sokolowski *et al.*, 1997, 1999). hnRNP C1/C2, polypyrimidine tract binding-protein (PTB), CstF-64 and human Fip-1 have been shown to bind to a U-rich region of the early 3′UTR of HPV-16, a sequence that enhances recognition of the early poly(A) signal pAE (Zhao *et al.*, 2005). The hnRNP A1 protein was shown to bind to splicing silencers in the HPV-16 L1-encoding region (Zhao *et al.*, 2004), whilst hnRNP E, hnRNP K and hnRNP H bind to the coding region of HPV-16 L2 (Collier *et al.*, 1998; Oberg *et al.*, 2005). The binding sites for hnRNP H coincide with downstream RNA elements that stimulate HPV-16 early polyadenylation (Oberg *et al.*, 2005). In HPV-31, the corresponding RNA sequence interacts with CstF-64 (Terhune *et al.*, 2001).

Proteins that can induce HPV-16 late gene expression from subgenomic HPV-16 expression plasmids have been identified recently (Somberg & Schwartz, 2010; Somberg *et al.*, 2008, 2009). PTB was found to induce production of the partially spliced L2 mRNA, suggesting that HPV-16 polyadenylation and/or splicing was targeted by PTB (Somberg *et al.*, 2008). In addition to PTB, the multifunctional adenovirus E4orf4 protein induced production of HPV-16 L1 mRNA (Somberg *et al.*, 2009). As E4orf4 interacts with cellular serine/arginine-rich (SR) proteins (in particular ASF/SF2) (Kanopka *et al.*, 1996, 1998), these results indicated that SR proteins regulate HPV-16 L1 mRNA splicing. We have shown previously that ASF/SF2 regulates HPV-16 mRNA splicing (Somberg & Schwartz, 2010), and Jia *et al.* (2009) showed that bovine papillomavirus type 1 and HPV-16 splicing is regulated by SRp20. In the case of HPV-16, one may therefore speculate that splicing of the late L1 mRNA is absent in mitotic cells or cervical cancer cells due to suboptimal levels of one or more of the SR proteins (Schwartz, 2008).

RESULTS

SRp30c activates HPV-16 late mRNA splicing and induces production of spliced L1 mRNA

In order to identify cellular factors that regulate HPV-16 late mRNA splicing, we co-transfected subgenomic HPV-16 plasmid pBEL (Fig. 1a) with cDNAs expressing the SR proteins SRp20, SRp30c, SC35, ASF/SF2, SRp40, SRp55 or SRp75. The expression of SR proteins or PTB from the plasmids used here has been demonstrated previously by our group (PTB: Somberg *et al.*, 2008; ASF/SF2: Somberg & Schwartz, 2010; SRp40, SRp55 and SRp75: Tranell *et al.*, 2010). Production of SRp20 and SC35 is shown in Fig. 1(b, lower panels). pBEL encoded viral early and late genes after a human cytomegalovirus (CMV) immediate-early promoter (Fig. 1a), but primarily produced the early E4 mRNA upon transfection of HeLa cells, as described previously (Rush *et al.*, 2005; Somberg & Schwartz, 2010; Somberg *et al.*, 2008, 2009; Zhao *et al.*, 2004, 2005). HPV-16 late L1 and L2 mRNAs were undetectable (Fig. 1b). This is due to the efficient use of the early splice site SA3358 and early poly(A) signal pAE (Rush *et al.*, 2005; Somberg & Schwartz, 2010; Zhao *et al.*, 2005) and to suppression of the late splice sites SD3632 and SA5639 (Fig. 1a) (Somberg & Schwartz, 2010; Zhao *et al.*, 2005). pBEL was co-transfected with 1 μg each of plasmids expressing the various SR proteins indicated in Fig. 1(b). Cytoplasmic RNA was extracted and subjected to Northern blotting with an L1 probe (Fig. 1a). The results revealed that overexpression of SRp30c induced late gene expression, primarily L1/L1i mRNA (where L1i mRNA represents a 'mutation-induced' late message spliced from the position 880 5′-splice site directly to the position 5639 3′-splice site), but not L2 mRNA (Fig. 1b). None of the other SR proteins induced HPV-16 late gene expression (Fig. 1b). In addition, SR protein 9G8, a splicing regulatory protein, failed to induce HPV-16 late gene expression under these conditions (data not shown). The induction of spliced L1 mRNA by SRp30c is in contrast to the induction of partially spliced L2 mRNA by PTB (hnRNP I) (Fig. 1b), as reported previously (Somberg *et al.*, 2008). Transfection of pBEL with serially diluted SRp30c plasmid revealed that SRp30c induced L1 mRNA splicing in a dose-dependent manner (Fig. 1c). A Western blot showed the levels of overexpression of SRp30c in HeLa cells (Fig. 1d). We also transfected pBEL with SRp30c in other cell lines such as 293T cells and C33A with similar results (data not shown). Our results indicated that overexpression of SRp30c activates splicing from either SD3632 or SA5639, or both. We concluded that SRp30c is an activator of HPV-16 late L1 mRNA splicing.

SRp30c inhibits splicing to early splice site SA3358, thereby redirecting splicing to late splice site SA5639

Two different L1 mRNAs are detected by the L1 probe, the L1 and L1i mRNAs (Fig. 1a). These two mRNAs are not easily separated by Northern blotting. To determine whether SRp30c induced L1 or L1i mRNA, we analysed RNA extracted from HeLa cells transfected with pBEL and a serial dilution of SRp30c plasmid by RT-PCR. The results revealed that low levels of SRp30c overexpression primarily induced splicing of late mRNAs of the L1 type, whereas higher levels of SRp30c overexpression induced late mRNAs of the L1i type (Fig. 2a, b). These results confirmed that SRp30c could overcome suppression of late splice site SD3632 and/or SA5639. This production of L1i mRNAs at high concentrations of transfected SRp30c plasmid suggested that SRp30c induced L1 mRNA production by another mechanism. The results suggested that SRp30c also caused skipping of the exon between SA3358 and SD3632

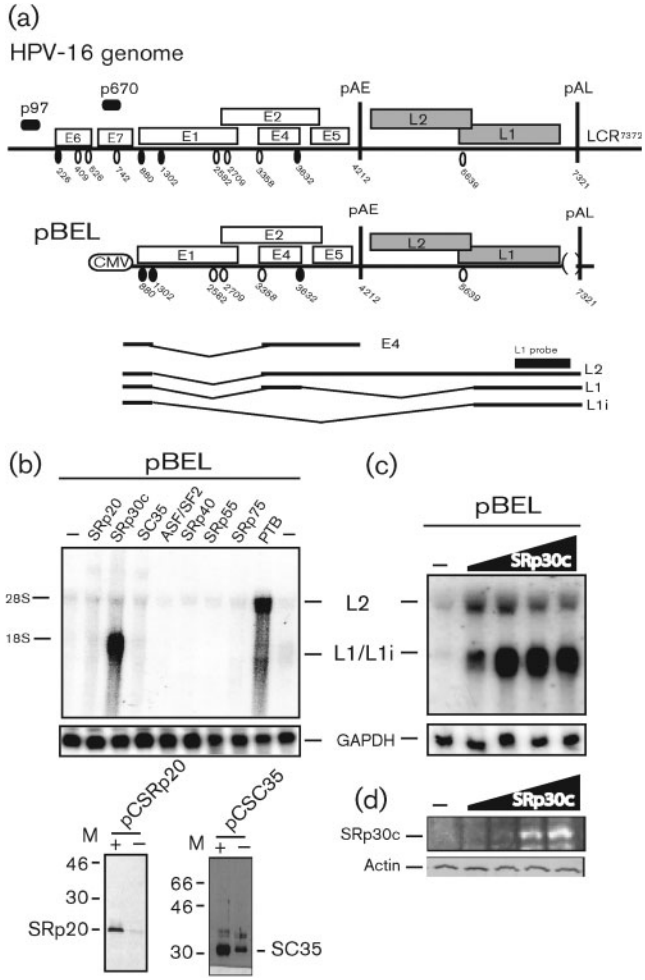

Fig. 1. (a) Schematic representation of the HPV-16 genome and subgenomic HPV-16 expression plasmid pBEL (Zhao *et al.*, 2004). The early and late viral promoters p97 and p670 are indicated. Numbers indicate nucleotide positions of 5′-splice sites (filled ovals), 3′-splice sites (open ovals) or the early and late poly(A) signals pAE and pAL, respectively. A subset of mRNAs that may be produced by pBEL are indicated (Baker & Calef, 1997). The position of the L1 probe is shown. CMV, Human cytomegalovirus promoter; LCR, long control region. (b) Upper panel: Northern blot of cytoplasmic RNA extracted from HeLa cells transfected with pBEL in the absence or presence of plasmids expressing SRp20, SRp30c, SC35, ASF/SF2, SRp40, SRp55, SRp75 or PTB and probed with the L1 probe. The positions of the L2 and L1/L1i mRNAs, and 28S and 18S rRNAs are indicated. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was detected as a loading control. Lower panel: Western blots demonstrating the production of SC35 and SRp20 from expression plasmids. (c) Northern blot of cytoplasmic RNA extracted from HeLa cells transfected with pBEL and a 2-fold serial dilution of SRp30c expression plasmid, starting at 4 μg, probed with the L1 probe. GAPDH was detected as a loading control. (d) Western blot demonstrating overexpression of SRp30c in HeLa cells, using anti-SRp30c antibody N-13 or anti-actin antibody I-19 as a loading control (Santa Cruz Biotechnology).

in the E4-encoding region of HPV-16 (Fig. 2a, b). To confirm these results, the same RNA samples were subjected to Northern blotting and analysed with a probe for E4 mRNAs (Fig. 2c). Low levels of SRp30c marginally affected early E4 mRNAs, and induced expression of L1 mRNA (Fig. 2c). These L1 mRNAs were of the L1 type containing the central exon between SA3358 and SD3632, as they were detected by the E4 probe. In contrast, higher levels of SRp30c reduced levels of both E4 and L1 mRNAs (Fig. 2c), whilst at the same time shifting expression to L1i mRNAs. Hybridization of the same filter to an L1 probe, which detects both L1 and L1i mRNAs, revealed that high levels of SRp30c inhibited E4 and L1 mRNA production but induced high levels of L1i mRNAs, which were detected by the L1 probe (Fig. 2d) but not by the E4 probe (Fig. 2c). Taken together, these results indicated that SRp30c induced HPV-16 late gene expression by activation of late splice sites and suppression of early splice sites, i.e. SA3358. To confirm that this also occurred in the absence of the late genes, serially diluted SRp30c was co-transfected with pBearly (Fig. 2a), which lacks the late region of pBEL, but contains the early splice sites SD880 and SA3358 that are also present in pBEL. Northern blotting revealed that low levels of SRp30c had little or no effect on E4 mRNAs (Fig. 2c), whereas high levels of SRp30c inhibited production of early E4 mRNAs that were spliced to SA3358 (Fig. 2c). These results indicated that SRp30c inhibited SA3358. In addition, we transfected serially diluted SRp30c (0.125–1.0 μg) with plasmid pBELM (Fig. 2a), in which the splicing silencers at SA5639 had been inactivated. SRp30c induced L1 mRNA production from this plasmid in a manner similar to the induction of L1 from pBEL (low levels of SRp30c induced primarily L1 mRNAs, whereas high levels of SRp30c induced L1i mRNAs) (Fig. 2d and data not shown). Taken together, these results demonstrated that SRp30c induces HPV-16 late gene expression by activating SA5639 and/or SD3632 to produce spliced L1 mRNAs, and by inhibiting early splice site SA3358 to produce L1i mRNA by an exon-skipping mechanism.

To demonstrate further that the exon-skipping mechanism was specific for SRp30c, we compared the effects of SRp30c and adenovirus E4orf4 (AdE4orf4) on L1 mRNA production. In contrast to SRp30c, AdE4orf4 induced primarily L1 mRNA and not L1i mRNA (Fig. 2e), as reported previously (Somberg *et al.*, 2009). This was confirmed by RT-PCR (Fig. 2b). These results indicated that AdE4orf4 solely induced late mRNA splicing from SD3632 to SA5639, which confirmed our previously published results on AdE4orf4 (Somberg *et al.*, 2009). Although the same splicing event could be activated by SRp30c, SRp30c also inhibited splicing to SA3358 at high concentrations, thereby indirectly inducing splicing directly from SD880 to SA5639. Quantitative PCR using SYBR Green on the same cDNA samples revealed that overexpression of SRp30c (0.125 μg) caused a 4.1-fold reduction in cycle threshold value. Quantification of the L1 and L1i peaks demonstrated that the L1i peak was 49 % of the L1 + L1i

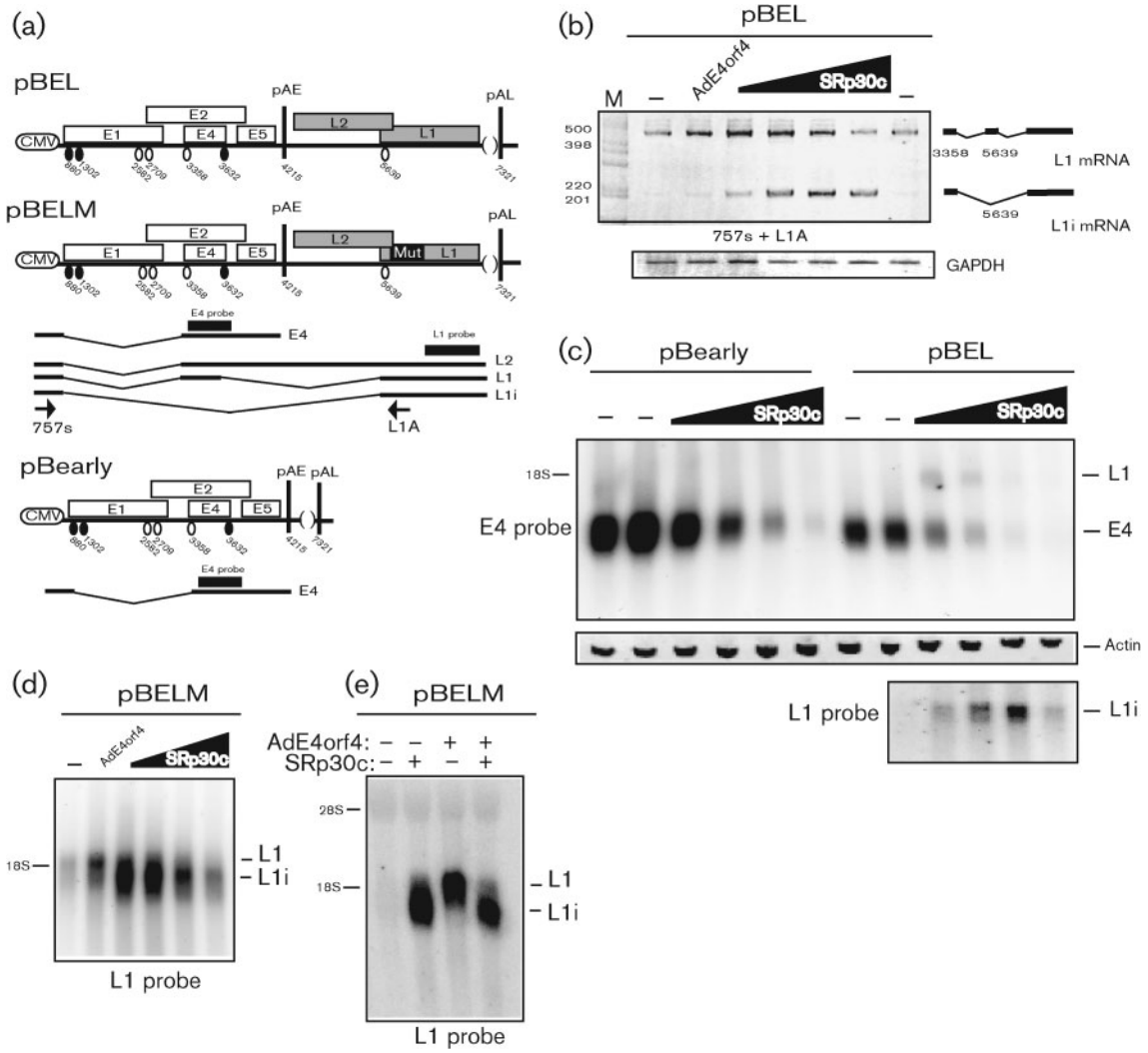

Fig. 2. (a) Schematic representation of the subgenomic HPV-16 expression plasmids pBEL, pBELM and pBearly (Zhao *et al.*, 2004, 2005). Mut indicates a number of nucleotide substitutions that inactivated splicing silencers downstream of SA5639 in plasmid pBELM (Zhao *et al.*, 2004). Numbers indicate nucleotide positions of 5′-splice sites (filled ovals), 3′-splice sites (open ovals) or the early and late poly(A) signals pAE and pAL, respectively. A subset of mRNAs that may be produced by the plasmids is indicated. The positions of the E4 and L1 probes and RT-PCR primers 757s and L1A are shown. (b) RT-PCR with primers 757s and L1A on cDNA of cytoplasmic RNA extracted from HeLa cells transfected with the indicated plasmids. pBEL was transfected with a twofold serial dilution of SRp30c expression plasmid, starting at 4 μg. GAPDH was amplified from the same cDNAs. PCR products representing mRNAs are indicated. M, Molecular size markers (bp). (c–e) Northern blots of cytoplasmic RNA extracted from HeLa cells transfected with pBEL, pBELM or pBearly in the absence or presence of plasmids expressing SRp30c or E4orf4. SRp30c expression plasmid was twofold serially diluted prior to transfection, starting at 4 μg. HPV-16 mRNAs detected by the E4 or L1 probe are indicated, and the positions of the 28S and 18S rRNAs are indicated. Actin was detected as a loading control.

mRNAs at low levels of SRp30c (0.125 μg), whilst the L1i peak shifted to 71 % of the L1 + L1i mRNAs in the presence of high levels of SRp30c (1 μg). Co-transfection of pBELM with both SRp30c and AdE4orf4 resulted in high levels of L1i (Fig. 2e), which was expected, as SRp30c would also inhibit SA3358 in the presence of E4orf4. We concluded that SRp30c induces production of both L1 and L1i mRNA.

These results suggested that SRp30c binds directly to the HPV-16 sequence between SA3358 and SD3632. To test whether SRp30c interacts physically with this sequence, we performed UV cross-linking of radiolabelled HPV-16 RNA (nt 3407–3627) to glutathione *S*-transferase (GST)-fused SRp30c (Fig. 3a). A Coomassie blue-stained acrylamide gel demonstrated the purified GST–SRp30c (Fig. 3b). Whilst the recombinant GST–SRp30c was cross-linked efficiently

to the HPV-16 RNA sequence, GST alone was not (Fig. 3c). The interaction between the HPV-16 RNA and SRp30c was sequence-specific, as a 4-fold excess of unlabelled HPV-16 RNA competed well with the probe for SRp30c, whilst little or no competition was observed using a 20-fold excess of unlabelled HPV-1 M5 RNA (Fig. 3d). We concluded the SRp30c binds directly to the HPV-16 sequence between SA3358 and SD3632. We cannot exclude the possibility that SRp30c binds to other positions in the HPV-16 genome.

SRp30c induces exon skipping by interfering with splicing enhancers downstream of SA3358 in an RS domain-independent manner

Our results demonstrated that wild-type (WT) SRp30c displayed properties similar to those of a *trans*-dominant ASF/SF2 protein (Somberg & Schwartz, 2010), i.e. it competed with WT ASF/SF2 for binding to the splicing enhancer downstream of SA3358, thereby inhibiting SA3358 and redirecting splicing to SA5639. This *trans*-dominant ASF/SF2 mutant, named ASFDRS, has been shown previously to inhibit splicing to SA3358 (Somberg & Schwartz, 2010). If induction of L1i mRNA by SRp30c was the result of competition with ASF/SF2 for binding to the splicing enhancer downstream of SA3358, an SRp30c mutant lacking the RS domain should also be able to induce L1i mRNA production. This prompted us to investigate the effect on HPV-16 late gene expression of an SRp30c mutant that lacked the RS domain (SRp30cDRS). Transfection of SRp30cDRS with pBELM followed by analysis of L1 mRNA production revealed that both WT SRp30c and SRp30cDRS induced production of L1/L1i mRNA to similar levels (Fig. 4a). Interestingly, SRp30c and SRp30cDRS induced similar levels of L1i mRNA from pBELM, demonstrating that the major L1i-inducing mechanism of both SRp30c and SRp30cDRS was inhibition of SA3358. RT-PCR analysis confirmed that both proteins induced HPV-16 late mRNAs of the L1i type, indicative of skipping of the internal exon SA3358–SD3632 in the E4-encoding region (Fig. 4b). Similar results were obtained with ASFDRS

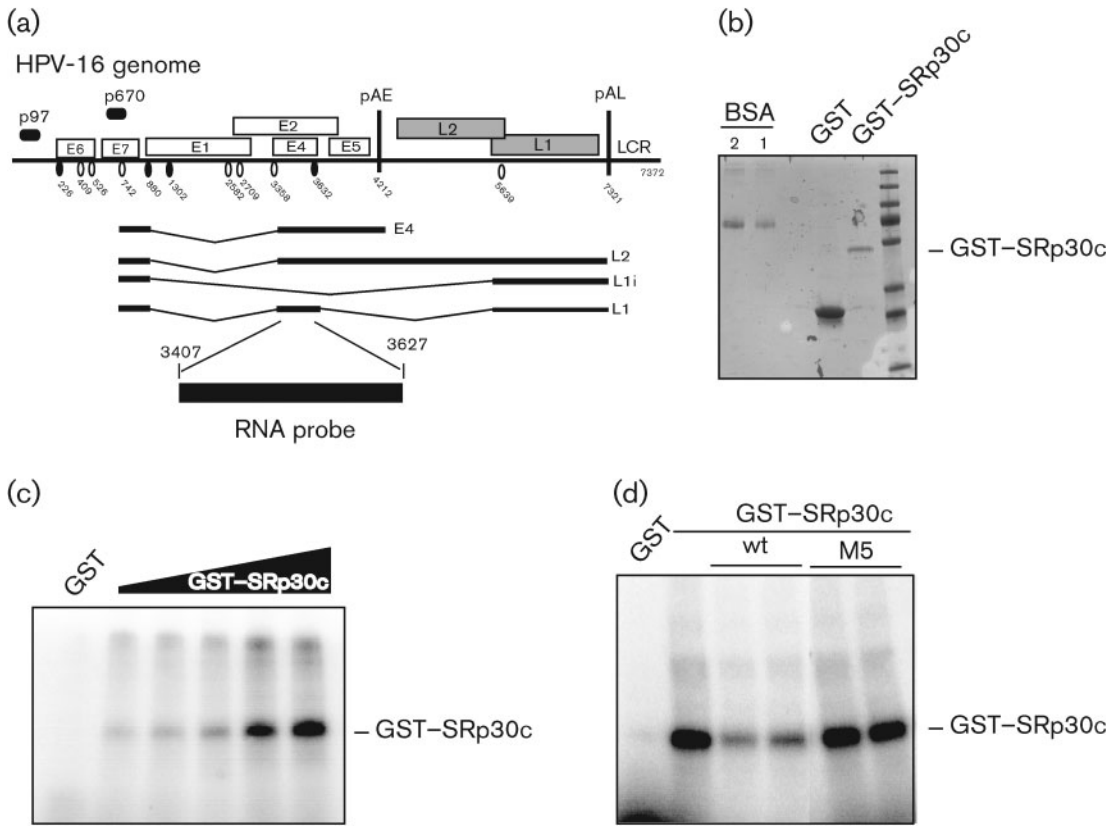

Fig. 3. (a) Schematic representation of the HPV-16 genome and position of the *in vitro*-synthesized HPV-16 RNA probe used for UV cross-linking. (b) Coomassie blue-stained acrylamide gel with purified recombinant GST or GST–SRp30c protein. BSA (2 and 1 μg, as indicated) was also loaded on the gel. (c) UV cross-linking of 2 μg GST or twofold serially diluted GST–SRp30c (starting at 2 μg) to the *in vitro*-synthesized, radiolabelled HPV-16 RNA probe shown in (a). (d) UV cross-linking of GST or GST–SRp30c to the radiolabelled HPV-16 RNA probe shown in (a) in the presence of a fourfold excess of cold HPV-16 RNA probe (wt) or a 20-fold excess of mutant HPV-1 AU-rich RNA element (M5).

(Fig. 4a, b), as reported previously (Somberg & Schwartz, 2010), whereas overexpression of WT ASF/SF2 at these levels showed no detectable effect on L1 mRNA expression, as expected (Fig. 4a). As an additional control, we monitored the effect of SRp55 on HPV-16 late gene expression. The results revealed that neither SRp55 nor SRp55DRS affected HPV-16 late gene expression (Fig. 4a). Overexpression of ASF/SF2, SRp30c and SRp55 was monitored by Western blotting (Fig. 4d). Collectively, our results demonstrated that SRp30c could activate L1i mRNA production by an exon-skipping mechanism.

Induction of HPV-16 L1 mRNA splicing from SD3632 to SA5639 requires the RS domain of SRp30c

Although we had confirmed that SRp30c could induce L1i production from pBELM by an exon-skipping mechanism, results on the WT pBEL plasmid shown in Fig. 2 demonstrated clearly that SRp30c also activated splicing from either late splice site SA5639 or SD3632, or both, as L1 mRNAs were produced from pBEL in the presence of exogenous SRp30c. These results suggested direct activation of one or both HPV-16 late splice sites. Next, we

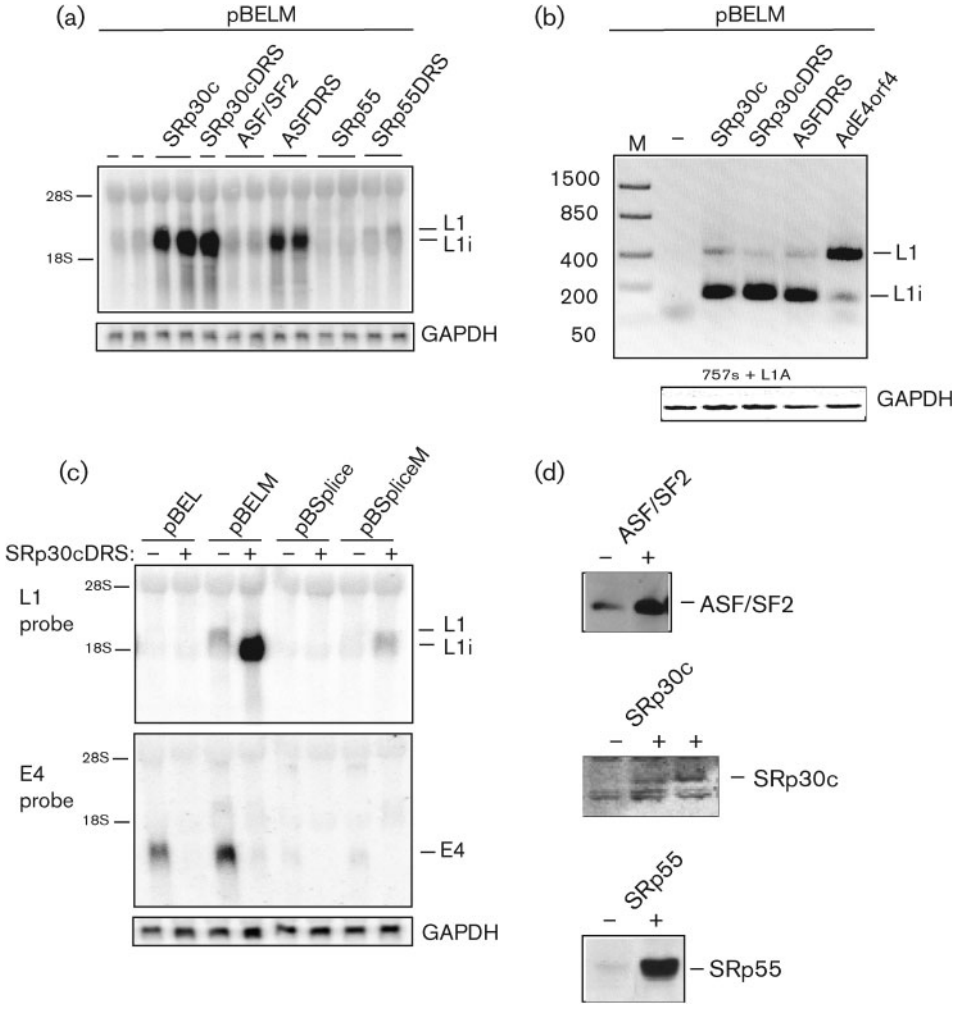

Fig. 4. (a) Northern blot of cytoplasmic RNA extracted from HeLa cells transfected with pBELM in the absence or presence of plasmids expressing SRp30c, SRp30cDRS, ASF/SF2, ASFDRS, SRp55 or SRp55DRS. Blots were probed with the L1 or GAPDH probe. The L1/L1i and GAPDH mRNAs and the 28S and 18S rRNAs are indicated. (b) RT-PCR with primers 757s and L1A (see Fig. 2a) on cDNA synthesized from HeLa cells transfected with pBELM and the indicated SR protein expression plasmids. cDNAs representing L1 or L1i mRNAs are indicated. M, Molecular size markers (bp). (c) Northern blots of cytoplasmic RNA extracted from HeLa cells transfected with pBEL, pBELM, pBSplice and pBSpliceM in the absence or presence of SRp30cDRS. Blots were probed with the L1 or E4 probe. For positions of probes, see Fig. 2(a). E4, L1, L1i and GAPDH mRNAs are indicated in the figure. (d) Western blots with antibody 1H4 specific for various SR proteins (Santa Cruz Biotechnology) or antiserum #65 (Somberg & Schwartz, 2010) demonstrating overexpression of various SR proteins.

investigated whether the RS domain of SRp30c was required for activation of splicing from WT plasmid pBEL. The results revealed that SRp30cDRS could not induce L1 mRNA production from pBEL, in which SA5639 was suppressed by downstream splicing silencers in the L1-encoding region (Fig. 4c), whereas high levels of L1 mRNA were induced from pBELM, in which the splicing silencers had been inactivated, as expected (Fig. 4c). However, SRp30cDRS did inhibit expression of E4 mRNA from both pBEL and pBELM (Fig. 4c). These results demonstrated that SRp30cDRS could inhibit SA3358, but, in contrast to WT SRp30c, it could not activate splicing to WT SA5639 in pBEL. We concluded that activation of late splice sites SD3632 and/or SA5639 by SRp30c required an intact RS domain, whereas inhibition of SA3358 did not.

To demonstrate that WT SRp30c could activate splice sites SD3632 and/or SA5639, and did not solely induce L1 mRNA production indirectly by inhibiting SA3358, we investigated the effect of SRp30c on HPV-16 plasmids pBSplice and pBSpliceM (Fig. 5a) (Somberg *et al.*, 2008). These plasmids are derived from pBEL and pBELM (Fig. 2a), respectively, and both contain a large deletion in the early region that removes all early splice sites including SA3358, whereas the two late HPV-16 splice sites SD3632 and SA5639 are intact (Fig. 5a). mRNAs that could potentially be produced from pBSplice and pBSpliceM are shown below the schematic representations of the plasmids in Fig. 5(a). We have shown previously that none of these plasmids expresses detectable levels of L1 mRNAs upon transfection of HeLa cells. This is presumably an effect of the efficient suppression of splice sites SD3632 and SA5639 in HeLa cells. However, co-transfection of pBSplice with SRp30c activated splicing from SD3632 to SA5639 to produce spliced L1 mRNAs (Fig. 5b). These results clearly demonstrated that SRp30c alleviates inhibition of SD3632 and/or SA5639. In addition, SRp30c activated splicing from pBSpliceM (Fig. 5a), in which splicing silencers at SA5639 had been inactivated, demonstrating that SRp30c activated splicing from SD3632 (Fig. 5b). In contrast to WT SRp30c, mutant SRp30cDRS did not induce L1 mRNA expression from pBSplice (Fig. 4c) and only low levels of L1 mRNAs were produced from pBSpliceM in the presence of SRp30cDRS (Fig. 4c). These results supported the idea that SRp30cDRS induced L1 mRNA expression solely by inhibiting SA3358, thereby redirecting splicing to SA5639. We concluded that activation of splicing between SD3632 and SA5639 by SRp30c required an intact RS domain and that WT SRp30c could both activate splicing from SD3632 and inhibit splicing to SA3358.

SRp30c interferes with splicing regulatory sequences at SD3632

To investigate whether SRp30c could activate splicing from pBspliceM when splicing inhibitory sequences at SD3632 had been deleted, we overexpressed SRp30c with a number of HPV-16 deletion mutants (Fig. 5a). SRp30c did not have a major effect on L1 mRNA production if sequences upstream (pMT1SD) or downstream (pBSpM3703–4530) of SD3632 were deleted and suppression of splicing at SD3632 was alleviated (Fig. 5c, d). In contrast, plasmids with smaller deletions downstream of SD3632 as in pMT3, or deletions in L2 as in pBSpM4288–4530 (Fig. 5a), which did not activate L1 (Fig. 5b), responded well to SRp30c (Fig. 5c, d). Taken together, these results indicated that SRp30c activates splicing of late splice site SD3632 by interfering with splicing inhibitory RNA elements in an RS domain-dependent manner.

Evidence that SRp30c overcomes suppression of both HPV-16 late splice sites, SD3632 and SA5639

Our results demonstrated that SRp30c activated splicing from the late 5′-splice site SD3632. We also wished to determine whether SRp30c acted directly on the late 3′-splice site SA5639. To test the effect of SRp30c on each splice site individually, we optimized SD3632 in the context of WT SA5639, and vice versa, in the background of plasmid pBEL (Fig. 6a). SD3632 was optimized either by improving the complementarity of SD3632 to U1 small nuclear ribonucleoprotein (U1snRNP) up to 100 %, or by introducing the T9 deletion upstream of SD3632, which has been shown to activate SD3632 (Rush *et al.*, 2005; Somberg *et al.*, 2008), resulting in plasmids pOPSD and pBELT9, respectively (Fig. 6a). If SRp30c activated only SD3632 and not SA5639, these plasmids should not respond to SRp30c. However, co-transfection of SRp30c with either pOPSD or pBELT9 resulted in induction of L1 mRNA production (Fig. 6b). RT-PCR analysis revealed that the induced mRNAs were primarily L1 mRNAs (Fig. 6c) and not L1i mRNA, demonstrating that SRp30c induced splicing from the optimized SD3632 to WT SA5639. Interestingly, SRp30c induced both L1i and L1 mRNA from pBEL, confirming that both SD3632 and SA5639 were activated by SRp30c (Fig. 6c). The lack of L1i mRNA production from pBELT9 and pOPSD is probably the result of a more efficient exon definition of the SA3358–SD3632 exon as a result of the improved SD3632 in these plasmids. As expected, SRp30c also induced L1 mRNA production from pBELM in which splicing silencers at SA5639 had been inactivated (Fig. 6b), as also seen in Fig. 2(d, e) and Fig. 4(a–c). In contrast, optimization of both SD3632 and SA5639, as in plasmids pT9 and pOPSDM, resulted in constitutive, high-level production primarily of L1 mRNAs (Fig. 6b, c). These L1 mRNA levels were not increased further by SRp30c, as expected (Fig. 6b, c). Therefore, SRp30c could activate splicing from both late splice sites, SD3632 and SA5639, on the HPV-16 genome. Taken together, our results established firmly that SRp30c induced L1 mRNA production by activating splicing from the two late splice sites, SD3632 and SA5639, and induced L1i mRNA production by activating SA5639, whilst at the same time inhibiting the competing splice site, SA3358.

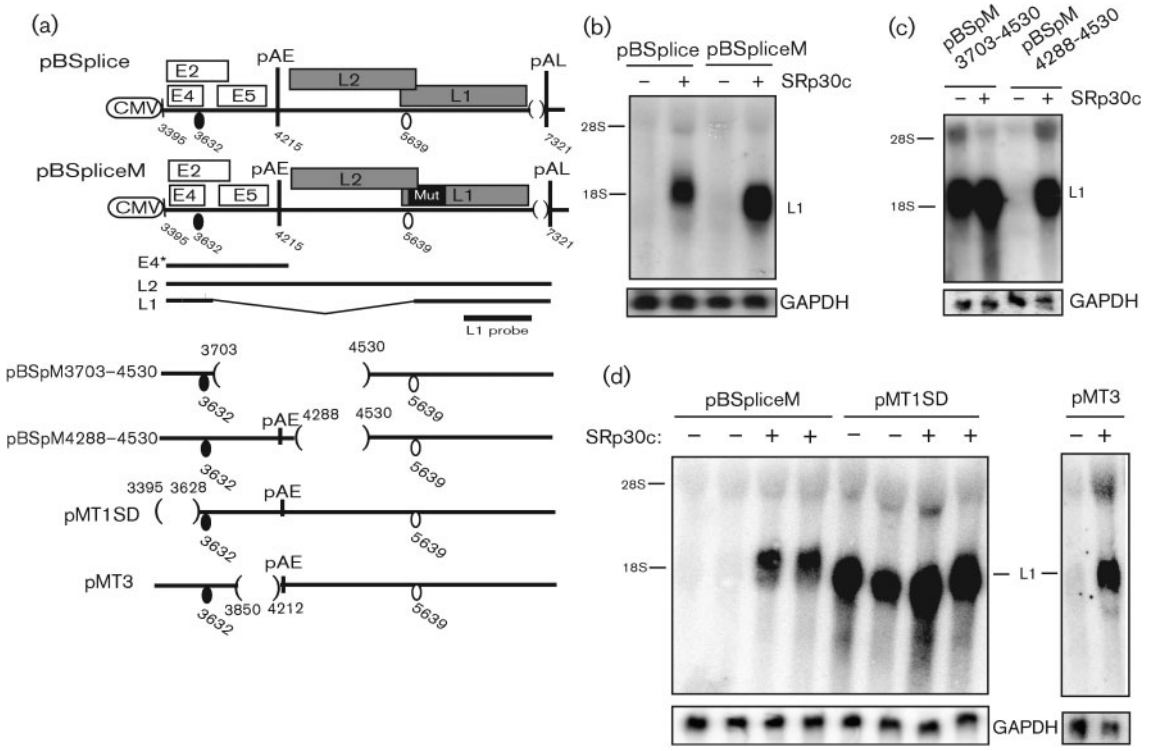

Fig. 5. (a) Schematic representation of subgenomic HPV-16 expression plasmids pBSplice and pBSpliceM (Somberg *et al.*, 2008). Mut indicates a number of nucleotide substitutions that inactivate splicing silencers downstream of SA5639 in plasmid pBSpliceM. mRNAs that may be produced by the plasmids are indicated. Deletion mutants pBSpM3703–4530, pBSpM4288–4530, pMT1SD and pMT3 are indicated. Numbers represent nucleotide positions of splice sites, poly(A) signals or borders of deletions and refer to the HPV-16R sequence (Baker & Calef, 1997). The position of the L1 probe is indicated. (b–d) Northern blots of cytoplasmic RNA extracted from HeLa cells transfected with the indicated plasmids in the absence or presence of SRp30c. Blots were probed with the L1 or GAPDH probes. The positions of the L1 and GAPDH mRNAs and 18S and 28S rRNAs are indicated. Note that in (d), the sizes of the bands differ as a result of the deletions.

DISCUSSION

SRp30c stimulated splicing of the late HPV-16 mRNAs, thereby promoting production of the spliced L1 mRNA. In HPV-16-infected cells, levels of SRp30c may therefore increase in response to differentiation, or factors that antagonize SRp30c, such as ASF/SF2, may decrease. In the latter case, it would be the relative levels of splicing factors that determine the outcome of the splicing events, rather than absolute levels of SRp30c (Fig. 7). We have shown previously that ASF/SF2 levels decrease in terminally differentiated cells and increase in high-grade cervical lesions and cervical cancer (Fay *et al.*, 2009). It would therefore be of interest to compare levels of individual SR proteins in individual cells with those of L1 protein levels in HPV-16-infected squamous epithelium. However, we are not aware of antibodies against SRp30c that work in immunohistochemistry.

The adenovirus E4orf4 protein decreases the phosphorylation of SR proteins in adenovirus-infected cells, thereby inactivating these proteins and causing an alteration in adenovirus alternative splicing that leads to induction of adenovirus late gene expression (Kanopka *et al.*, 1996, 1998). E4orf4 primarily binds ASF/SF2 (Estmer Nilsson *et al.*, 2001). Interestingly, adenovirus E4orf4 protein also induces HPV-16 late gene expression (Somberg *et al.*, 2009), indicating that dephosphorylation of SR proteins (or ASF/SF2, which is the primary partner of E4orf4) can activate HPV-16 late gene expression. Similarly to adenovirus, vaccinia virus infection also causes dephosphorylation of SR proteins in infected cells (Huang *et al.*, 2002). We have shown that SR proteins decrease in abundance as cervical epithelial cells differentiate (Fay *et al.*, 2009), perhaps inducing HPV-16 late gene expression in a manner similar to the inactivation of SR proteins by dephosphorylation by adenovirus E4orf4. Whether down-regulation of expression of SR proteins in differentiating cervical epithelium alters the relative ratios of active SR proteins in the infected cells, or whether SRp30c is unaffected or increases in abundance in differentiated cells, remains to be investigated.

High levels of SRp30c resulted in more L1i mRNAs at the expense of the L1 mRNAs. As the only difference between

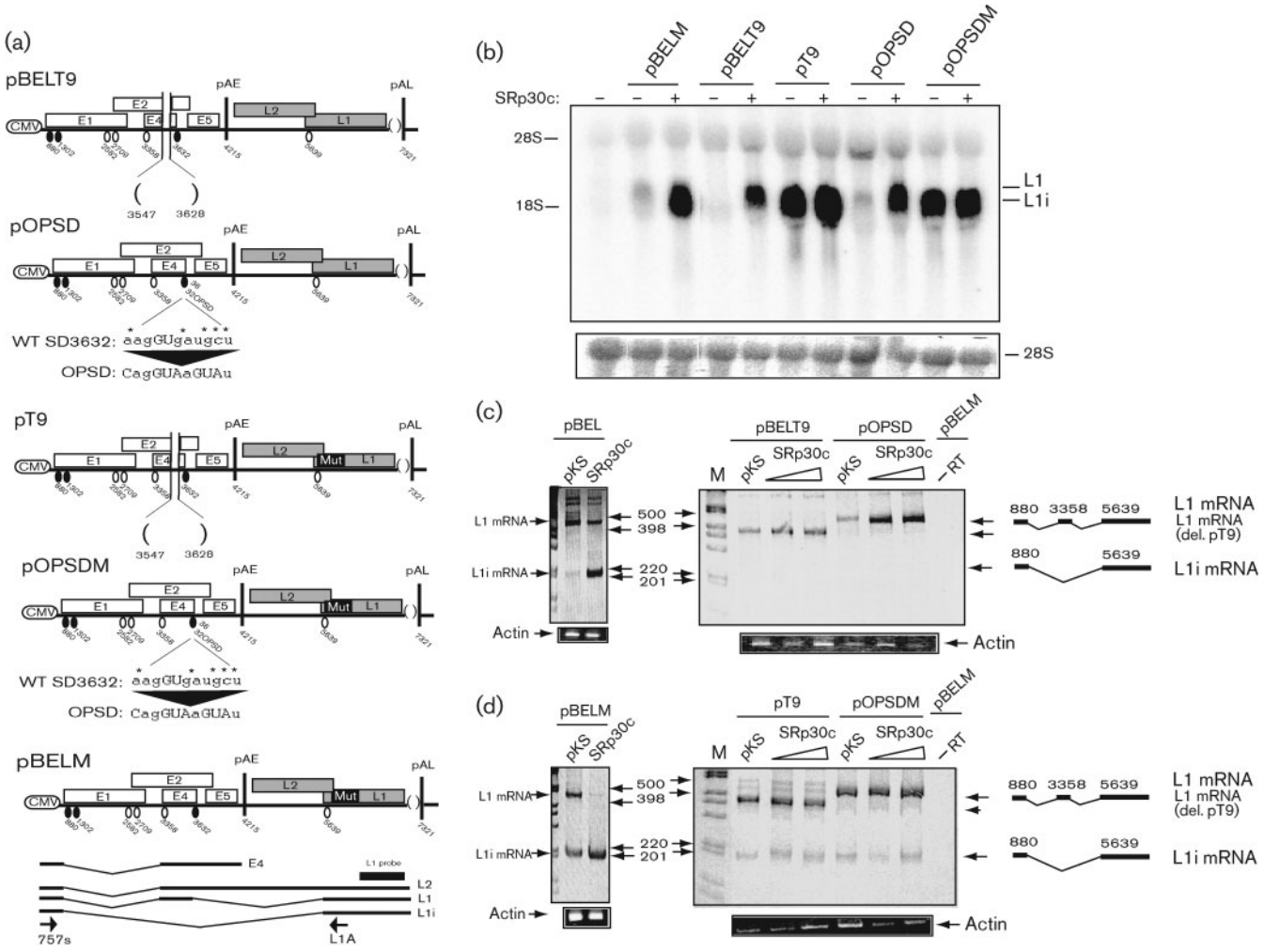

Fig. 6. (a) Schematic representation of the subgenomic HPV-16 expression plasmids pBELT9, pOPSD, pT9, pOPSDM and pBELM (Rush *et al.*, 2005; Zhao *et al.*, 2004). Numbers represent nucleotide positions of splice sites, poly(A) signals or borders of deletions. A subset of mRNAs that may be produced by the plasmids are indicated. The positions of the L1 probe and RT-PCR primers 757s and L1A are shown. (b) Northern blot of cytoplasmic RNA extracted from HeLa cells transfected with pBELM, pBELT9, pT9, pOPSD or pOPSDM in the absence or presence of SRp30c. The blot was hybridized to the L1 probe. L1 and L1i mRNAs and 28S and 18S rRNAs are indicated. Lane 1 shows the results of a mock transfection. Detection of 28R rRNA was used as a loading control. (c, d) RT-PCR with primers 757s and L1A on cDNA synthesized from cytoplasmic RNA extracted from HeLa cells transfected with pBEL, pBELT9 or pOPSD (c), or pBELM, pT9 or pOPSDM (d), in the presence of empty pBluecriptKS$^-$ vector (pKS) or SRp30c expression plasmid. The positions of bands representing L1 and L1i mRNAs are indicated. Note that L1 mRNAs produced by pBELT9 or pT9 were smaller than those produced by pOPSD and pOPSDM due to the 80 nt deletion in the E4-coding region. Detection of actin was used as a loading control.

these two mRNAs was that L1i mRNAs lack the 275 nt exon between SA3358 and SD3632, both L1 mRNAs should be translated efficiently into L1 protein. In addition, both L1 and L1i mRNAs have been observed in HPV-16-infected cells (Baker & Calef, 1997; Milligan *et al.*, 2007). The role of SRp30c during HPV-16 infection may therefore be to regulate HPV-16 late mRNA splicing and to determine the ratio between L2 and L1/L1i mRNAs.

Overexpression of SRp30c has been shown to cause redistribution of Y box-binding protein 1 (YB-1) into the nucleus, presumably allowing YB-1 to regulate mRNA splicing and/or transcription (Raffetseder *et al.*, 2003). Interestingly, YB-1 also interacts with hnRNP K and we have previously reported that hnRNP K binds directly to HPV-16 late mRNAs, immediately upstream of late 3′-splice site SA5639 (between SD3632 and SA5639) (Collier *et al.*, 1998). Other investigators have found SRp30c and poly(C)-binding protein (PCBP; also known as hnRNP E) present in the same RNA–protein complex that was formed on the 3′UTR of the human preprorenin mRNA (Skalweit *et al.*, 2003). We have shown previously that PCBP,

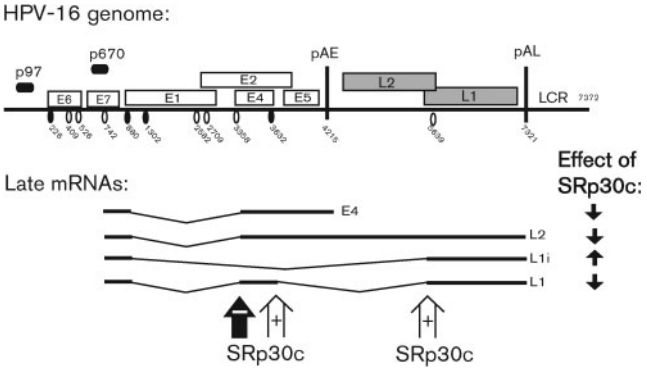

Fig. 7. Low levels of SRp30c stimulate splicing from both HPV-16 SD3632 and SA5639, which are used exclusively by the late L1 mRNAs. High levels of SRp30c promote exclusion of the exon between SA3358 and SD3632, indicating that SRp30c inhibits SA3358. High levels of SRp30c therefore favour production of L1i mRNAs over L1, L2 and E4 late HPV-16 mRNAs.

similarly to hnRNP K, binds specifically to the HPV-16 L2-encoding region, immediately upstream of HPV-16 late 3′-splice site SA5639 (Collier *et al.*, 1998). One may therefore speculate that SRp30c could exert its enhancing effect on HPV-16 late mRNA splicing at least partly through proteins that bind to both YB-1 and HPV-16 mRNAs, such as hnRNP K and PCBP.

METHODS

Plasmid constructs. The following plasmids have been described previously: pBEL (Zhao *et al.*, 2004), pBELM (Zhao *et al.*, 2004), pBearly (Zhao *et al.*, 2005), pBSplice (Somberg *et al.*, 2008), pBSpliceM (Somberg *et al.*, 2008), pMT3 (Somberg *et al.*, 2008), pBSpM4288–4530 (Somberg *et al.*, 2008), pBELT9 (Rush *et al.*, 2005), pT9 (Rush *et al.*, 2005), pOPSD (Rush *et al.*, 2005), pOPSDM (Rush *et al.*, 2005), pT7wtexon4 (Somberg & Schwartz, 2010) and pMT1SD (Somberg & Schwartz, 2010).

pBSpliceD3703-4530 was created by PCR-amplifying a sequence from pBR-HPV-16 with primers E4S (Zhao *et al.*, 2004) and M3A (see Supplementary Table S1, available in JGV Online). The PCR fragment was then subcloned into pCR2.1-TOPO, released with *Bss*HII and *Apa*I and subcloned into pBSpliceM. CMV-SRp30cDRS and CMV-SRp55DRS were created by PCR-amplifying sequences from CMV-SRp30c or CMV-SRp55 with primers 30cS and 30cΔRSA, and SRp55S and SRp55RRMAS, respectively (Supplementary Table S1). The PCR fragments were then cleaved with *Bss*HII and *Xho*I and subcloned into pC0806 (Collier *et al.*, 2002). pGST-SRp30c was created by PCR-amplifying the sequence from pSRp30c with primers SRP30C-S and SRP30C-AS (Supplementary Table S1). The PCR fragment was then cleaved with *Bam*HI and *Eco*RI and subcloned into pGEX-GST, resulting in pGST-SRp30c.

pASF/SF2, pASFDRS, pSRp30c, pSRp55 and pAdE4orf4 were kind gifts from Dr Göran Akusjärvi (IMBIM, Uppsala University, Sweden).

Transfection and cell culture. HeLa cells were cultured in Dulbecco's modified Eagle's medium containing 10 % heat-inactivated FCS, 100 U penicillin ml^{-1} and 100 μg streptomycin ml^{-1}. Transfections were carried out using Fugene 6 (Roche Molecular Biochemicals) according to the manufacturer's protocol, as described previously (Somberg *et al.*, 2008). All plasmids were transfected in a minimum of three independent experiments with similar results.

RNA extraction, Northern blotting and radiolabelled DNA probe synthesis. Cytoplasmic RNA was extracted using IsoB/NP-40 buffer, as described previously (Wiklund *et al.*, 2002). Northern blotting was carried out by size separation of 10 μg cytoplasmic RNA on 1.2 % agarose gels containing 2.2 M formaldehyde, followed by transfer overnight to nitrocellulose filters subsequently hybridized with an L1, E4 or GAPDH probe. The L1 probe was excised from pBEL with *Bam*HI and *Xho*I, whereas the E4 probe was generated by PCR using primers E4S and K1 (Somberg & Schwartz, 2010). The GAPDH probe was generated as described previously (Somberg & Schwartz, 2010). DNA probes were radiolabelled with [α-^{32}P]dCTP using a Decaprime kit (Ambion).

RT-PCR. Cytoplasmic RNA (200 ng) was reverse-transcribed at 42 °C using SuperScript II and random hexamers according to the manufacturer's protocol (Invitrogen). The cDNA (2 μl) was amplified in a 100 μl PCR using primers 757s and L1A or L1aM, as described previously (Zhao *et al.*, 2004). GAPDH cDNA was amplified as a control using previously described primers (Johansson *et al.*, 2005).

Purification of recombinant GST proteins and UV cross-linking. Purification of recombinant GST-fusion proteins using glutathione–Sepharose was performed as specified by the manufacturer (GE Healthcare). Plasmid pT7wtexon4 was linearized with *Xba*I and subjected to *in vitro* transcription with T7 RNA polymerase in the presence of [α-^{32}P]UTP. The RNA probe (10^5 c.p.m.) was UV cross-linked to GST and GST-fusion proteins, as described previously (Sokolowski *et al.*, 1999).

ACKNOWLEDGEMENTS

We wish to thank the members of the groups of Göran Magnusson, Göran Akusjärvi, Catharina Svensson and S. S. for providing reagents and for discussions at laboratory meetings. This research was sponsored by the Technological Sector Research Study III (Ireland), Science Foundation Ireland, Swedish Cancer Society, Swedish Research Council/Medicine and Linneus Support from the Swedish Research Council to Uppsala RNA Research Center.

REFERENCES

Baker, C. & Calef, C. (1997). Maps of papillomavirus mRNA transcripts. In *Human Papillomaviruses: a Compilation and Analysis of Nucleic Acid and Amino Acid Sequences*. Edited by S. R. Billakanti, C. E. Calef, A. D. Farmer, A. L. Halpern & G. L. Myers. Los Alamos, NM: Los Alamos National Laboratory.

Collier, B., Goobar-Larsson, L., Sokolowski, M. & Schwartz, S. (1998). Translational inhibition *in vitro* of human papillomavirus type 16 L2 mRNA mediated through interaction with heterogenous ribonucleoprotein K and poly(rC)-binding proteins 1 and 2. *J Biol Chem* **273**, 22648–22656.

Collier, B., Öberg, D., Zhao, X. & Schwartz, S. (2002). Specific inactivation of inhibitory sequences in the 5′ end of the human papillomavirus type 16 L1 open reading frame results in production of high levels of L1 protein in human epithelial cells. *J Virol* **76**, 2739–2752.

Cumming, S. A., Repellin, C. E., McPhillips, M., Radford, J. C., Clements, J. B. & Graham, S. V. (2002). The human papillomavirus

type 31 late 3′ untranslated region contains a complex bipartite negative regulatory element. *J Virol* **76**, 5993–6003.

Doorbar, J. (2005). The papillomavirus life cycle. *J Clin Virol* **32** (Suppl. 1,), S7–S15.

Estmer Nilsson, C., Petersen-Mahrt, S., Durot, C., Shtrichman, R., Krainer, A. R., Kleinberger, T. & Akusjärvi, G. (2001). The adenovirus E4-ORF4 splicing enhancer protein interacts with a subset of phosphorylated SR proteins. *EMBO J* **20**, 864–871.

Fay, J., Kelehan, P., Lambkin, H. & Schwartz, S. (2009). Increased expression of cellular RNA-binding proteins in HPV-induced neoplasia and cervical cancer. *J Med Virol* **81**, 897–907.

Goraczniak, R. & Gunderson, S. I. (2008). The regulatory element in the 3′-untranslated region of human papillomavirus 16 inhibits expression by binding CUG-binding protein 1. *J Biol Chem* **283**, 2286–2296.

Graham, S. V. (2008). Papillomavirus 3′ UTR regulatory elements. *Front Biosci* **13**, 5646–5663.

Howley, P. M. & Lowy, D. R. (2001). Papillomaviruses and their replication. In *Fields Virology*, 4th edn, pp. 2197–2229. Edited by D. M. Knipe & P. M. Howley. Philadelphia, PA: Lippincott Williams & Wilkins.

Huang, T. S., Nilsson, C. E., Punga, T. & Akusjärvi, G. (2002). Functional inactivation of the SR family of splicing factors during a vaccinia virus infection. *EMBO Rep* **3**, 1088–1093.

Jia, R., Liu, X., Tao, M., Kruhlak, M., Guo, M., Meyers, C., Baker, C. C. & Zheng, Z. M. (2009). Control of the papillomavirus early-to-late switch by differentially expressed SRp20. *J Virol* **83**, 167–180.

Johansson, C., Zhao, H., Bajak, E., Granberg, F., Pettersson, U. & Svensson, C. (2005). Impact of the interaction between adenovirus E1A and CtBP on host cell gene expression. *Virus Res* **113**, 51–63.

Kanopka, A., Mühlemann, O. & Akusjärvi, G. (1996). Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature* **381**, 535–538.

Kanopka, A., Mühlemann, O., Petersen-Mahrt, S., Estmer, C., Öhrmalm, C. & Akusjärvi, G. (1998). Regulation of adenovirus alternative RNA splicing by dephosphorylation of SR proteins. *Nature* **393**, 185–187.

Koffa, M. D., Graham, S. V., Takagaki, Y., Manley, J. L. & Clements, J. B. (2000). The human papillomavirus type 16 negative regulatory RNA element interacts with three proteins that act at different posttranscriptional levels. *Proc Natl Acad Sci U S A* **97**, 4677–4682.

Longworth, M. S. & Laimins, L. A. (2004). Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol Mol Biol Rev* **68**, 362–372.

McPhillips, M. G., Veerapraditsin, T., Cumming, S. A., Karali, D., Milligan, S. G., Boner, W., Morgan, I. M. & Graham, S. V. (2004). SF2/ASF binds the human papillomavirus type 16 late RNA control element and is regulated during differentiation of virus-infected epithelial cells. *J Virol* **78**, 10598–10605.

Milligan, S. G., Veerapraditsin, T., Ahamet, B., Mole, S. & Graham, S. V. (2007). Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells. *Virology* **360**, 172–181.

Oberg, D., Collier, B., Zhao, X. & Schwartz, S. (2003). Mutational inactivation of two distinct negative RNA elements in the human papillomavirus type 16 L2 coding region induces production of high levels of L2 in human cells. *J Virol* **77**, 11674–11684.

Oberg, D., Fay, J., Lambkin, H. & Schwartz, S. (2005). A downstream polyadenylation element in human papillomavirus type 16 L2 encodes multiple GGG motifs and interacts with hnRNP H. *J Virol* **79**, 9254–9269.

Raffetseder, U., Frye, B., Rauen, T., Jürchott, K., Royer, H.-D., Jansen, P. L. & Mertens, P. R. (2003). Splicing factor SRp30c interaction with Y-box protein-1 confers nuclear YB-1 shuttling and alternative splice site selection. *J Biol Chem* **278**, 18241–18248.

Rush, M., Zhao, X. & Schwartz, S. (2005). A splicing enhancer in the E4 coding region of human papillomavirus type 16 is required for early mRNA splicing and polyadenylation as well as inhibition of premature late gene expression. *J Virol* **79**, 12002–12015.

Schwartz, S. (2008). HPV-16 RNA processing. *Front Biosci* **13**, 5880–5891.

Skalweit, A., Doller, A., Huth, A., Kähne, T., Persson, P. B. & Thiele, B. J. (2003). Posttranscriptional control of renin synthesis: identification of proteins interacting with renin mRNA 3′-untranslated region. *Circ Res* **92**, 419–427.

Sokolowski, M. & Schwartz, S. (2001). Heterogeneous nuclear ribonucleoprotein C binds exclusively to the functionally important UUUUU-motifs in the human papillomavirus type-1 AU-rich inhibitory element. *Virus Res* **73**, 163–175.

Sokolowski, M., Zhao, C., Tan, W. & Schwartz, S. (1997). AU-rich mRNA instability elements on human papillomavirus type 1 late mRNAs and c-*fos* mRNAs interact with the same cellular factors. *Oncogene* **15**, 2303–2319.

Sokolowski, M., Furneaux, H. & Schwartz, S. (1999). The inhibitory activity of the AU-rich RNA element in the human papillomavirus type 1 late 3′ untranslated region correlates with its affinity for the elav-like HuR protein. *J Virol* **73**, 1080–1091.

Somberg, M. & Schwartz, S. (2010). Multiple ASF/SF2 sites in the human papillomavirus type 16 (HPV-16) E4-coding region promote splicing to the most commonly used 3′-splice site on the HPV-16 genome. *J Virol* **84**, 8219–8230.

Somberg, M., Zhao, X., Fröhlich, M., Evander, M. & Schwartz, S. (2008). Polypyrimidine tract binding protein induces human papillomavirus type 16 late gene expression by interfering with splicing inhibitory elements at the major late 5′ splice site, SD3632. *J Virol* **82**, 3665–3678.

Somberg, M., Rush, M., Fay, J., Ryan, F., Lambkin, H., Akusjärvi, G. & Schwartz, S. (2009). Adenovirus E4orf4 induces HPV-16 late L1 mRNA production. *Virology* **383**, 279–290.

Terhune, S. S., Hubert, W. G., Thomas, J. T. & Laimins, L. A. (2001). Early polyadenylation signals of human papillomavirus type 31 negatively regulate capsid gene expression. *J Virol* **75**, 8147–8157.

Tranell, A., Fenyö, E. M. & Schwartz, S. (2010). Serine- and arginine-rich proteins 55 and 75 (SRp55 and SRp75) induce production of HIV-1 *vpr* mRNA by inhibiting the 5′-splice site of exon 3. *J Biol Chem* **285**, 31537–31547.

Wiklund, L., Sokolowski, M., Carlsson, A., Rush, M. & Schwartz, S. (2002). Inhibition of translation by UAUUUAU and UAUUUUUAU motifs of the AU-rich RNA instability element in the HPV-1 late 3′ untranslated region. *J Biol Chem* **277**, 40462–40471.

Zhao, X., Rush, M. & Schwartz, S. (2004). Identification of an hnRNP A1-dependent splicing silencer in the human papillomavirus type 16 L1 coding region that prevents premature expression of the late L1 gene. *J Virol* **78**, 10888–10905.

Zhao, X., Öberg, D., Rush, M., Fay, J., Lambkin, H. & Schwartz, S. (2005). A 57-nucleotide upstream early polyadenylation element in human papillomavirus type 16 interacts with hFip1, CstF-64, hnRNP C1/C2, and polypyrimidine tract binding protein. *J Virol* **79**, 4270–4288.

Zhao, X., Rush, M., Carlsson, A. & Schwartz, S. (2007). The presence of inhibitory RNA elements in the late 3′-untranslated region is a conserved property of human papillomaviruses. *Virus Res* **125**, 135–144.

Zheng, Z. M. & Baker, C. C. (2006). Papillomavirus genome structure, expression, and post-transcriptional regulation. *Front Biosci* **11**, 2286–2302.

zur Hausen, H. (2002). Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer* **2**, 342–350.