Published online 31 March 2022

Nucleic Acids Research, 2022, Vol. 50, No. 7 ***3867–3891***
https://doi.org/10.1093/nar/gkac213

hnRNP G/RBMX enhances HPV16 E2 mRNA splicing through a novel splicing enhancer and inhibits production of spliced E7 oncogene mRNAs

Chengyu Hao[2], Yunji Zheng[2,4], Johanna Jönsson[1,2], Xiaoxu Cui[2], Haoran Yu[2], Chengjun Wu[3], Naoko Kajitani[1,2] and Stefan Schwartz[1,2,*]

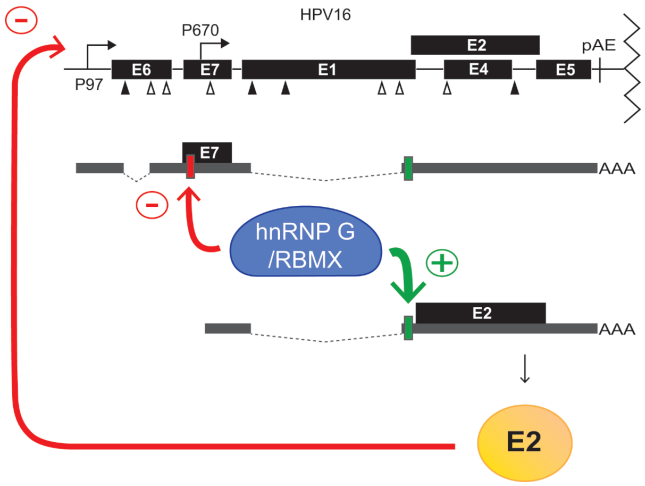
[1]Department of Medical Biochemistry and Microbiology (IMBIM), Uppsala University, BMC-B9, 751 23 Uppsala, Sweden, [2]Department of Laboratory Medicine, Lund University, BMC-B13, 221 84 Lund, Sweden, [3]School of Biomedical Engineering, Dalian University of Technology, Liaoning IC Technology Key Lab, 116024 Dalian, China and [4]School of Pharmacy, Binzhou Medical University, 264003 Yantai, China

Received November 23, 2021; Revised March 11, 2022; Editorial Decision March 15, 2022; Accepted March 18, 2022

ABSTRACT

Human papillomavirus type 16 (HPV16) E2 is an essential HPV16 protein. We have investigated how HPV16 E2 expression is regulated and have identified a splicing enhancer that is required for production of HPV16 E2 mRNAs. This uridine-less splicing enhancer sequence (ACGAGGACGAGGACAAGGA) contains 84% adenosine and guanosine and 16% cytosine and consists of three 'AC(A/G)AGG'-repeats. Mutational inactivation of the splicing enhancer reduced splicing to E2-mRNA specific splice site SA2709 and resulted in increased levels of unspliced E1-encoding mRNAs. The splicing enhancer sequence interacted with cellular RNA binding protein hnRNP G that promoted splicing to SA2709 and enhanced E2 mRNA production. The splicing-enhancing function of hnRNP G mapped to amino acids 236–286 of hnRNP G that were also shown to interact with splicing factor U2AF65. The interactions between hnRNP G and HPV16 E2 mRNAs and U2AF65 increased in response to keratinocyte differentiation as well as by the induction of the DNA damage response (DDR). The DDR reduced sumoylation of hnRNP G and pharmacological inhibition of sumoylation enhanced HPV16 E2 mRNA splicing and interactions between hnRNP G and E2 mRNAs and U2AF65. Intriguingly, hnRNP G also promoted intron retention of the HPV16 E6 coding region thereby inhibiting production of spliced E7 oncogene mRNAs.

GRAPHICAL ABSTRACT



INTRODUCTION

A subset of the human papillomaviruses (HPV) with tropism for mucosal epithelial cells can establish persistent infections that may progress to cancer (1). HPV16 is the most common high risk HPV types and is associated with various anogenital cancers as well as head and neck cancer (1,2). The life-cycle of HPV is strictly linked to cell differentiation in the squamous epithelium with expression of the HPV early genes in the lower and middle layers of the stratified epithelium (3–6). Essential early proteins include E6 and E7 that prevent apoptosis and induce proliferation of the HPV-infected cell (7,8). Simultaneous expression of the two HPV DNA replication factors E1 and E2 ensures replication of the HPV DNA genome (9–12). As cell dif-

*To whom correspondence should be addressed. Tel: +46 18 4714444; Mob: +46 73 9806233; Fax: +46 18 4714673; Email: stefan.schwartz@imbim.uu.se

ferentiation proceeds, a switch from the HPV early to the late gene expression program is induced (3,4). This switch includes a promoter switch, a polyA site switch and a shift in HPV alternative mRNA splicing (13). The HPV replication and transcription factor E2 plays a key role in this switch as it downregulates the HPV early promoter and inhibits the HPV early polyadenylation signal to induce HPV late gene expression (14). Apart from its role as a DNA replication factor, HPV E2 is also a transcriptional regulator of the HPV early genes, including the E6 and E7 oncogenes (10–12,15–18). Low levels of papillomavirus E2 protein stimulate transcription from the HPV early promoter, whereas high levels of E2, shut down the early promoter and E6 and E7 expression (10–12). Thus, E2 can induce apoptosis in cancer cells (19,20). It is therefore not surprising that the HPV E2 gene is inactivated in many HPV-driven cancers (21). Beyond its role as a transcriptional activator, E2 also interacts with Serine-Arginine rich proteins (22,23) and inhibits in vitro splicing (23) and with polyadenylation factors to inhibit polyadenylation in cells as well as *in vitro* (14). Furthermore, during division of the HPV-infected cell, E2 contributes to proper partition of the HPV genomes to both daughter cells (24). Despite the many and important roles of E2 in the papillomavirus life cycle as well as the suggested role in carcinogenesis, regulation of E2 expression has not been studied in detail. HPV16 E2 mRNAs can be made from both HPV16 early and late promoters (25–27). It is the HPV16 mRNA that is initiated at late promoter P670 and is spliced from SD880 to SA2709 that produces the highest levels of E2 protein of various alternatively spliced E2-encoding mRNAs (28). An alternative upstream 3′-splice site named SA2582 can be used as well, but the resulting mRNAs produce less E2 protein than those spliced to SA2709 (28). Despite the essential role of E2 in the HPV life cycle, it is currently not known how the major E2 splice site SA2709 is selected for splicing and how it is regulated.

Previous research has shown that some HPV16 splice sites are controlled by adjacent cis-acting RNA elements that either enhance or reduce splicing to a specific splice site, so called splicing enhancer and silencer elements (13,29,30). These RNA elements interact with cellular RNA binding proteins that control the utilization efficiency of each splice site, thereby playing a major role in the regulated expression of the HPV genes at the level of RNA processing (31). Recent results have shown that the HPV16 mRNAs contain hot spots for RNA binding proteins that are often located near HPV16 splice sites (32). Many of these hot spots interact with the ubiquitous cellular RNA binding protein hnRNP L in an Akt kinase dependent manner (32). Other HPV16 regulatory RNA elements bind directly to hnRNP A1 and A2 and control either HPV16 early or late gene expression (33–36). The many RNA binding proteins that interact with HPV16 mRNAs may either cooperate or counteract each other to control HPV16 mRNA splicing and gene expression (37). One may speculate that the HPV16 E2-specific splice site SA2709 is under control of cis-acting HPV16 RNA elements and cellular trans-acting, RNA binding proteins, which prompted the investigation described herein.

MATERIALS AND METHODS

Plasmids

The following plasmids have been described previously: pBEL (38), pBELsL (39,40), pC97ELsLuc (39,40), pHPV16AN (40), hnRNP G (41), phnRNP A1 (33), phnRNP A2 (33), phnRNP D (40), pX856F (33), p556F (33) and pCL086 (42). The RBM14 expression plasmid pCDNA3.2/V5-DESTRBM14-V5 was purchased from Addgene (#69823), the c-myc-tagged, pCMV6entry-derived RBM15 expression plasmid pRBM15b (RC222622) and the c-myc-tagged, pCMV6entry-derived hnRNP M expressing plasmid (RC220513) were purchased from Origene.

To construct pD200 that has a deletion between HPV16 nucleotide positions 2731 and 2922, HPV16 sequences were first PCR-amplified from plasmid pC97EL using primers CMVSAL and E2asxba (PCR-primers for plasmid constructions are listed in Supplementary Table S1). The PCR fragment was digested with SalI and CsiI and subcloned into pC97ELsLuc digested with the same two enzymes thereby generating pD200 that has a deletion in the E2 coding sequence and a newly introduced XbaI site immediately upstream of CsiI. To generate pBELAS, the sequence between positions 2731 and 2922 was PCR amplified with primers ANTI-F and ANTI-R followed by digestion with XbaI and CsiI and insertion in antisense orientation between XbaI and CsiI in pD200. To generate pBELXC, the sequence between positions 2731 and 2922 was PCR amplified with primers XBA2730 and E42A followed by digestion with XbaI and CsiI and insertion into pD200. These subcloning steps generated a pBEL plasmid with an XbaI site inserted at position 2721. To generate pBELCM, the sequence between positions 2731 and 2922 was codon-modified downstream of the E2 ATG at position 2756 and inserted between XbaI and CsiI in pD200. To generate plasmid p2758, HPV16 sequences were PCR-amplified with primers X2758F and E42A followed by digestion with XbaI and PteI and insertion into pD200. The PteI site has previously been inserted at HPV16 nucleotide position 3400 in pBEL (38). To generate plasmid p15, HPV16 sequences were PCR-amplified with primers CMVSAL and 2830as followed by digestion with SalI and CsiI and subcloning into pBEL. To generate plasmid p16, HPV16 sequences were PCR-amplified with primers XBA2730 and 2880as followed by digestion with XbaI and CsiI and subcloning into pBELD200. To generate plasmids p2780, p2830, p2880, p3030, p3130 and p3211, HPV16 sequences were first PCR amplified with primers XBA2780, XBA2830, XBA2880, XBA3030, XBA3130 or XBA3211 and E42A, digested with XbaI and PteI and inserted into plasmid pBEL. To generate pM1, pM2, pM3, pM12, pM13, pM23 and pMALL, HPV16 sequences were amplified with PCR-primers M2758-1S, M2758-2S, M2758-3S, M2758-12S, M2758-13S, M2758-23S or M2758-ALLS and E42A, digested with XbaI and PteI and inserted into plasmid pD200.

To construct HPV16 subgenomic plasmid pBELEN, pBELsluc plasmid was cut with restriction enzymes CsiI and XhoI, followed by blunting of sticky DNA ends

with Klenow fragment and re-ligation. To construct pBELENdE1, pBELEN was cut with restriction enzymes AdeI and NsiI to delete sequences in the E1 coding region, followed by blunting of sticky DNA ends with Klenow fragment and re-ligation. To generate pBELENdEN and pBELENdE1dEN, a synthetic DNA fragment (Eurofins Genomics) ranging from HPV16 nucleotide position 2481 to 2922 carrying an 18-nucleotide deletion of HPV16 sequences between nucleotides 2731 and 2750 was inserted into pBELEN and pBELENdE1, respectively, using restriction enzymes NsiI and CsiI. To generate pC97ELdEN a synthetic KasI-CsiI fragment (Eurofins Genomics) with an 18-nucleotide deletion of HPV16 sequences between nucleotides 2731 and 2750 was digested with KasI and CsiI and inserted into pC97ELsLuc. To generate pHPV16ANdEN an SdaI-CsiI-DNA fragment was excised from pC97ELdEN and inserted into pHPV16AN digested with the same restriction enzymes.

To construct HPV16 subgenomic plasmids pX765, pX656 and pX478, primer B97S was used in combination with primer X765A, X656A or X478A to amplify HPV16 sequences that were digested with PteI and XhoI and inserted between the PteI and XhoI sites in CMV promoter driven empty vector pCL086 (42). HPV16 subgenomic plasmid pHAE6E7Flag encodes HPV16 E6 and E7 genes (HPV16 nucleotide positions 104–855) (HPV16 nucleotide positions refer to the HPV16 reference sequence HPV16R (GeneBank: K02718.1)). The E6 open reading frame was fused to an HA tag sequence at the 5′-end and the E7 open reading frame was fused to a flag tag sequence at the 3′-end. pHAE6E7Flag was generated by insertion of a PCR-amplified DNA fragment produced with PCR primers: 'sense with HA tag' and 'antisense with flag tag' into pCL086 (42) using PteI and XhoI. pBEE2 was constructed by first PCR-amplifying E2 sequences from the pHPV16AN plasmid with primers 2917S and Xhol-E2-Rv followed by digestion of PCR-fragments with CsiI and XhoI and subcloning into pBEL digested with the same enzymes. pBEE2-END was constructed by first PCR-amplifying E2 sequences from pC97ELdEN plasmid with primers 2088S and E2Qas followed by digestion of PCR-fragments with CsiI and KpnI and subcloning into pBEE2 digested with the same enzymes. Plasmid p16E1-3xF encodes the HPV16 E1 gene (HPV16 nucleotide positions 865–2811) fused with a 3xFLAG tag sequence at the 3′-end and was constructed by a 2-step PCR amplification, first with primers PteI-E1-Fw and E1-3xFLAGmiddle-Rv and next with primers PteI-E1-Fw and 3xFLAG-middle-XhoI-Rv, followed by cloning into plasmid vector pCL086 using PteI and XhoI.

To generate pGM1, pGM2, pGM4, pGM5, pGM6, pGM7, pGM8, pGM9 and pGM10, hnRNP G sequences were PCR amplified from wild type hnRNP G plasmid with primers PteI-FLAG-hnRNPG-Fw and NTDas, RBDstartas, GM4as (aa83), GM5as (aa95), GM6as (aa127), GM7as (aa186), GM8as (aa265), GM9as (aa286) or GM10as (aa310) followed by digestion of PCR-fragments with PteI and XhoI and subcloning into pCL086 digested with the same enzymes. To generate pGM3, hnRNP G sequences were PCR amplified from wild type hnRNP G plasmid with primers RRMendS and hnRNPG RBas2 followed by digestion of PCR-fragments with PteI and XhoI and subcloning into pCL086 digested with the same enzymes. To generate pGM11, hnRNP G sequences were PCR amplified from wild type hnRNP G plasmid with primers PteI-flag-aa109-Fw and Xhol-aa236as followed by digestion of PCR-fragments with PteI and XhoI and subcloning into pCL086 (42) digested with the same enzymes.

Cells

HeLa cells, C33A2 cells and SiHa cells were cultured in Dulbecco's Modified Eagle Medium supplemented with 10% bovine calf serum and 1% penicillin-streptomycin. The C33A2 cell line is derived from the HPV-negative cervical cancer cell line C33A and contains the subgenomic HPV16 plasmid pBELsLuc stably integrated into the genome. pBELsLuc contains a gene segment encoding poliovirus 2A internal ribosome entry site together with the *Metridia longa* secreted luciferase (sLuc) gene in the L1 coding region of HPV16 (40). Induction of HPV16 late gene expression results in the appearance of sLuc in the cell culture medium. HPV16-infected tonsillar cancer cell line HN26 has been described previously (43). Briefly, the HN26 cells are derived from a tumor of a 48-year-old nonsmoking man with non-keratinizing, HPV16-positive tonsil oral squamous cell carcinoma, stage T2N0M0. The HN26 cells contain episomal HPV16 DNA and have an intact p53 gene. HN26 cells were cultured in RPMI 1640 medium (HyClone) with 10% iron-supplemented bovine calf serum (HyClone), 5%MEM Non-essential Amino Acid Solution (Sigma Aldrich) and 5% sodium pyruvate (Sigma Aldrich). The HPV16-immortalized keratinocyte cell line 3310 cell line has been described previously and was generated by stable transfection of normal neonatal human foreskin keratinocytes (nHFK) with HPV16 genome plasmid pHPV16ANE2fs (44). The 3310 cells were cultured in EpiLife medium (Gibco) supplemented with 1% human keratinocyte growth supplement (HKGS, Gibco) and 0.2% Gentamicin/Amphotericin (Gibco). Differentiation of 3310 cells was induced by addition of $CaCl_2$ at a final concentration of 2.4 mM in the keratinocyte culture medium for 24 hrs. For treatment of C33A2 cells or HN26 cells with melphalan, cell culture medium was replaced with medium containing indicated concentrations of melphalan for indicated time periods. Melphalan hydrochloride (Y0001158, Sigma) was dissolved in DMSO and DMSO void of melphalan was used as a control in all experiments.

Transfections

Transfections of HeLa cells were carried out using Turbofect according to the manufacturer's instructions (Thermo Scientific). Turbofect was mixed with plasmid DNA and incubated at room temperature for 15 min prior to drop-wise addition to 60-mm plates with subconfluent HeLa cells. Cells were harvested at 20hrs post transfection. Each plasmid was transfected in triplicate, in a minimum of two independent experiments.

Subcellular fractionation

Nuclear and cytoplasmic extracts were prepared from HeLa cells grown in 10 cm dishes. Cells were harvested at 24 h post-transfection by scraping and then spun at 1500 rpm for

5 min in an Eppendorf centrifuge. Cell pellets were resuspended in 250 ul ice-cold Buffer l (10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EGTA, 1 mM DTT and ptotease inhibitor) and allowed to swell for 15 min on ice. Then, 250 ul of ice-cold NP40-Buffer I (10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EGTA, 1 mM DTT, 0.8% NP-40 and protease inhibitor) was added. When lysis was observed under the microscope, samples were spun at 2000 rpm for 10min. Supernatants were collected and 5 M NaCl was added to a final concentration of 137 mM and stored on ice until use (cytoplasmic extracts). Pellets were resuspended in ice-cold RIP buffer and incubated with rotation for 30 min, followed by freezing and thawing and passage twice through a needle. Finally, samples were centrifuged at 14000 rpm for 20 min and the supernatants collected as nuclear extracts.

siRNAs and siRNA transfections

siRNA knock downs were carried out using DharmaFECT™ transfection reagent according to the manufacturer's instructions. Briefly, the siRNA was diluted to 40 nM final concentration in 250 ul serum free medium, and the mixture was added to 250 ul of serum free medium with 5 ul transfection reagent. The mixture was incubated at room temperature for 20 min prior to the addition of the mixture to a 60mm plate with subconfluent C33A2, HeLa or SiHa cells. siRNA to hnRNPG was ON-TARGET plus SMART pool Human RBMX (L-011691-01-0005 Dharmacon™). The scrambled control (scr) was siGENOME Control pool non-targeting #2 (D-001206-14-20, Dharmacon™).

RNA extraction and RT-PCR

Total RNA was extracted using TRI Reagent and Direct-zol RNA MiniPrep kit (ZYMO Research) according to the manufacturer's protocol. 1 μg of total RNA was reverse transcribed in a 20 μl reaction at 37°C using M-MLV Reverse Transcriptase (Invitrogen) and random primers (Thermo Scientific). One microliter of cDNA was subjected to PCR amplification. In some experiments, RT-PCR conditions were optimised to detect both HPV16 unspliced (E1) and spliced (E2) mRNAs by prolonging the extension time during PCR cycling. RT-PCR primers are listed in Supplementary Table S2. cDNA representing spliced actin or unspliced actin mRNAs were amplified with primers actin-s or actin-S1 and actin-a (Supplementary Table S2). To determine the identity of the various RT-PCR bands we either sequenced each gel-purified band directly using HPV16 primers, or after cloning of each gel-purified band into TOPO-plasmid. The inserts in the TOPO plasmids were determined using primers flanking the insert and sequencing was performed on multiple clones from each cloning reaction.

Real-time quantitative PCR (qPCR)

qPCR was performed on 1μl of cDNA prepared as described above in a MiniOpticon (Bio-Rad) using the SsoAdvanced SYBR Green Supermix (Bio-Rad) according to the manufacturer's instructions. For quantitation of E2 mRNA by RT-qPCR, primers 773S and 2715E2as were used (Supplementary Table S2) and results were normalized to GAPDH mRNA levels determined with primers gapdhf and gapdhr (Supplementary Table S2).

RNA-mediated protein pull-down assay

Nuclear extracts were prepared according to the procedure described previously (28). Briefly, the cells were lysed by using lysis buffer A (10 mM HEPES,10 mM KCl, 0.1 mM EGTA, 1 mM DTT, 0.4% NP40 pH 7.9 and protease inhibitors) for cytoplasmic proteins, then using buffer B (10 mM HEPES, 400 mM NaCl, 1 mM EGTA, 1 mM EDTA, 1 mM DTT pH7,9 and protease inhibitor) for nuclear proteins. The nuclear extracts were mixed with streptavidin-coated magnetic beads (Dynabeads M-280 Streptavidin, Invitrogen) carrying biotin-labelled single stranded RNA oligonucleotides in binding buffer (10 mM HEPES, 130 mM NaCl, 2.5 mM $MgCl_2$, 1 mM EGTA, 1 mM EDTA, 1 mM DTT, 10% glycerol pH 7.9 and protease inhibitor). RNA oligos are listed in Supplementary Table S3. The mixtures were incubated at room temperature with rotation for 1 h, followed by washing 10 times with 1 ml wash buffer (10 mM HEPES, 200 mM NaCl, 2.5 mM $MgCl_2$, 0.5% Triton X-100, 1 mM DTT pH 7.9). Proteins were eluted by boiling of the beads in SDS-PAGE loading buffer and subjected to SDS-PAGE followed by western blot analysis with indicated antibodies.

Preparation of proteins for mass spectrometry

Protein extracts were subjected to SDS-polyacrylamide gel electrophoresis followed by either Western blotting that was performed as described previously, or by staining with Silver statin (SilverQuestTM staining Kit, Invitrogen). Bands were excised and subjected to liquid chromatography–mass spectrometry (LC–MS) analysis at the SCIBLU Proteomics Resource Centre at Lund university. Primary and secondary antibodies used for western blotting are listed in Supplementary Table S4. Filters were stained with the Clarity Western ECL Substrate (BioRad) or the Super Signal West Femto chemiluminescence substrate (Pierce).

Immunoprecipitation and western blotting

Proteins for western blotting were extracted from cells using the radioimmunoprecipitation assay (RIPA) buffer (50 mM Tris pH 7.6, 150 mM NaCl, 0.5% Na-DOC, 1% NP-40, 1 mM DTT) with 30 min incubation on ice and occasional vortexing followed by centrifugation to remove cell debris. Immunoprecipitation was performed by over-night incubation of 400 ul of cell extract with Dynabeads and 1–2 ug of antibody followed by washing three times in RIPA buffer as described previously. Protein extracts were subjected to SDS-polyacrylamide gel electrophoresis and Western blotting that was performed as described previously. Primary and secondary antibodies used for Western blotting are listed in Supplementary Table S4. Filters were stained with the Clarity Western ECL Substrate (Bio Rad) or the Super Signal West Femto chemiluminescence substrate (Pierce).

RNA-protein immunoprecipitation assay (RIP)

Cells were lysed by resuspension in 1ml RIP lysis buffer (200 mM Tris, pH 8, 137mM NaCl, 1% NP-40, 2 mM EDTA, protease inhibitor and RNase inhibitor) followed by incubation on a rotator at 4°C for 30 min. Cell debris was removed by centrifugation. For immunoprecipitations, 1 ug of the indicated antibody or normal mouse IgG was incubated overnight at 4°C in 0.5 ml of lysate. Antibodies are listed in Supplementary Table S4. 50 ul (1.5 mg) of Dynabeads Protein A (#10002D, Invitrogen) were washed three times with lysis buffer and added to the cell extracts with antibody. The beads were washed six times with wash buffer (200mM Tris, pH 8, 137 mM NaCl, 1% NP-40, 2 mM EDTA) and RNA was extracted from the immunoprecipitations using Trizol-chloroform, dissolved in water and analyzed by RT-PCR directly. RT-PCR primers are listed in Supplementary Table S2. The RNA was ethanol-precipitated and dissolved in 20 ul of water. 10 ul of immunoprecipitated RNA was reverse transcribed using M-MLV reverse transcriptase (Invitrogen) and random hexamer primers (Thermo Scientific) according to the protocol of the manufacturer. 1 ul of cDNA were subjected to PCR amplification using primers indicated in each figure.

Application of small molecule substances to cells

C33A2 or HN26 cells were grown in 10-cm dishes. Cell culture medium was replaced with medium containing indicated concentrations of inhibitors for 21hrs or indicated time points. All inhibitors were dissolved in dimethyl sulfoxide (DMSO) (Sigma) and DMSO in the absence of inhibitor was used as a control in all experiments. Inhibitors were: Melphalan Hydrochloride (#Y0001158, European pharmacopoeia reference standard), 2-D08/S8696 (Selleckchem) (sumoylation-inhibitor) and TAK-981/S8829 (Selleckchem) (sumoylation-inhibitor).

sLuc assay

The *Metridia longa* secreted luciferase activity in the cultured medium of the C33A2 cells was monitored with the help of the Ready To Glow secreted luciferase reporter assay (Clontech) according to the instructions of the manufacturer.

Cell viability assay

HeLa cells were seeded in 96-well plates and incubated at 37°C in CO_2 incubator for 24 h. Transfections were performed with a total amount of 100 ng plasmid DNA per well using 0.3 ul Fugene6 (Promega). The transfected cells were incubated for 24 h followed by addition of 10 ul of WST-1 cell proliferation reagent (Roche) to each well and incubation for 2 h at 37°C in 5% CO_2. The absorbance was monitored at 450 nm in a Tristar LB941.

Quantitations

The softeware used to determine band intensity in western blots and RT-PCR gels is 'Image Lab 6.0.1' and quantitations were performed with the software 'Prism GraphPad 8.4.0'.

RESULTS

RNA sequences downstream of HPV16 3′-splice site SA2709 control production of E2 mRNAs

The HPV16 E2 protein is primarily produced from an mRNA that is spliced from HPV16 5′-splice site SD880 to HPV16 3′-splice site SA2709 (Figure 1A) (28). Figure 1A shows previously identified, HPV16 alternatively spliced mRNAs utilizing SA2709 (25). mRNAs spliced to the upstream 3′-splice site SA2582 can also produce E2 protein, but to a lesser extent than mRNAs spliced to SA2709 (28). To investigate how production of the E2 mRNA that is spliced from HPV16 SD880 to SA2709 is regulated, we used an HPV16 subgenomic reporter plasmid named pBEL (Figure 1B). Plasmid pBEL has the potential to produce E1 mRNAs that are unspliced, E2 mRNAs that are spliced from SD880 to either SA2582 or SA2709 and E1∧E4 mRNAs spliced from SD880 to SA3358 (Figure 1C). These mRNAs may be monitored by RT-PCR using primers indicated in Figure 1C. To determine if the major E2 splice site SA2709 was controlled by downstream RNA elements, we deleted 190 nucleotides (nt) between HPV16 positions 2731 and 2922 downstream of SA2709 in pBEL, resulting in plasmid pD200 (Figure 1B, upper panel). HeLa cells were transfected with pBEL and pD200 and total RNA was extracted and analyzed by RT-PCR. The RT-PCR reaction was optimized to detect both unspliced and spliced mRNAs in the same reaction despite their size difference. The results revealed that the 190nt deletion in plasmid pD200 (nt2731-nt2922) reduced splicing to SA2709 (Figure 1D). In contrast, the levels of E2 mRNAs spliced to SA2582 and the levels of unspliced E1-mRNAs increased (Figure 1D). PCR in the absence of reverse transcription did not yield amplified DNA fragments (Figure 1D). mRNAs spliced from SD880 to SA3358 were unaffected by the deletion downstream of SA2709 (Figure 1D). RT-PCR analysis of serially diluted cDNA from the transfection suggested that the 190nt deletion reduced E2 mRNA levels >128-fold, as the intensity of the RT-PCR band from the 128-fold dilution of cDNA from the pBEL transfection was similar to the intensity of the RT-PCR band from undiluted cDNA from the pD200 transfection (Figure 1E). This was in line with the RT-qPCR analysis of spliced E2 mRNAs that revealed that pD200 produced 166-fold less E2 mRNA than pBEL (Figure 1F). Note that different PCR-conditions were used for RT-PCR and RT-qPCR. To determine if the 190nt sequence between HPV16 nt positions 2731 and 2922 directed splicing to SA2709 in a sequence-dependent manner, we turned this 190nt element upside down (plasmid pBELAS) (Figure 1B, upper panel). For the exact sequence of the insert, see Supplementary Figure S1A. pBELAS produced lower E2 mRNA levels than pBEL (Figure 1G), indicating that splicing enhancement was dependent on the primary sequence between positions 2731 and 2922. In contrast, reintroduction of the wild type sequence into pD200, resulting in plasmid pBELXC, efficiently restored E2 RNA levels to those produced by pBEL (compare E2 mRNAs levels from pBELXC and pBEL in Figure 2C). Since the majority of the 190nt RNA sequence that enhanced splicing to SA2709 overlapped the E2 protein coding sequence, we also codon modified the E2 coding region up to position 2931, resulting

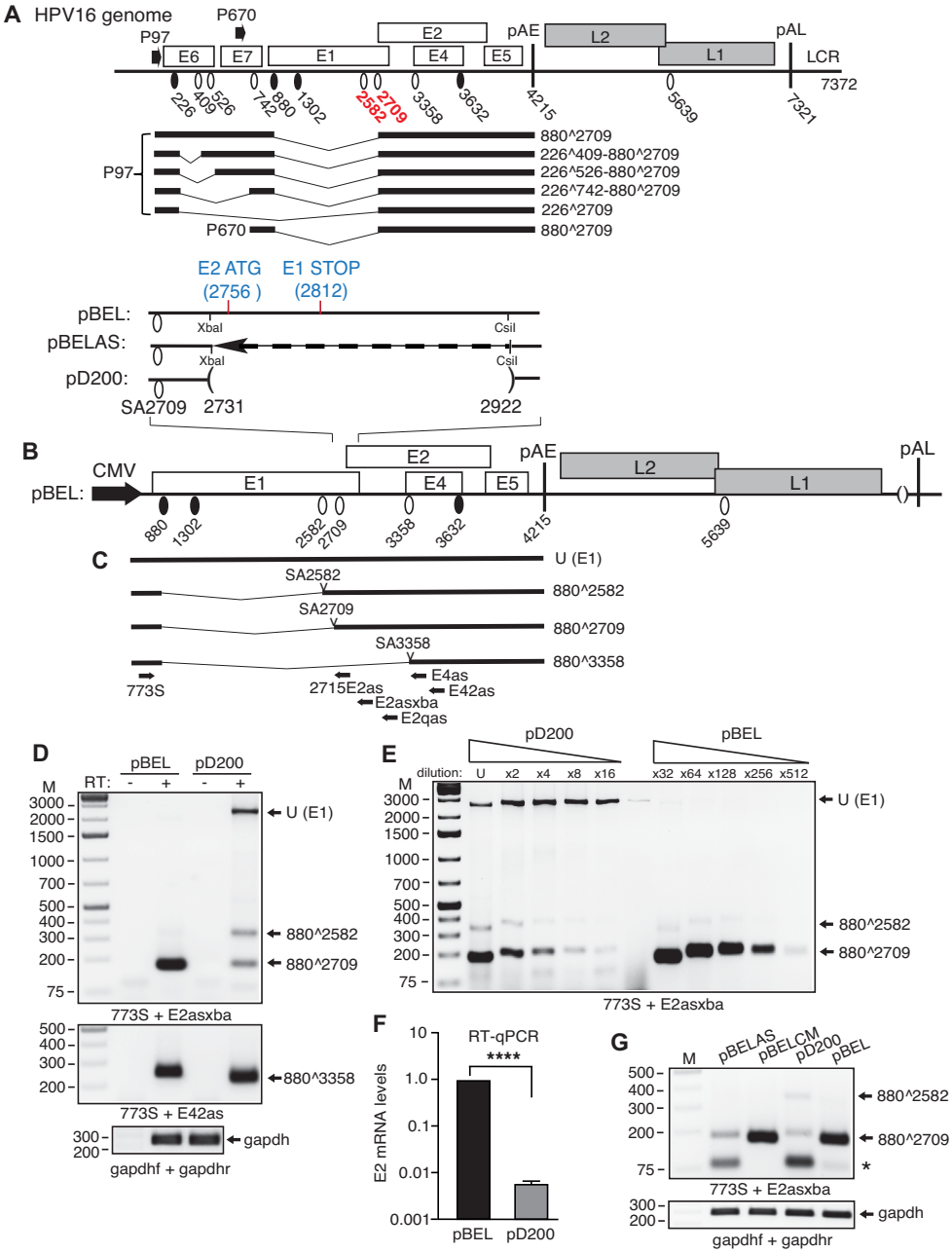


Figure 1. (**A**) Linearized HPV16 genome (numbers refer to the HPV16 reference strain GeneBank: K02718.1). Early and late genes are indicated. P97: HPV16 early promoter. P670: HPV16 late promoter. Black oval: 5′-splice site/splice donor. White oval: 3′-splice site/splice acceptor. Splice acceptors SA2582 and SA2709 located upstream of the E2 open reading frame are indicated in red. pAE: HPV16 early polyadenylation site. pAL: HPV16 late polyadenylation site. LCR: long control region. A subset of alternatively spliced HPV16 E2 mRNAs all utilizing HPV16 3′-splice site SA2709. (**B**) Middle panel**:** Schematic representation of the HPV16 subgenomic reporter plasmid pBEL. Transcription of the HPV16 sequences in the pBEL plasmid is driven by the human cytomegalovirus immediate early promoter (CMV). Early and late genes are indicated. P97: HPV16 early promoter. P670: HPV16 late promoter. Black oval: 5′-splice site/splice donor. White oval: 3′-splice site/splice acceptor. pAE: HPV16 early polyadenylation site. pAL: HPV16 late polyadenylation site. Numbers refer to the HPV16 reference strain (GeneBank: K02718.1). Upper panel: Blow up of the region downstream of HPV16 3′splice site SA2709. The deletion introduced between positions 2731 and 2922 in pBEL to produce pD200 is shown, the introduction of a unique XbaI restriction site at position 2731 is indicated and the insertion of antisense sequences between 2731 and 2922 is indicated with an arrow (pBELAS). (**C**) The major HPV16 mRNAs produced from pBEL are spliced from HPV16 5′-splice site SD880 to either 3′-splice site SA2582, SA2709 or SA3358 and polyadenylated at pAE. (**D**) RT-PCR with indicated primers on RNA extracted from HeLa cells transfected with pBEL or pD200. RT-PCR was performed in the absence (−) or presence (+) of RT as indicated. Unspliced E1 encoding mRNAs U (E1) mRNAs and mRNAs spliced between SD880 and SA2582, SA2709 or SA3358 are indicated to the right. Note that the two gapdh samples represent controls for pBEL and pD200 performed in the presence of RT. M, molecular weight marker. (**E**) PCR on 2-fold serially diluted cDNA obtained from RNA extracted from HeLa cells transfected with pD200 or pBEL. For pD200, cDNA was undiluted, 2-, 4-, 8- and 16-fold diluted. For pBEL, cDNA was 32-, 64-, 128- and 256- and 512-fold diluted. (**F**) RT-qPCR with primers 773S and 2715E2as was performed as described in Materials and Methods to quantitate relative E2 mRNA levels produced by pBEL and pD200 upon transfection of HeLa cells. $P < 0.0001$. (**G**) RT-PCR with the indicated primers on RNA extracted from HeLa cells transfected with pBELAS, pBELCM, pD200 or pBEL.

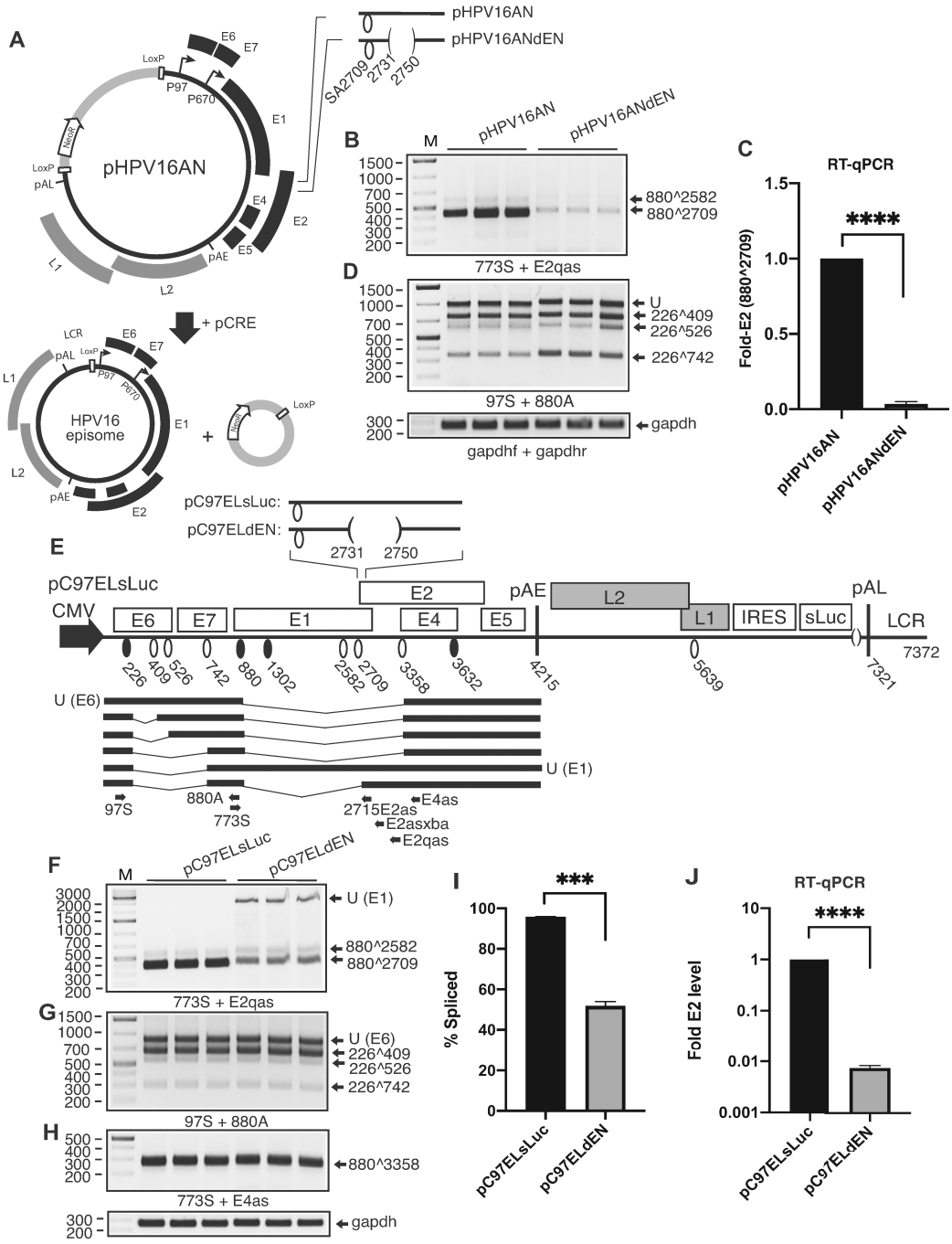


Figure 2. (**A**) Schematic representation of genomic HPV16 plasmid pHPV16AN. LoxP sites and HPV16 early (P97) and late (P670) promoters and early (pAE) and late (pAL) poly(A) signals are indicated. Positions of PCR primers 16S and 16A are indicated. The effect of the cre recombinase on pHPV16AN in transfected cells is illustrated. (**B**) RT-PCR with indicated primers on RNA extracted from HeLa cells transfected with pHPV16AN or pHPV16ANdEN in the presence of cre-expressing plasmid pCAGGS-nlscre. Transfections were performed in triplicates. M, molecular weight marker. (**C**) RT-qPCR with primers 773S and 2715E2as was performed as described in Materials and Methods to quantitate relative E2 mRNA levels produced by pHPV16AN or pHPV16ANdEN upon transfection of HeLa cells. $P < 0.0001$. (**D**) RT-PCR with indicated primers on RNA extracted from HeLa cells transfected with pHPV16AN or pHPV16ANdEN. Transfections were performed in triplicates. (**E**) Schematic representation of the HPV16 subgenomic pC97ELsLuc reporter plasmid. Transcription of the HPV16 sequences in the pC97ELsLuc plasmid is driven by the human cytomegalovirus immediate early promoter (CMV). HPV16 splice sites are indicated. Numbers refer to the HPV16 reference strain (GeneBank: K02718.1). Early and late polyadenylation signals pAE and pAL are indicated. A subset of HPV16 alternatively spliced, early mRNAs are indicated. Arrows represent RT-PCR primers. IRES, internal ribosome entry site; sLuc, secreted luciferase. The deletion of the splicing enhancer downstream of splice site SA2709 is indicated at the top. (**F–H**) RT-PCR on RNA extracted from HeLa cells transfected with the indicated plasmids. RT-PCR primers are indicated below each gel image and spliced mRNAs represented by the amplicons are indicated to the right. (**I**) Densitometric quantification of electrophoresed RT-PCR products were performed as described in Materials and Methods. Percentage spliced E2 mRNA over unspliced/intron retained E1 mRNAs is shown. (**J**) RT-qPCR with primers 773S and 2715E2as was performed as described in Materials and Methods to quantitate relative E2 mRNA levels produced by pC97ELsLuc and pC97ELdEN upon transfection of HeLa cells. $P < 0.0001$.

in pBELCM. For the exact sequence of the insert, see Supplementary Figure S1B. However, these mutations did not affect the splicing-enhancing activity of the 190nt sequence (Figure 1G), suggesting that these mutations failed to inactivate the splicing enhancer element or that the splicing enhancer was located upstream of the E2 ATG. Combined, our results indicated that a splicing enhancer sequence was located downstream of HPV16 nucleotide position 2731, downstream of E2 mRNA splice site SA2709.

An RNA sequence located between HPV16 E2 mRNA splice site SA2709 and the E2 AUG is required for efficient splicing to SA2709

To map the borders of the putative splicing enhancer located between HPV16 nucleotide positions 2731 and 2922, we introduced deletions downstream of SA2709 in HPV16 plasmid pBEL (Supplementary Figure S2A and B). The various deletion mutants were transfected into HeLa cells, RNA was extracted and RT-PCR performed with the primers 773S and E2A displayed in Figure 1C. The RT-PCR results demonstrated that all deletions starting immediately downstream of position 2731 reduced splicing to SA2709 (Supplementary Figure S2C and D). In addition, these deletions induced production of E2 mRNAs spliced to SA2582 and of unspliced E1 mRNAs (Supplementary Figure S2C and D). In contrast, deletions introduced further down, as in plasmids p15 and p16 (Supplementary Figure S2B), did not reduce splicing to SA2709 (Supplementary Figure S2E). We also confirmed that splicing to HPV16 3′-splice site SA3358 was unaffected by the various deletions in these plasmids (Supplementary Figure S2E). Taken together, these results demonstrated that sequences enhancing splicing to SA2709 were located upstream of nucleotide position 2830. The splicing enhancer would therefore be located between HPV16 nucleotide positions 2731 and 2830. To map the location of the enhancer element further, we constructed plasmid p2758 that was extended to nucleotide position 2758 and carried a deletion between 2758 and 2922 (Supplementary Figure S2B). Splicing to SA2709 in p2758 was similar to the wild type pBEL plasmid (Supplementary Figure S2F), demonstrating that the splicing enhancer sequence was intact in plasmid p2758, which essentially mapped the splicing enhancer to a 28 nt sequence located between HPV16 nucleotide positions 2731 and 2758. These results were confirmed using smaller HPV16 expression plasmids named pBELEN and pBELENdEN (Supplementary Figure S3A and B). The only difference between pBELEN and pBELENdEN is the deletion of the 28-nucleotide putative splicing enhancer located between nucleotide position 2731 and 2758 in pBELENdEN (Supplementary Figure S3B). The major mRNAs produced by these plasmids are displayed in Supplementary Figure S3C. The majority of the mRNAs produced from pBELEN were spliced between SD880 and SA2709, as were the mRNAs produced from the parental plasmid pBEL and plasmid p2758 (Supplementary Figure S3D). In contrast, plasmid pBELENdEN that lacked the splicing enhancer between nucleotide positions 2731 and 2758, produced significantly lower levels of mRNAs spliced to SA2709, comparable to those produced by pD200 (Supplementary Figure S3D). mRNAs produced from the latter plasmids were either unspliced or generated by redirection of splicing to the competing 3′-splice site SA2582 (Supplementary Figure S3D). These results confirmed the presence of a splicing enhancer located between HPV16 nucleotide positions 2731 and 2758, downstream of the HPV16 E2 3′-splice site SA2709.

The splicing enhancer downstream of HPV16 3′-splice site SA2709 is required for production of HPV16 E2 mRNA from the episomal form of the HPV16 genome

Next we wished to determine if the RNA sequence between nucleotide positions 2731 and 2758 played a significant role in the production of HPV16 E2 mRNAs from the episomal form of the full-length HPV16 genome. We deleted the enhancer located between nucleotide positions 2731 and 2750 in the pHPV16AN plasmid (Figure 2A). This plasmid contains two loxP sites flanking the HPV16 genome from which the episomal genome can be released with the help of a cre-recombinase produced from the co-transfected pCRE expression plasmid (Figure 2A). As can be seen from the RT-PCR of the E2 mRNAs spliced from HPV16 5′-splice site SD880 to 3′-splice site SA2709, plasmid pHPV16ANdEN (in which the enhancer was deleted) produced lower levels of the E2 mRNAs than the wild type pHPV16AN plasmid (Figure 2B). RT-qPCR analysis of the E2 mRNA levels revealed a >30-fold reduction in E2 mRNA levels (Figure 2C). Splicing in the E6/E7-coding region was not substantially affected by the deletion downstream of splice site SA2709 (Figure 2D). Similar results were obtained using HPV16 reporter plasmid pC97ELsLuc that contains the HPV16 genome inserted downstream of the strong CMV promoter (Figure 2E). Deletion of the enhancer downstream of SA2709 significantly reduced the E2 mRNA levels spliced to SA2709 (Figure 2F) and altered the percentage spliced E2 mRNA to E1 mRNAs with retained intron (Figure 2I), but did not affect splicing within the E6/E7 coding region (Figure 2G), nor to the downstream splice site SA3358 (Figure 2H). RT-qPCR revealed a 125-fold reduction in E2 mRNA levels spliced to SA2709 as a consequence of the deletion (Figure 2J). Thus, our results demonstrated that the HPV16 sequence between positions 2731 and 2750 of the HPV16 genome was required for efficient production of HPV16 E2 mRNAs also in the context of the entire HPV16 genome.

Inactivation of the HPV16 E2 mRNA splicing enhancer by nucleotide substitutions

Analysis of the sequence between nucleotide positions 2731 and 2758 revealed that it harbored three AC(G/A)AGG repeats (totally 18 nucleotides between positions 2732 and 2749), or alternatively three AC(G/A)AGGA repeats between positions 2732 and 2750 (totally 19 nucleotides) (Supplementary Figure S4A, B and C). To investigate if the three repeats constitute the splicing enhancer, we introduced point mutations at all 19 positions in plasmid p2758, resulting in plasmid pMALL (Supplementary Figure S4C). In general, the nucleotides were mutated to the complementary nucleotide when possible, at the same time avoiding to create long uninterrupted polypyrimidine tracts as they

could potentially interact with known RNA-binding proteins such as U2AF65 and hnRNP I. Analysis of RNA extracted from HeLa cells transfected with plasmid pMALL or p2758 revealed that the substitutions reduced splicing to SA2709 and resulted in an increase in unspliced E1 mRNAs (E1) (Supplementary Figure S4D). As expected, splicing to the downstream HPV16 splice site SA3358 was unaffected by these mutations (Supplementary Figure S4D). These results demonstrated that the HPV16 splicing enhancer sequence encoding the three AC(G/A)AGG-repeats directed splicing to SA2709 in a sequence-specific manner. The RNA sequence encoding the three repeats was a U-less and purine-rich sequence (84% adenosine and guanosine) with three cytosines (16% cytosine) (Supplementary Figure S4C). To analyze the enhancer sequence further we mutated one or two of the three AC(G/A)AGG-repeats at a time (Supplementary Figure S4C). Mutations in motifs #1 and #3 reduced splicing to SA2709, whereas mutations in motif #2 appeared to have only a minor effect on the splicing enhancer (Supplementary Figure S4E). Mutations of two or three repeats simultaneously also reduced splicing to SA2709 (Supplementary Supplementary S4E). Analysis of additional duplicates followed by quantitation revealed that nucleotide substitutions in all three motifs or motifs 1 and 3 significantly reduced E2 mRNA splicing ($P < 0.001$) (Supplementary Figure S4F and H). Analysis of longer exposures also revealed that the splicing-maimed mutants pMALL, pM13 and pM3 produced significantly higher levels of E1 mRNAs with retained intron than wild type plasmids, although for the E1 mRNAs $P < 0.01$ (Supplementary Figure S4G and I). Combined, these results demonstrated that all three AC(G/A)AGG repeats contributed the enhancer activity of the E2 splicing enhancer and indicated that repeats #1 and #3 played a more significant role in this enhancement.

Identification of proteins that interact with the splicing enhancer downstream of E2 mRNA splice site SA2709

We next used biotinylated RNA oligos representing the splicing enhancer sequence from position 2732 to 2758, or a single AC(G/A)AGG-motif or mutants of both (Figure 3A) to identify cellular protein partners of the splicing enhancer. The biotinylated RNA oligos were attached to streptavidin coated magnetic beads and used for pull down of proteins from nuclear and cytoplasmic extracts prepared from HeLa cells. Proteins interacting with the RNA oligos were eluted and separated on polyacrylamide gels that were silver stained. Multiple bands of various molecular weight were specifically pulled down with RNAs representing wild type enhancer RNA sequences, compared with mutant RNAs (Figure 3B). These bands were cut out from the gels and sent for mass spectrometric analysis. For each band, multiple proteins were identified. RNA binding proteins represented by various peptides and multiple scores for each peptide were selected for confirmatory Western blots. As can be seen, some proteins were efficiently pulled down in an RNA sequence-specific manner (hnRNP A1, hnRNP A2, hnRNP G, hnRNP M, RBM14 and RBM15) (Figure 3C). In contrast, hnRNP L and hnRNP U that were not specifically identified by mass spectrometry, were not pulled down by these RNA oligonucleotides (Figure 3C), supporting the specificity of the pull-down reactions. Proteins that represented general splicing factors were pulled down less efficiently (SF3b, U2AF65) (Figure 3C), but with specificity. Possibly because they may be pulled down on the basis of protein-protein interactions rather than direct binding to RNA since splice site SA2709 was absent from the RNA oligonucleotides. We concluded that multiple RNA binding proteins were interacting specifically with the HPV16 wild type RNA enhancer sequence.

hnRNP G enhances splicing to HPV16 E2 mRNA splice site SA2709

To determine the effect of the identified RNA binding proteins on HPV16 E2 mRNA splicing, we co-transfected expression plasmids for hnRNP A1, hnRNP A2, hnRNP G, hnRNP M, RBM14 or RBM15 with HPV16 reporter plasmid pBELENdE1 (see Supplementary Figure S3B for a schematic representation of this HPV16 reporter plasmid). This plasmid has a deletion in the E1 coding region that reduces the size of the unspliced E1 mRNAs thereby facilitating simultaneous detection of the unspliced (dE1) mRNA and the spliced E2 mRNAs in the same RT-PCR reaction. Locations of RT-PCR primers in HPV16 reporter plasmid pBELENdE1 are shown in Supplementary Figure S3C. As can be seen from the results, overexpression of hnRNP G enhanced production of the E2 mRNAs (Figure 3D and E), whereas hRNP A1 and A2 strongly inhibited splicing (Figure 3D and E), while hnRNP M, RBM14 and RBM15 appeared to display a weak inhibitory effect on HPV16 E2 mRNA splicing (Figure 3D and E). We speculated that hnRNP G acted through the E2 splicing enhancer to promote splicing to HPV16 3′-splice site SA2709. To provide further support for this idea, we used subgenomic HPV16 E1/E2 reporter plasmids pBELEN and pBELENdE1 that both contained an intact splicing enhancer at SA2709 (Supplementary Figure S3B) paired with pBELENdEN and pBELENdE1dEN that both lacked the same enhancer (Supplementary Figure S3B). Co-transfection of the four plasmids individually with hnRNP G expression plasmid revealed that hnRNP G enhanced splicing to SA2709 from plasmids pBELEN and pBELENdE1 that contained an intact splicing enhancer downstream of SA2709 (Figure 3F). This was particularly obvious with plasmid pBELENdE1 from which unspliced mRNA was readily detected by RT-PCR (Figure 3F). In contrast, hnRNP G did not enhance splicing to SA2709 and production of E2 mRNAs from plasmids pBELENdEN and pBELENdE1dEN that both contained a 19-nucleotide deletion encompassing the splicing enhancer downstream of SA2709 (Figure 3F). To determine if hnRNP G also affected the levels of E1 or E2 proteins, we transfected plasmid pBEE2 that encodes the wild type E2 enhancer or plasmid pBEE2-DEN that lacks the enhancer (Supplementary Figure S5A and B) into HeLa cells in the absence or presence of hnRNP G plasmid. These plasmids produced either spliced E2 encoding mRNAs or intron retained E1 encoding mRNAs (Supplementary Figure S5C). pBEE2 produced E2 protein as expected whereas pBEE2-DEN produced undetectable levels of E2 protein (Supplementary Figure S5D). The levels of E2 protein pro-

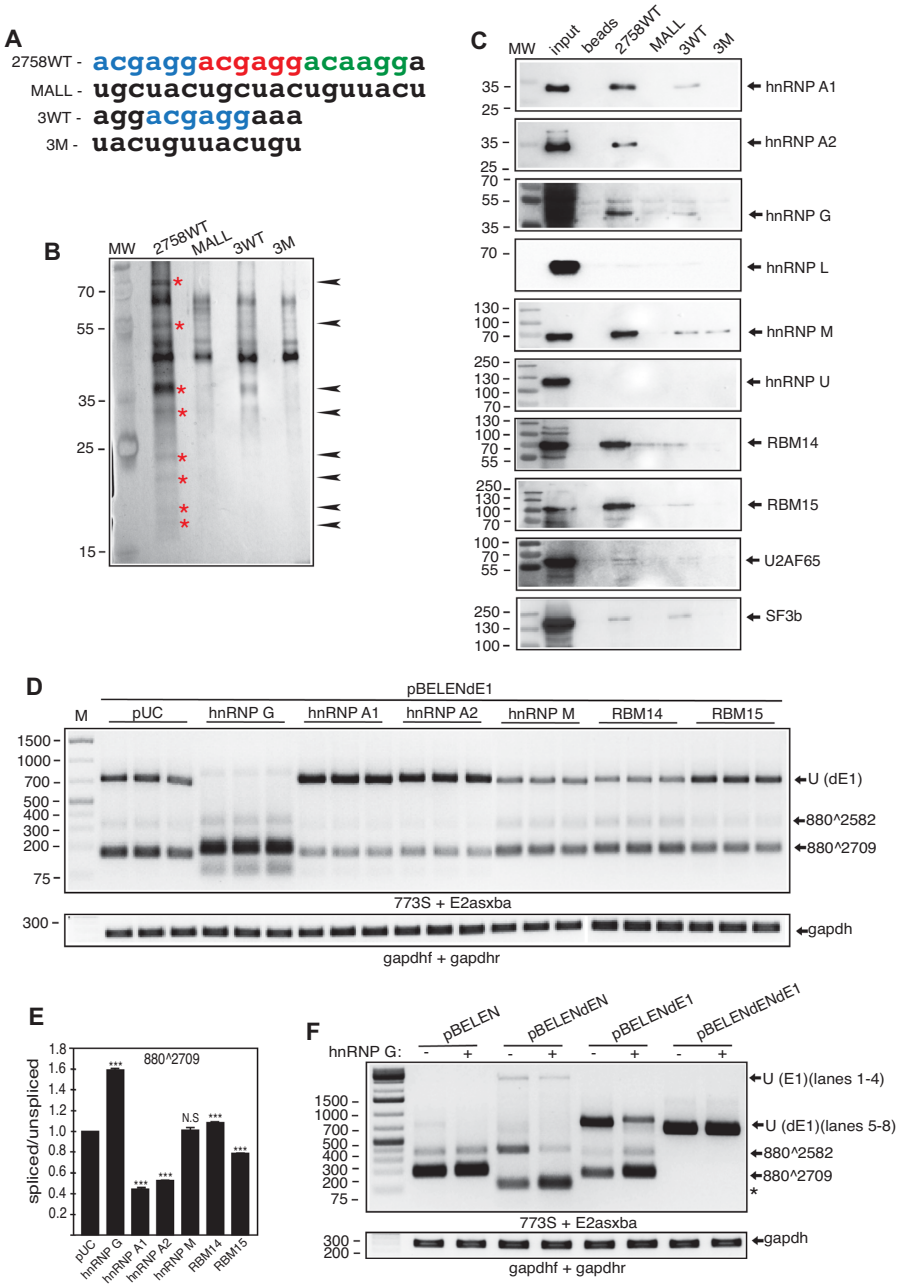


Figure 3. (**A**) Sequences of biotinylated HPV16 RNA oligonucleotides used in RNA-mediated protein pull-downs of proteins from HeLa cell nuclear extracts. The three AC(G/A)AGG - repeats that constitute the enhancer are colored in the wild type RNA sequence (2758WT). MALL, all nucleotide positions in the wild type enhancer are mutated; 3WT, contains one AC(G/A)AGG – repeat; 3M, contains one mutant enhancer motif. Names of RNA oligonucleotides are listed to the left. (**B**) Silver stained SDS-polyacrylamide gel displaying proteins pulled down by the biotinylated RNA oligos shown in (A). Bands indicated with red stars and arrows were cut out from the gel and subjected to mass spectrometry. MW, molecular weight marker. (**C**) Western blotting of factors pulled down by Streptavidin-coated magnetic beads carrying biotinylated HPV16 RNA oligo 2758WT, MALL, 3WT or 3M. Filters were stained with antibodies to proteins identified by mass spectrometry (indicated to the right) and by antibodies to hnRNP L and hnRNP U that were not identified by mass spectrometry and served as negative controls. (**D**) RT-PCR with primers 773S and E2asxba on RNA extracted from HeLa cells transfected with HPV16 subgenomic plasmid pBELENdE1 in the presence of empty pUC plasmid or plasmids expressing either of the indicated proteins. Unspliced U (dE1) mRNA and mRNAs spliced between SD880 and SA2582 or SA2709 are indicated to the right of the gel image. Transfections were performed in triplicates. Locations of RT-PCR primers in HPV16 reporter plasmid pBELENdE1 are shown in Supplementary Figure S3C. M, molecular weight marker. (**E**) Densitometric quantification of RT-PCR bands in (A). The quantitations are displayed as ratios between cDNA bands representing HPV16 mRNAs spliced between SD880 and SA2709 and cDNA bands representing unspliced HPV16 mRNAs (U (dE1)) in each transfected sample. Splicing ratio monitored with pBELENdE1 cotransfected with pUC was set as 1. Quantitations were performed on triplicate samples. Statistically significant differences between 880∧2709 cDNA levels in samples transfected with pBELENdE1 and plasmids expressing either of the indicated proteins and samples transfected with pBELENdE1 and empty pUC plasmid are shown. ***, $P < 0.001$; N.S., not significantly different. (**F**) RT-PCR with primers 773S and E2asxba on RNA extracted from HeLa cells transfected with HPV16 subgenomic plasmids pBELEN, pBELENdEN, pBELENdE1 or pBELENdENdE1 in the absence (−) or presence (+) of hnRNP G plasmid. Unspliced U (dE1) mRNA and mRNAs spliced between SD880 and SA2582 or SA2709 are indicated to the right of the gel image. *, position of primer-dimer products.

duced from the pBEE2 plasmid encoding wild type E2 enhancer increased in the presence of hnRNP G, whereas pBEE2-DEN produced undetectable levels of E2 protein also in the presence of hnRNP G (Supplementary Figure S5D). Since HPV16 E2 has been shown to induce apoptosis, our results predicted that plasmid pBEE2 would decrease cell viability, whereas pBEE2-END should not. Transfection of HeLa cells with pBEE2 decreased HeLa cell viability, whereas pBEE2-END did not (Supplementary Figure S5E). Co-transfection of pBEE2 with hnRNP G decreased HeLa cell viability further, as expected (Supplementary Figure S5E). In contrast, overexpression of hnRNP G reduced the levels of E1 protein produced from the E1 expression plasmid p16E1-3XF (Supplementary Figure S5F–H), as expected, since E1 is produced from the unspliced, intron retained mRNA and hnRNP G enhances splicing to E2 splice site SA2709. We concluded that hnRNP G acted on the splicing enhancer downstream of SA2709 to enhance splicing to the HPV16 E2 mRNA 3′-splice site SA2709, thereby enhancing production of HPV16 E2 protein at the expense of HPV16 E1 protein.

hnRNP G exerts a splicing inhibitory effect on HPV16 mRNAs when the HPV16 E6 and E7 coding sequences are present on the pre-mRNAs

To investigate the role of the cellular RNA-binding protein hnRNP G in the control of HPV16 early gene expression further, we transfected HeLa cells with pHPV16AN, that encodes the full-length HPV16 genome, in the absence or presence of hnRNP G plasmid. To create episomal HPV16 DNA, the pHPV16AN plasmids that contain loxP sites flanking the HPV16 genome was co-transfected with cre-expressing plasmid, as described previously (Figure 2A). Surprisingly, overexpression of hnRNP G did not enhance HPV16 E2 mRNA production from pHPV16AN (Figure 4A). Furthermore, RT-PCR analysis revealed that hnRNP G efficiently inhibited splicing of the E6/E7 mRNAs produced from the episomal HPV16 genome, favoring unspliced E6 mRNAs with retained intron over the spliced E7 mRNAs (Figure 4B and C). To confirm these results, we also determined the effect of hnRNP G on HPV16 subgenomic reporter plasmid pC97ELsLuc in which the CMV promoter was inserted upstream of E6/E7 in place of the early HPV16 promoter (Figure 4D). In line with the results obtained with the full-genome plasmid pHPV16AN, overexpression of hnRNP G did not enhance HPV16 E2 mRNA production from pC97ELsLuc (Figure 4E). hnRNP G inhibited E6/E7 mRNA splicing also from plasmid pC97ELsLuc (Figure 4F and G). Co-transfection of pC97ELsLuc with plasmids expressing hnRNP A1 and hnRNP A2, re-produced previously published results, i.e. inhibition of splicing between SD226 and SA409 followed by production of primarily mRNAs spliced to SA742 (hnRNP A2) or unspliced, E6-encoding mRNAs (hnRNP A1) (Figure 4F and G) (33). To provide support for the functional significance of the inhibitory effect of hnRNP G on production of HPV16 spliced E7 mRNA, we also monitored the levels of pRB in cells transfected with pC97ELsLuc in the absence or presence of hnRNP G. As expected, pRB levels were reduced in cells transfected with pC97ELsLuc compared with cells transfected with empty pUC plasmid as a result of E7 production from pC97ELsLuc (Supplementary Figure S6A and B). The levels of pRB were restored in cells transfected by pC97ELsLuc and hnRNP G since hnRNP G inhibited production of spliced E7 mRNAs from pC97ELsLuc (Supplementary Figure S6A and B). Finally, the inhibitory effect of hnRNP G on E6/E7 mRNA splicing was dependent on the amount of transfected hnRNP G plasmid (Supplementary Figure S6C and D) and could be overcome by co-expression of hnRNP A2 that stimulated splicing to SA742 (Supplementary Figure S6E and F), suggesting that the splicing inhibitory effect is subordinated when hnRNP A2 enhances splicing. Combined, our results demonstrated that hnRNP G inhibited production of spliced HPV16 E7 mRNAs.

Since overexpression of hnRNP G appeared not to enhance E2 mRNA production from pC97ELsLuc, we knocked down hnRNP G by siRNA in cells transfected with pC97ELsLuc – then analyzed E2 mRNA levels. Transfection of siRNAs to hnRNP G resulted in reduced hnRNP G protein levels (Figure 4H) and reduced splicing to HPV16 E2 splice site SA2709 and caused accumulation of unspliced E1 mRNAs (Figure 4I). Knock down of hnRNP G also resulted in enhanced splicing of the E6/E7 mRNAs as shown by an increase in levels of the major E7 mRNA spliced between SD226 and SA409 at the expense of the intron-containing E6 mRNA (Figure 4J). Knock down of hnRNP G in the C33A2 cell line that contains integrated copies of the HPV16 reporter plasmid pBELsLuc (Supplementary Figure S7A, B and C) that lacks the E6/E7 coding region resulted in reduced splicing to SA2709 and enhanced splicing to the alternative 3′-splice site SA2582 as well as in accumulation of unspliced E1 mRNAs (Supplementary Figure S7D–F). Locations of RT-PCR primers in HPV16 reporter plasmid pBELsLuc are shown in Supplementary Figure S7C. Knock down of hnRNP G in the HPV16-positive SiHa cervical cancer cell line caused an increase in levels of the spliced E7 mRNA (226∧409) at the expense of the intron-containing E6 mRNA (Supplementary Figure S7G and H). Taken together, our results demonstrated that hnRNP G is required for splicing to HPV16 SA2709 and for production of E2 mRNAs, but when the E6- and E7-coding regions were present on the HPV16 pre-mRNAs, overexpression of hnRNP G exerted a splicing inhibitory effect on the HPV16 E6 and E7 mRNAs. These results suggested that hnRNP G interacted with sequences in the E6 and E7 coding regions in addition to the splicing enhancer downstream of HPV16 E2 splice site SA2709. To reconcile these results, we speculated that hnRNP G may have different effects on HPV16 mRNAs produced from the HPV16 early promoter P97 and containing E6/E7 sequences compared to mRNAs produced from the late promoter P670 and lacking E6/E7 sequences (for the location of the two HPV16 promoters P97 and P670, se Figure 1A).

hnRNP G inhibits splicing of HPV16 E6 and E7 oncogene mRNAs independently of RNA sequences downstream of the E2 splice site SA2709

To investigate if hnRNP G acted directly on sequences in the HPV16 E6/E7 coding region (Figure 5A), we analyzed

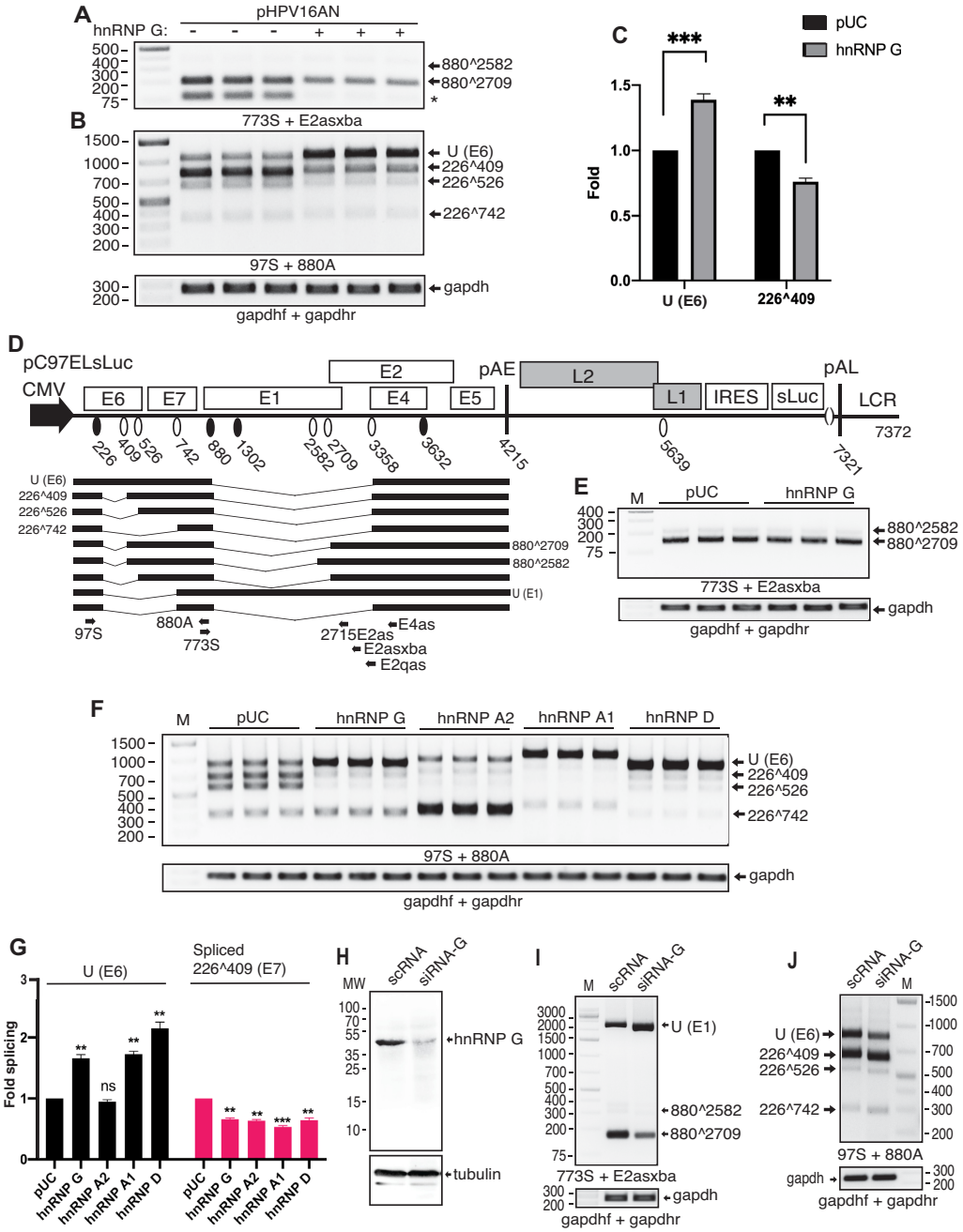


Figure 4. (**A**, **B**) RT-PCR with indicated primers on RNA extracted from HeLa cells transfected with pHPV16AN and cre-expressing plasmid in the absence (−) or presence (+) of phnRNP G. Triplicate transfections are shown. (A) mRNAs spliced between SD880 and SA2582 or SA2709 are indicated to the right of the gel image. *, position of primer-dimer products. (**B**) Unspliced/intron retained U (E6) mRNA and mRNAs spliced between SD880 and SA409, SA526 or SA742 are indicated to the right of the gel image. (**C**) Densitometric quantification of RT-PCR bands (U (E6) and 226∧409) in (**B**) in the absence or presence of hnRNP G. The The quantitations are displayed as relative levels of U (E6) cDNAs or 226∧409 in the absence or presence of hnRNP G. (**D**) Schematic representation of the HPV16 subgenomic pC97ELsLuc reporter plasmid. A subset of HPV16 alternatively spliced, early mRNAs are indicated. Arrows represent RT-PCR primers. IRES, internal ribosome entry site; sLuc, secreted luciferase. (**E**) RT-PCR with indicated primers on RNA extracted from HeLa cells co-transfected with pC97ELsLuc and empty pUC plasmid or hnRNP G-expression plasmid. mRNAs spliced between SD880 and SA2582 or SA2709 are indicated to the right of the gel image. (**F**) RT-PCR with primers 97S and 880A on RNA extracted from HeLa cells co-transfected with pC97ELsLuc and empty pUC plasmid or hnRNP-expressing plasmids hnRNP G, hnRNP A2, hnRNP A1 or hnRNP D. Unspliced U (E6) mRNA and mRNAs spliced between SD880 and SA409, SA526 or SA742 are indicated to the right of the gel image. (**G**) Densitometric quantification of RT-PCR bands (U (E6) and 226∧409) in (H) in the absence or presence of hnRNP G. The quantitations are displayed as relative levels of U (E6) cDNAs or 226∧409 cDNAs in the absence or presence of hnRNP G. p-values are indicated. (**H**) Western blotting of hnRNP G in HeLa cells transfected with scrambled siRNAs (scRNA) or siRNAs to hnRNP G (siRNA-G). (**I**, **J**) HeLa cells were transfected with HPV16 subgenomic expression plasmid pC97ELsLuc and scrambled siRNA (scRNA) or siRNAs to hnRNP G, (siRNA-G) and RNA was extracted. (I) RT-PCR was performed with primers 773S and E2asxba. Unspliced/intron retained U (E1) mRNAs and mRNAs spliced between SD880 and SA2582 or SA2709 are indicated to the right. (J) RT-PCR was performed with primers 97S and 880A. Unspliced U (E6) mRNA and mRNAs spliced between SD880 and SA409, SA526 or SA742 are indicated to the right of the gel image.

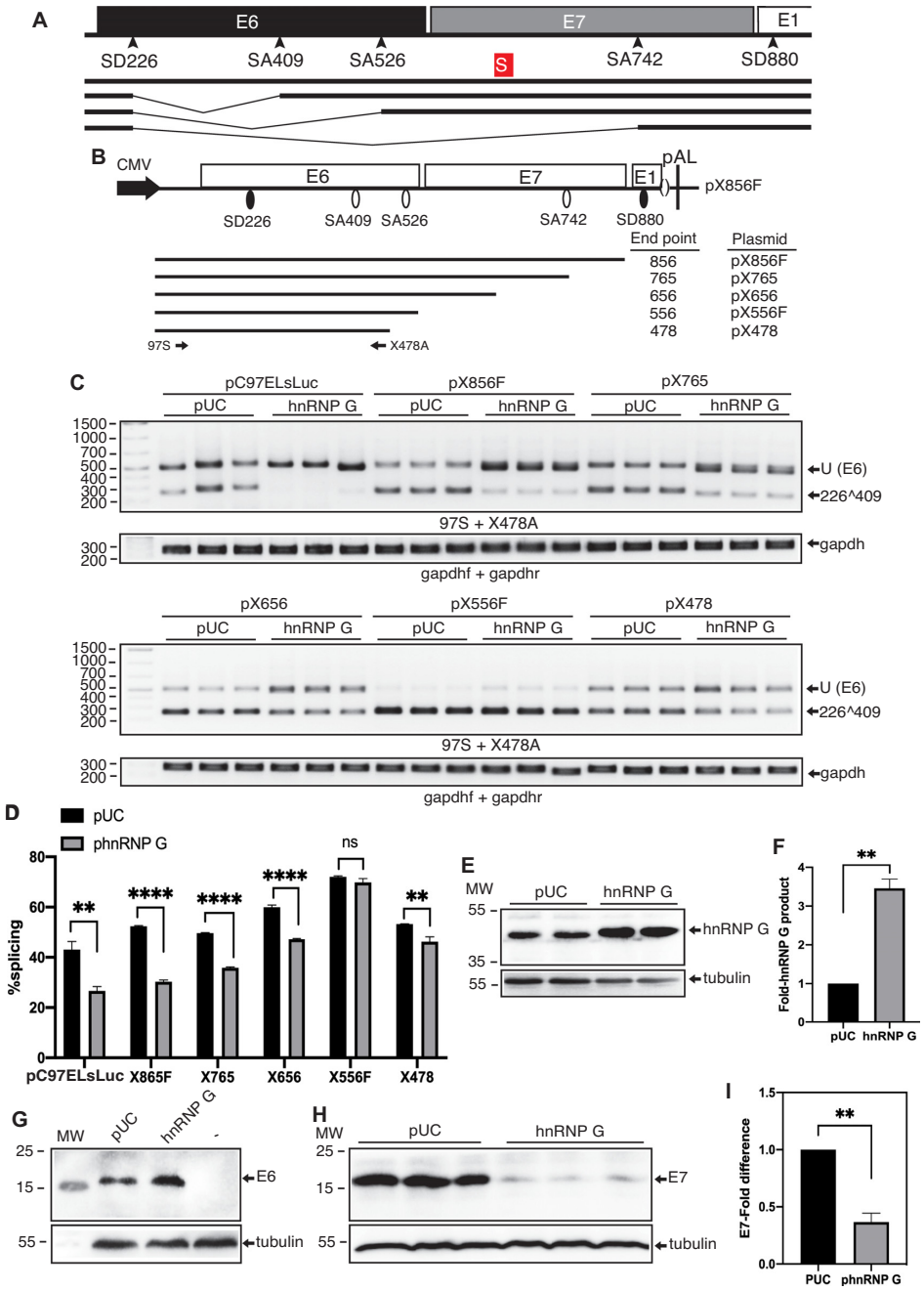


Figure 5. (**A**) Schematic representation of the 5′-end of the HPV16 genome encoding E6, E7 and part of E1. HPV16 splice sites SD226, SA409, SA526, SA742 and SD880 are indicated. A previously identified splicing silencer is indicated as a red box with 'S'. (**B**) Schematic representation of the HPV16 subgenomic expression plasmid pX856F in which HPV16 sequences encoding E6, E7 and part of E1 have been inserted downstream of the CMV promoter. HPV16 splice sites SD226, SA409, SA526, SA742 and SD880 are indicated. Regions and endpoints of sequences included in various HPV16 subgenomic expression plasmids are indicated. Locations of the RT-PCR primers 97S and X478A are indicated. (**C**) RT-PCR with primers 97S and X478A on RNA extracted from HeLa cells co-transfected with the indicated HPV16 plasmids and empty pUC plasmid or hnRNP G-expressing plasmid. Unspliced/intron retained U (E6) mRNA and mRNAs spliced between SD880 and SA409 are indicated to the right of the gel image. (**D**) Densitometric quantitation of triplicate samples in the gel images in (C). Quantitations are displayed as %-splicing in the graph. Mean values and standard deviations are shown. ns, not statistically significant difference. (**E**) Western blotting on cell extracts from HeLa cells co-transfected with pC97ElsLuc and pUC or hnRNP G plasmid. Cells were harvested 20hrs posttransfection. Blots were stained with antibody specific for hnRNP G or tubulin as described in Materials and Methods. (**F**) Densitometric quantification of the hnRNP G band in (E) in the absence or presence of transfected hnRNP G plasmid. The quantitations are displayed as relative levels of hnRNP G in the absence or presence of transfected hnRNP G plasmid. $P < 0.01$. (**G**) Western blotting on cell extracts from HeLa cells transfected with HPV16 reporter plasmid pHAE6E7Flag and empty pUC plasmid or hnRNP G expression plasmid. Blots were stained with antibody specific for HA-tag to detect HA-tagged HPV16 E6 protein or tubulin. (−), untransfected cells. (**H**) Western blotting on cell extracts from HeLa cells transfected with HPV16 reporter plasmid pHAE6E7Flag and empty pUC plasmid or hnRNP G expression plasmid. Blots were stained with antibody specific for FLAG-tag to detect FLAG-tagged HPV16 E7 protein or tubulin. Results from triplicate transfections are shown. (**I**) Densitometric quantification of the FLAG-E7 band in (H) in the absence or presence of transfected hnRNP G plasmid. The quantitations are displayed as relative levels of FLAG-E7 in the absence or presence of transfected hnRNP G plasmid. $P < 0.01$.

the effect of hnRNP G on a series of plasmids with nested deletions in E6 and E7 (Figure 5B). In this case, we used RT-PCR primer pair 97S and X478A that spanned the major splice sites SD226 and SA409 (Figure 5B). As can be seen, hnRNP G inhibited HPV16 E6/E7 mRNA splicing between SD226 and SA409 from plasmids that ended at HPV16 nucleotide positions 856, 765 and 656, but less so on plasmids that ended at positions 478 or 556 (Figure 5C and D). hnRNP G also inhibited splicing of E6/E7 mRNAs from the pC97ELsLuc plasmid as expected (Figure 5C and D). Plasmid pHAE6E7Flag encodes intact E6 and E7 coding regions and E6 is tagged with HA and E7 with Flag. Overexpression of hnRNP G was ∼3-fold (Figure 5E and F) and promoted E6 protein production (Figure 5G) while inhibiting E7 protein production (Figure 5H and I). hnRNP G acted on the E6/E7 coding region to inhibit HPV16 mRNA splicing thereby altering the balance in production of the two HPV16 oncogenes: the levels of the pro-mitotic E7 protein were reduced whereas the levels of the anti-apoptotic E6 protein were increased. We speculated that hnRNP G inhibited E6/E7 mRNA splicing by interacting with a previously identified splicing silencer in the E7 coding region indicated with red 'S' (Figure 5A) (33). We performed a pull-down assay with biotinylated RNA oligos spanning the splicing silencer (marked in red) or adjacent sequences (Supplementary Figure S8A). RNA oligos 604S1 and 620S1 encoded the splicing silencer and efficiently pulled down hnRNP G (Supplementary Figure S8B), whereas adjacent oligo 635S1 pulled down less hnRNP G (Supplementary Figure S8B). Furthermore, silencer mutant RNA oligos 604AM1 and 604AM2 pulled down less hnRNP G than the wild-type silencer RNA oligo 604S1 (Supplementary Figure S8B). These results indicated that hnRNP G interacted preferentially with a splicing silencer in the E7 coding region. We concluded that hnRNP G inhibited splicing of HPV16 E6 and E7 oncogene mRNAs independently of RNA sequences downstream of the E2 splice site SA2709. Thus, hnRNP G had a splicing inhibitory effect on HPV16 mRNAs when the E6/E7 coding sequences were present on the pre-mRNA, but had a splicing enhancing effect on the HPV16 E2 mRNAs when the E6/E7 coding sequences were absent from the HPV16 pre-mRNAs.

Different domains of the hnRNP G protein contribute to inhibition and enhancement of HPV16 mRNA splicing

To determine the contribution of each of various domains of the modular hnRNP G protein to the control of HPV16 mRNA splicing, we generated hnRNP G deletion mutants (Figure 6A). They were all expressed upon transfection of human cells (Figure 6B). To determine the effect of various deletion mutants of hnRNP G on HPV16 E2 mRNA splicing, we first co-transfected plasmids expressing hnRNP G or hnRNP G-mutants pGM1, pGM2 and pGM3 with HPV16 reporter plasmid pBELENdE1 (for a schematic representation of pBELENdE1 and locations of RT-PCR primers, see Supplementary Figure S3B and C) on which we could easily monitor the enhancing effect of hnRNP G on E2 mRNA splicing. The results revealed that the C-terminal mutant GM1 failed to enhance splicing of the E2 mRNAs (Figure 6C and E), whereas C-terminal mutant GM2 retained full splicing enhancing activity (Figure 6C and E), demonstrating that the immediate C-terminus encoding the C-RBD was dispensable for enhancement of HPV16 E2 mRNA splicing, whereas the central part of hnRNP G downstream of amino acid 236 was required for splicing enhancement (Figure 6A, C and E). Interestingly, deletion of the N-terminal RRM in GM3 only partially reduced the ability of hnRNP G to enhance E2 mRNA splicing (Figure 6C and E). To determine the splicing inhibitory effect on HPV16 E6/E7 mRNAs of the various hnRNP G mutants, they were co-transfected with HPV16 reporter plasmid pC97ELsLuc. Locations of RT-PCR primers in HPV16 reporter plasmid pC97ELsLuc are shown in Figure 4D. RT-PCR analysis of HPV16 E6/E7 mRNAs with primers 97S and 880A revealed that C-terminal deletions GM1 and GM2 inhibited E6/E7 mRNA splicing (Figure 6D and F). Surprisingly, inhibition of E6/E7 mRNA splicing by pGM1 was more efficient than inhibition by wild-type hnRNP G (Figure 6D and F), whereas the N-terminal deletion (GM3) inhibited E6/E7 mRNA splicing to a similar extent as the wild type hnRNP G (Figure 6D and F). In conclusion, the central and C-terminal region between hnRNP G amino acid positions 236 and 334 were necessary for enhancement of splicing to SA2709. In contrast, the N-terminal region of hnRNP G played a major role in the inhibition of E6/E7 mRNA splicing. Thus, the deletion mutants separated splicing-enhancing and splicing-inhibitory domains of hnRNP G.

To further map the region of hnRNP G that enhanced HPV16 E2 mRNA splicing, additional C-terminal deletion mutants of hnRNP G were generated (GM8, GM9, GM10) (Figure 6A). As can be seen, GM9 and GM10 fully enhanced E2 mRNA splicing from plasmid pBELENdE1, whereas the ability of GM8 to enhance E2 mRNA splicing was reduced (Figure 7A and B). Locations of RT-PCR primers in HPV16 reporter plasmid pBELENdE1 are shown in Supplementary Figure S3C. These experiments further mapped the splicing enhancing domain of hnRNP G to hnRNP G-sequences between amino acids 236 and 286. To further map the region of hnRNP G that inhibited HPV16 mRNA splicing, additional deletions were introduced in GM1 resulting in GM4, GM5, GM6 and GM7 (Figure 6A). With the exception of GM7 that inhibited HPV16 E6/E7 mRNA splicing, GM4, GM5 and GM6 failed to inhibit E6/E7 mRNA splicing (Figure 7C and D and supplementary Figure S9A and C), mapping the splicing inhibitory region of hnRNP G to hnRNP G-sequences between amino acids 127 and 186. However, an hnRNP G mutant encoding amino acids 109–236 (GM11) (Figure 6A) displayed only weak inhibitory activity (Supplementary Figure S9A and C) (the GM11 was expressed in HeLa cells (Supplementary Figure S9B)). Locations of RT-PCR primers in HPV16 reporter plasmid pC97ELsLuc are shown in Supplementary Figure S9D. Taken together, these results underscored the conclusion that enhancement of HPV16 E2 mRNA splicing and inhibition of HPV16 E6/E7 mRNA splicing were exerted by non-overlapping domains of hnRNP G.

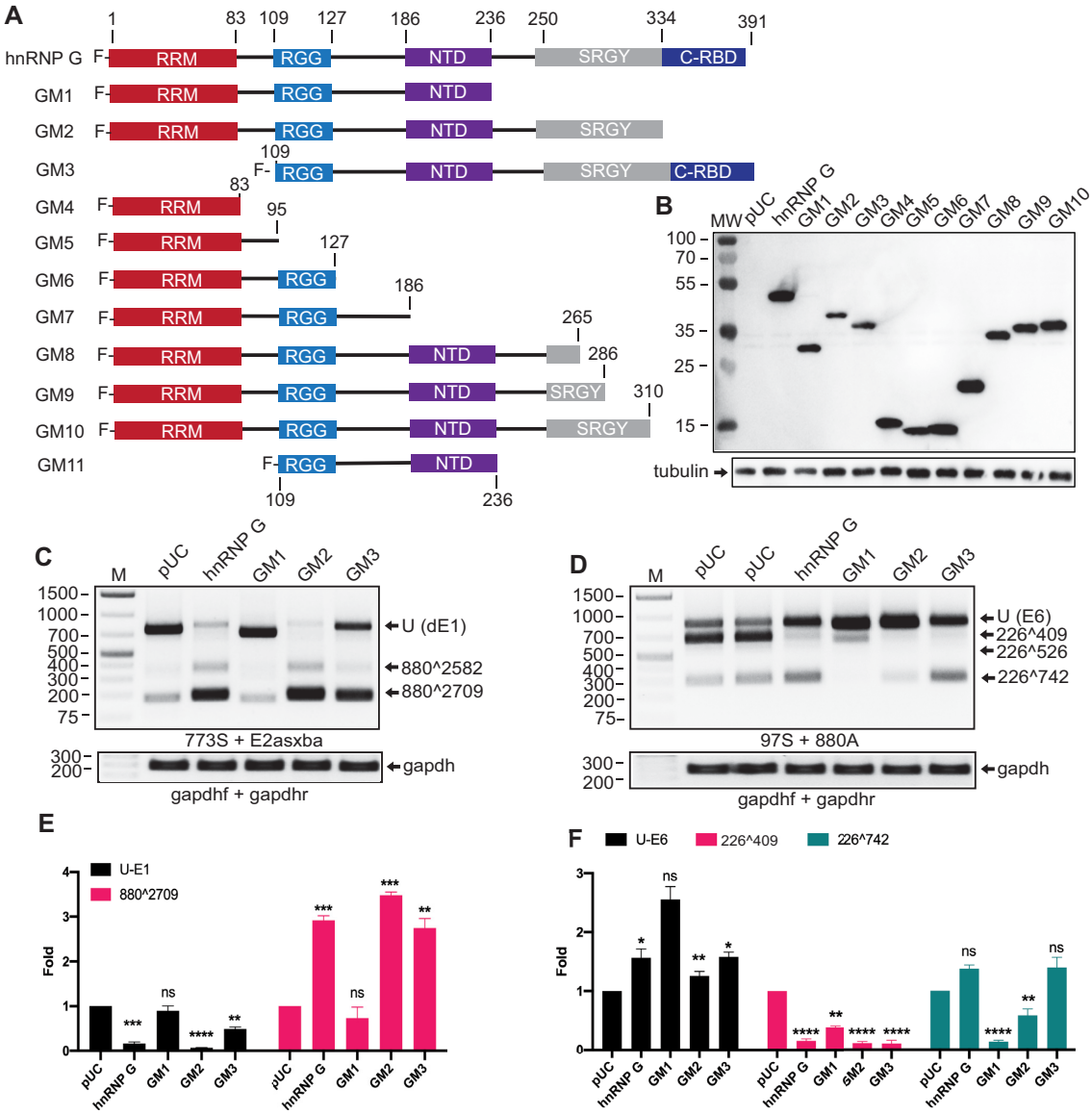


Figure 6. (**A**) Schematic representation of plasmids encoding wild type hnRNP G or various hnRNP G mutants. Amino acids positions in the hnRNP G proteins are indicated. RRM, RNA recognition motif; RGG, arginine-glycine-glycine motif-rich sequence (RGG-box); NTD, nascent RNA targeting domain; serine-arginine-glycine-tyrosine-motif rich domain (SRGY-box); C-RBD, C-terminal RNA binding domain. (**B**) Western blotting on cell extracts from HeLa cells transfected with empty pUC plasmid or plasmids expressing either of the various Flag-tagged hnRNP G proteins displayed in A. Blots were stained with antibody specific for Flag-tag to detect Flag-tagged hnRNP G proteins or with tubulin-specific antibody. (**C**) RT-PCR with primers 773S and E2asxba on RNA extracted from HeLa cells co-transfected with pBELENdE1 and empty pUC plasmid or plasmids expressing wild type or mutant hnRNP G. Unspliced/intron retained U (dE1) mRNAs and mRNAs spliced between SD880 and SA2582 or SA2709 are indicated to the right. Locations of RT-PCR primers in HPV16 reporter plasmid pBELENdE1 are shown in Supplementary Figure S3C. (**D**) RT-PCR with primers 97S and 880as on RNA extracted from HeLa cells cotransfected with pC97ELsLuc and empty pUC plasmid or plasmids expressing wild type or mutant hnRNP G. Unspliced/intron retained U (E6) mRNA and mRNAs spliced between SD880 and SA409, SA526 or SA742 are indicated to the right of the gel image. Locations of RT-PCR primers in HPV16 reporter plasmid pC97ELsLuc are shown in Figure 4D. (**E**) Densitometric quantification of RT-PCR bands (U (dE1) or 880∧2709) in (C). The The quantitations are displayed as relative levels of U (dE1) cDNAs or 880∧2709 in the absence or presence of hnRNP G. Transfections with empty pUC plasmid is set as 1. *P*-values are indicated. (**F**) Densitometric quantification of RT-PCR bands (U (E6), 226∧409 or 226∧742) in (D). The The quantitations are displayed as relative levels of U (E6) cDNAs, 226∧409 or 226∧742 in the absence or presence of hnRNP G. Transfections with empty pUC plasmid is set as 1. *P*-values are indicated.

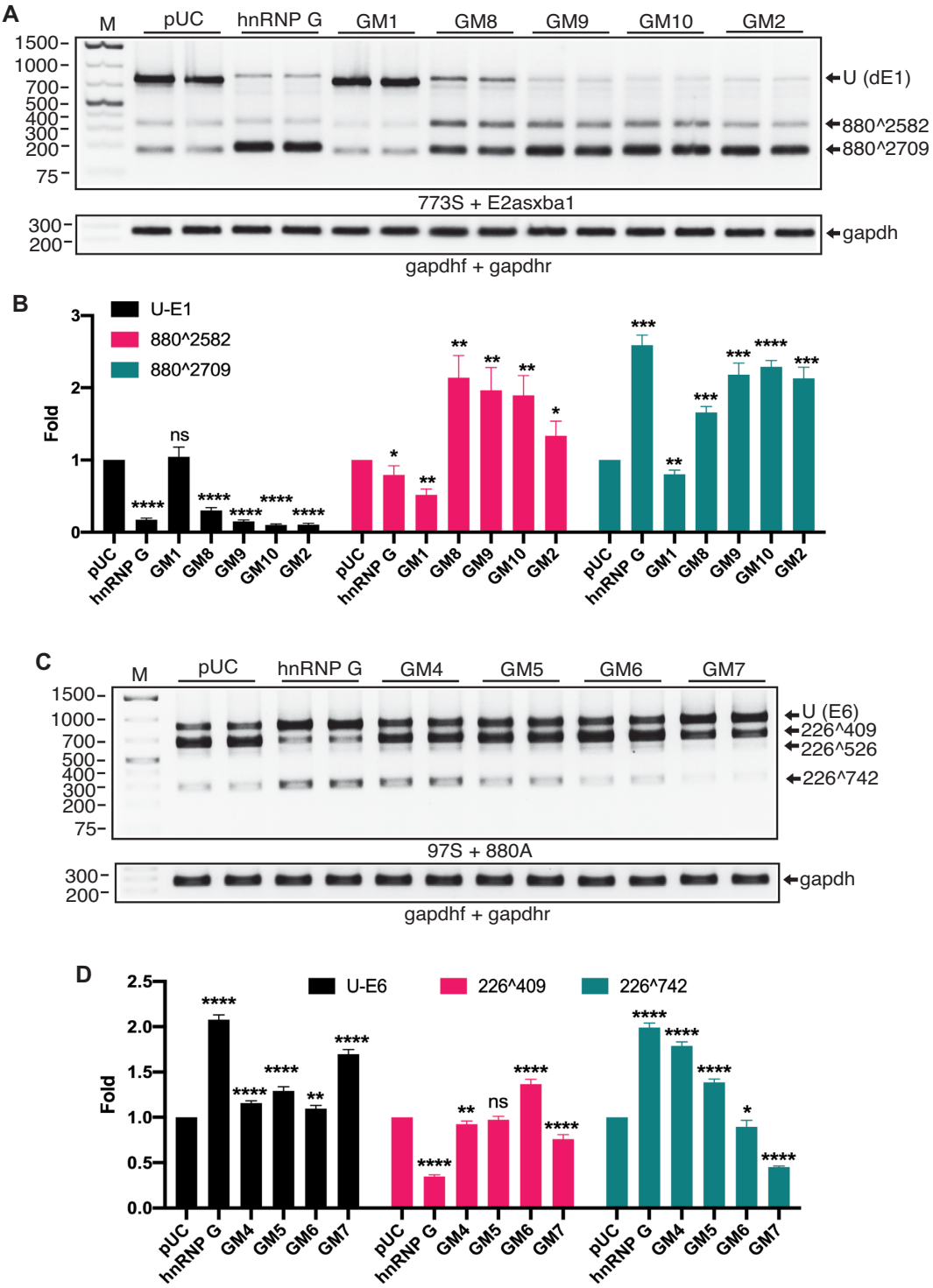


Figure 7. (**A**) RT-PCR with primers 773S and E2asxba on RNA extracted from HeLa cells co-transfected with pBELENdE1 and empty pUC plasmid or plasmids expressing wild type or mutant hnRNP G. Unspliced/intron retained U (dE1) mRNAs and mRNAs spliced between SD880 and SA2582 or SA2709 are indicated to the right. Locations of RT-PCR primers in HPV16 reporter plasmid pBELENdE1 are shown in Supplementary Figure S3C. (**B**) Densitometric quantification of RT-PCR bands (U (dE1) or 880∧2709) in (C). The The quantitations are displayed as relative levels of U (dE1) cDNAs or 880∧2709 in the absence or presence of hnRNP G. Transfections with empty pUC plasmid is set as 1. *P*-values are indicated. (**C**) RT-PCR with primers 97S and 880as on RNA extracted from HeLa cells cotransfected with pC97ELsLuc and empty pUC plasmid or plasmids expressing wild type or mutant hnRNP G. Unspliced/intron retained U (E6) mRNA and mRNAs spliced between SD880 and SA409, SA526 or SA742 are indicated to the right of the gel image. Locations of RT-PCR primers in HPV16 reporter plasmid pC97ELsLuc are shown in Figure 4D. (**D**) Densitometric quantification of RT-PCR bands (U (E6), 226∧409 or 226∧742) in (D). The quantitations are displayed as relative levels of U (E6) cDNAs, 226∧409 or 226∧742 in the absence or presence of hnRNP G. Transfections with empty pUC plasmid is set as 1. *P*-values are indicated.

hnRNP G interacts with splicing factor U2AF65 in a manner that correlates with splicing activation

To delineate the mechanism of action of hnRNP G we investigated if hnRNP G interacted with core splicing factors. We were particularly interested in U2AF65 since it is specific for 3′-splice sites and since hnRNP G stimulated splicing to HPV16 3′-splice site SA2709 through an adjacent splicing enhancer. We performed co-immunoprecipitation of hnRNP G with U2AF65 using extracts from untransfected HeLa cells and from HeLa cells transfected with both hnRNP G and U2AF65 plasmid. Samples were subjected to immunoprecipitation with either IgG or anti-hnRNP G antibody followed by western blotting with anti-U2AF65 antibody or anti-hnRNP A1 antibody as a control. As can be seen, a robust interaction between hnRNP G and U2AF65 was observed in untransfected cells as well as in transfected cells (Figure 8A). In contrast, we did not observe interactions between hnRNP A1 and U2AF65 in the same cell extracts (Figure 8A). Next, we investigated if co-immunoprecipitation of hnRNP G and U2AF65 correlated with the splicing-inhibitory or splicing enhancing functions of the various hnRNP G deletion mutants. Flag-tagged full-length hnRNP G or the various deletion mutants of hnRNP G (GM1, GM2 and GM3) were expressed in transfected cells (Figure 8B). The results revealed that U2AF65 co-immunoprecipitated with wild type hnRNP G as expected, and with hnRNP G mutants GM2 and pGM3, but not with GM1 (Figure 8C), demonstrating that interactions of hnRNP G mutants with U2AF65 corelated with the splicing enhancing function of these mutants (Table 1). Taken together, our results strongly indicated that hnRNP G promoted splicing to HPV16 E2 splice site SA2709 by recruiting U2AF65.

Activation of the cellular DNA damage machinery (DDR) enhances HPV 16 E2 mRNA splicing and interactions of hnRNP G with E2 mRNAs and with U2AF65

Since hnRNP G is a DNA-damage response (DDR) factor (45) and papillomaviruses utilize the DDR for E1/E2-dependent genome replication (6), as well as for control of HPV gene expression (46,47), we investigated if activation of the DDR affected HPV16 E2 gene expression via hnRNP G. We have previously shown that treatment of the HPV16 reporter cell line C33A2 (Supplementary Figure S7A) with melphalan activated the DDR and induced HPV16 late gene expression (46). Reporter cell line C33A2 contains the subgenomic HPV16 plasmid pBELsLuc integrated in the genome and has an sLuc reporter gene located in the late L1 coding region of HPV16 (Supplementary Figure S7A). Furthermore, the C33A2 cell line encodes a mutant, functionally inactive p53 gene which prevents melphalan from inducing apoptosis via DDR in this cell line. Thus, we could investigate the effect of the DDR on HPV16 in the absence of rapid onset of apoptosis. These cells were treated with various concentrations of the DNA-damaging and DDR-inducing drug melphalan, cell culture medium was collected at various time points for sLuc assay and cells were harvested for RNA extraction. As expected, melphalan induced HPV16 late gene expression in a time- and concentration dependent manner (Supplementary Figure S10A). RT-PCR analysis of HPV16 E2 mRNAs, revealed that E2 mRNAs were induced as well (Figure 8D and E). Despite the fact that 100uM melphalan induced the highest levels of sLuc, 50 uM melphalan was the optimal concentration for induction and detection of E2 mRNAs (Figure 8D and E). Locations of RT-PCR primers in HPV16 reporter plasmid pBELsLuc in the C33A2 cells are shown in Supplementary Figure S7C. The optimal time point for cell harvest was between 9–17hours post treatment with melphalan (Figure 8F). Western blot analysis of melphalan-treated C33A2 cells demonstrated that hnRNP G levels increased in response to DDR activation (Figure 8G and H). Immunoprecipitation of RNA-protein complexes from C33A2 cells treated with DMSO or melphalan with anti-hnRNP G antibody followed by RT-PCR with primers specific for HPV16 E2 mRNAs revealed that hnRNP G was associated with E2 mRNAs in these cells and that this association increased in the presence of melphalan (Figure 8I and J). A control experiment in which we showed that anti-hnRNP G antibody immunoprecipitated unspliced, intron retained HPV16 mRNAs from the nuclear fraction of HeLa cells transfected with HPV16 reporter plasmid pBEL, revealed that hnRNP G was associated with unspliced, intron retained HPV16 E1 pre-mRNAs in the nuclear fraction of the cells (Supplementary Figure S10B-D). Interactions between hnRNP G and U2AF65 also increased in the presence of melphalan (Figure 8K). Overexpression of hnRNP G enhanced the association of U2AF65 with HPV16 E2 mRNAs (Figure 8L and M). We concluded that hnRNP G was associated with HPV16 E2 mRNAs in these cells and that this association increased in response to DDR-activation. This is of significance in light of recent findings that hnRNP G is part of the DDR (45) and that HPV activates and utilizes the DDR (6). These results supported the idea that hnRNP G controls HPV16 E2 mRNA splicing.

Phosphorylation of hnRNP G is enhanced by DDR activation, whereas sumoylation is reduced

To investigate how the interaction of hnRNP G with HPV16 mRNAs is controlled, we first determined if hnRNP G was posttranslationally modified under the conditions and cells used here. Cell extracts were prepared from DMSO or melphalan-treated C33A2 cells and subjected to immunoprecipitation with antibodies to serine/threonine/tyrosine (S/T/Y)-phosphorylation or sumoylation, followed by Western blotting to hnRNP G. Results revealed that hnRNP G was phosphorylated and suomylated in these cells (Figure 9A and B) and that phosphorylation of hnRNP G increased in melphalan treated C33A2 cells (Figure 9A), whereas sumoylation of hnRNP G decreased (Figure 9B). Taken together, these results demonstrated that increased hnRNP G phosphorylation and decreased hnRNP G sumoylation correlated with increased interactions of hnRNP G with HPV16 E2 mRNAs.

Inhibition of sumoylation induced HPV16 E2 mRNA splicing and enhanced the association of hnRNP G with HPV16 E2 mRNAs and with U2AF65

Since a decrease in sumoylation of hnRNP G correlated with increased association of hnRNP G with HPV16 E2

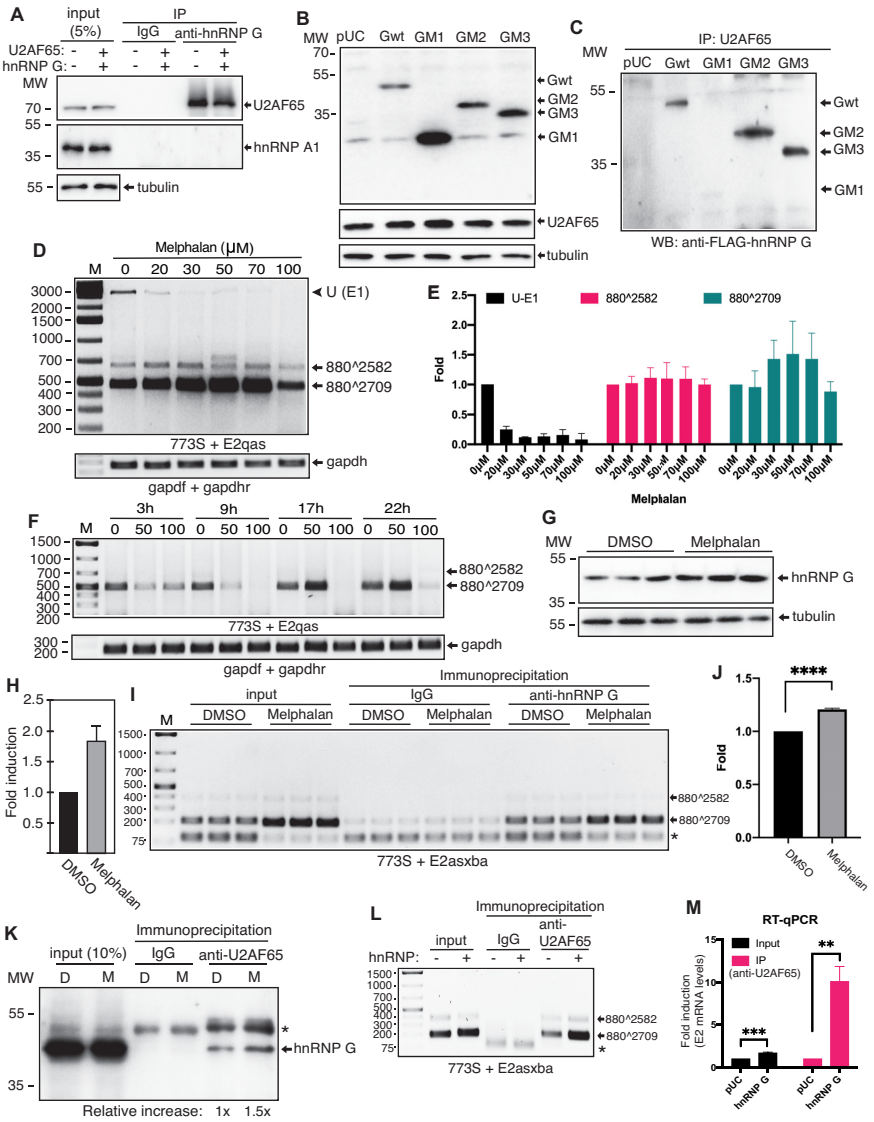


Figure 8. (**A**) Immunoprecipitation of proteins in cell extracts from untransfected (−) HeLa cells or from HeLa cells cotransfected with plasmids encoding U2AF65 (+) and hnRNP G (+) using either IgG or anti-hnRNP G antibody followed by Western blotting with antibodies to either U2AF65 or hnRNP A1. Input samples were subjected to Western blotting with antibodies to U2AF65, hnRNP A1, hnRNP G or tubulin. M, molecular weight marker. (**B**) Western blotting on cell extracts of HeLa cells transfected with plasmid encoding wild type, flag-tagged hnRNP G or the indicated flag-tagged hnRNP G mutants. Western blotting was performed with anti-Flag antibody, anti-U2AF65 antibody or anti-tubulin antibody. (**C**) Coimmunoprecipitation experiment on cell extracts from HeLa cells transfected with pUC-plasmid or Flag-tagged hnRNP G or the indicated Flag-tagged hnRNP G mutants using anti-U2AF65 antibody followed by Western blotting with anti-Flag antibody. (**D**) RT-PCR with primers 773S and E2qas on RNA extracted from C33A2 cells treated with the indicated concentrations of melphalan for 22 h. Locations of RT-PCR primers in HPV16 reporter plasmid pBELsLuc in the C33A2 cells are shown in Supplementary Figure S7B and C. (**E**) Densitometric quantitation of the RT-PCR bands in (D) that represent HPV16 unspliced/retained intron U (E1) mRNAs or mRNAs spliced from SD880 to SA2582 or to SA2709. (**F**) RT-PCR with primers 773S and E2qas on RNA extracted from C33A2 cells treated with the indicated concentrations of melphalan for the indicated time periods. cDNAs representing HPV16 mRNAs spliced between SD880 and SA2582 or SA2709 are indicated to the right. (**G**) Western blotting on cell extracts from C33A2 cells treated with DMSO or 50uM melphalan for 20hrs in triplicates. Blots were stained with antibody specific for hnRNP G or tubulin. (**H**) Densitometric quantitation of the Western blots in D. The quantitations are displayed as relative levels of hnRNP G levels in melphalan-treated cells over DMSO-treated cells. **(I)** C33A2 cells were treated with DMSO or melphalan for 21 h, lysed in RIPA buffer and subjected to immunoprecipitation with IgG or anti-hnRNP G antibody followed by extraction of the immunoprecipitated RNA and RT-PCR with HPV16-specific primers 773s and E2asxba. Input represents RT-PCR with the same primers on RNA extracted from 5% of the input in the immunoprecipitation incubation. cDNAs representing HPV16 mRNAs spliced between SD880 and SA2582 or SA2709 are indicated to the right. *, primer dimers**.** (**J**) Densitometric quantitation of the RT-PCR bands in (I) that represent HPV16 E2 mRNAs spliced from SD880 to SA2709. The quantitations are displayed as relative levels of 880∧2709 cDNAs obtained from anti-hnRNP G immunoprecipitated RNA-protein complexes obtained from melphalan-treated cells over DMSO-treated cells. (**K**) Immunoprecipitation of proteins in cell extracts from DMSO (D) or melphalan (M)-treated C33A2 cells using either IgG or anti-U2AF65 antibody followed by Western blotting with antibody to hnRNP G. M, molecular weight marker. (**L**) HeLa cells transfected with HPV16 reporter plasmid pBEL in the absence or presence of hnRNP G plasmid were lysed in RIPA buffer and subjected to immunoprecipitation with IgG or anti-U2AF65 antibody followed by extraction of the immunoprecipitated RNA and RT-PCR by HPV16 specific primers 773s and E2asxba. Input represents RT-PCR with the same primers on RNA extracted from 5% of the input in the immunoprecipitation incubation. M, molecular size marker; *, primer dimers. (**M**) RT-qPCR with primers 773S and E2asxba was performed as described in Materials and Methods on the RNA analyzed in (L). **, $P < 0.001$.

Table 1. Summary of the splicing inhibitory effect of wild type hnRNP G protein or mutant hnRNP G proteins GM1, GM2 or GM3 on HPV16 E6/E7 mRNA splicing and of the activation of HPV16 E2 mRNA splicing and of co-immunoprecipitation with U2AF65. Number of '+' indicates degree of E6/E7 mRNA splicing inhibition, E2 mRNA splicing activation or co-immunoprecipitation of U2AF65. -, no effect on mRNA splicing or no detectable co-immunoprecipitation

hnRNP G	Inhibition E6/E7 mRNA splicing	Activation E2 mRNA splicing[a]	U2AF65 coIP
wt	+	+++	+
GM1	+++	-	-
GM2	++	+	+
GM3	+	++	+

[a]in the absence of upstream E6/E7 sequences.

mRNAs and increased E2 mRNA levels, we investigated if pharmacological inhibition of sumoylation in C33A2 cells could induce HPV16 E2 mRNA splicing. Sumoylation inhibitors 2-D08 and TAK-981 significantly induced higher levels of E2 mRNAs in C33A2 cells (Figure 9C and D and Supplementary Figure S10E and F). Locations of RT-PCR primers in HPV16 reporter plasmid pBELsLuc in the C33A2 cells are shown in Supplementary Figure S7C. To ascertain that hnRNP G was associated with the increased E2 mRNA levels, we performed RIP experiments. The results demonstrated an increase in the association of hnRNP G with HPV16 E2 mRNAs in the presence of the sumoylation inhibitor 2-D08 or TAK-981 (Figure 9E and Supplementary Figure S10G). Furthermore, treatment of the C33A2 cells with 2-D08 or TAK-981 also enhanced interactions between hnRNP G and the splicing factor U2AF65 (Figure 9F and Supplementary Figure S10H). Taken together, these results suggested that HPV16 E2 mRNA splicing was mediated by interactions between hnRNP G and the HPV16 E2 mRNA splicing enhancer and the splicing factor U2AF65, and that these interactions were regulated by protein sumoylation.

Association between HPV16 E2 mRNAs and hnRNP G in HPV16-immortalized human keratinocytes increases in response to cell differentiation

HPV16 gene expression is intimately linked to cell differentiation (4,6,48). We therefore investigated if calcium-induced differentiation of the previously described HPV16-immortalized human keratinocyte 3310 cell line (44) affected the interactions between hnRNP G and HPV16 E2 mRNAs. As can be seen, immunoprecipitation of RNA-protein complexes from extracts of 3310 cells with anti-hnRNP G antibody demonstrated that hnRNP G was associated with HPV16 E2 mRNAs in 3310 cells (Figure 10A and B). Treatment of the 3310 cells with calcium induced E2 mRNA production and enhanced association of HPV16 E2 mRNAs with hnRNP G (Figure 10C and E). Locations of RT-PCR primers in the HPV16 genome are shown in Figure 10J. RT-PCR on the involucrin differentiation marker confirmed that calcium induced cell differentiation in the 3310 cells (Figure 10D). We concluded that hnRNP G was associated with HPV16 E2 mRNAs in HPV16-immortalized

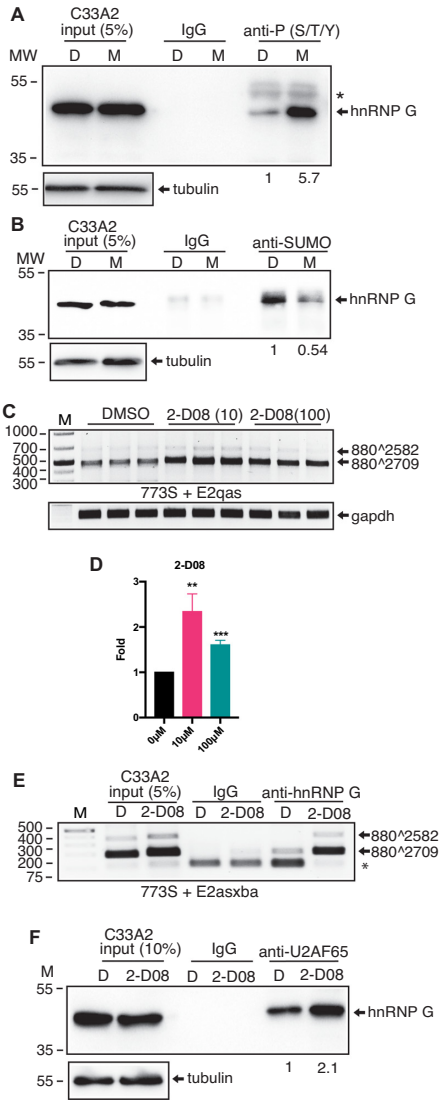


Figure 9. (**A**) Immunoprecipitation of proteins in cell extracts from DMSO (D) or melphalan (M) treated C33A2 cells using either IgG or anti-P(S/T/Y) antibody followed by western blotting with antibody to hnRNP G. Input samples represent 5% of the sample volume used for immunoprecipitation. MW, molecular weight marker; *, unknown band. (**B**) Immunoprecipitation of proteins in cell extracts from DMSO (D) or melphalan (M) treated C33A2 cells using either IgG or anti-sumo antibody followed by western blotting with antibody to hnRNP G. (**C**) C33A2 cells were treated with DMSO (D) or 10- or 100uM of sumoylation inhibitor 2-D08 followed by RNA extraction and RT-PCR with HPV16-specific primers 773s and E2qas. Locations of RT-PCR primers in HPV16 reporter plasmid pBELsLuc in the C33A2 cells are shown in Supplementary Figure S7B and C. (**D**) Densitometric quantification of the RT-PCR band in (C) representing HPV16 880∧2709 spliced mRNAs. The quantitations are displayed as relative levels of 880∧2709 cDNAs in 2-D08-treated cells over DMSO-treated cells. *P*-values are indicated. (**E**) C33A2 cells were treated with DMSO (D) or 10- or 100 uM of sumoylation inhibitor 2-D08 followed by lysis in RIPA buffer and immunoprecipitation with IgG or anti-hnRNP G antibody. RNA was extracted from the Immunoprecipitated RNA-protein complexes and subjected to RT-PCR with HPV16-specific primers 773s and E2asxba. Locations of RT-PCR primers in HPV16 reporter plasmid pBELsLuc in the C33A2 cells are shown in Supplementary Figure S7B and C. *, primer dimer. (**F**) Immunoprecipitation of proteins in cell extracts from DMSO (D) or sumoylation inhibitor 2-D08-treated C33A2 cells using either IgG or anti-U2AF65 antibody followed by western blotting with antibody to hnRNP G. Input samples represent 10% of the sample volume used for immunoprecipitation.

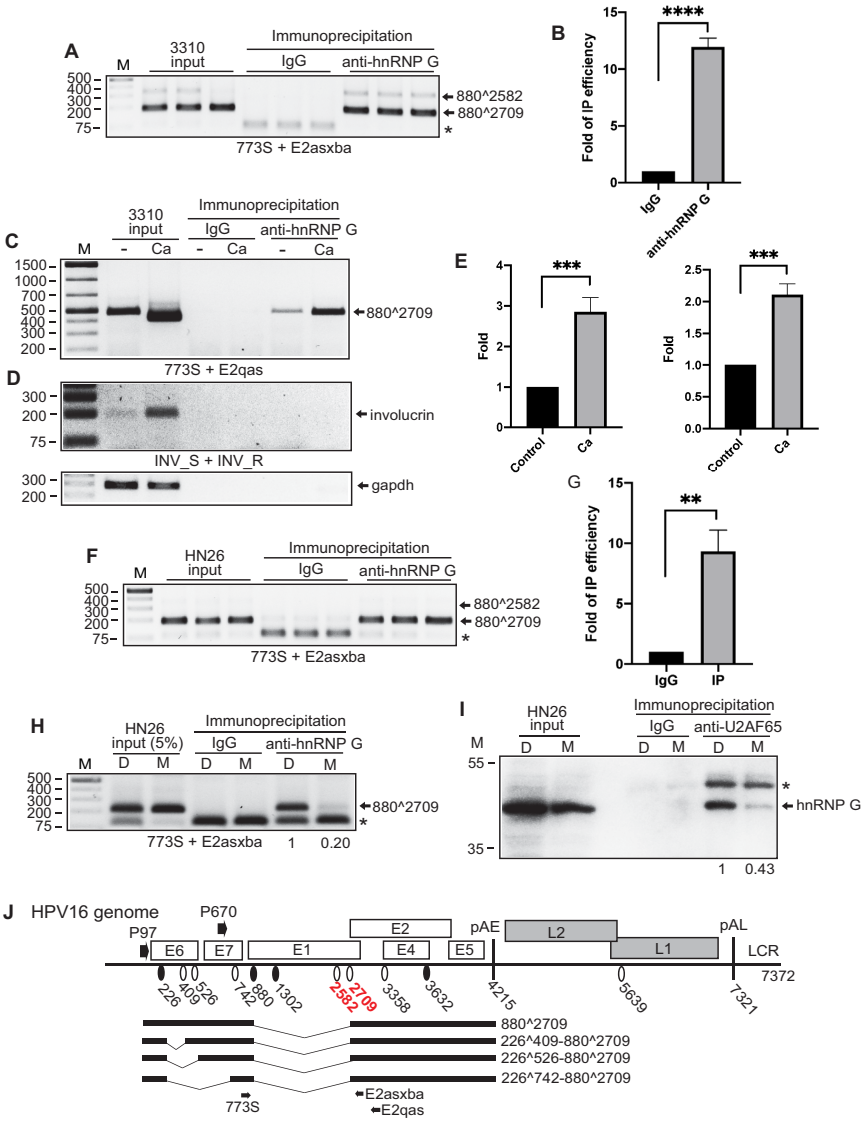


Figure 10. (**A**) HPV16-immortalized human keratinocyte cell line 3310 cells were lysed in RIPA buffer and subjected to immunoprecipitation with IgG or anti-hnRNP G antibody followed by extraction of the immunoprecipitated RNA and RT-PCR by HPV16 specific primers 773s and E2asxba. Input represents RT-PCR with the same primers on RNA extracted from 5% of the input in the immunoprecipitation incubation. M, molecular size marker; *, primer dimers. (**B**) Densitometric quantitation of the RT-PCR bands in (A) that represent HPV16 E2 mRNAs spliced from SD880 to SA2709. The quantitations are displayed as relative levels of 880∧2709 cDNAs obtained from anti-hnRNP G immunoprecipitated RNA–protein complexes over IgG immunoprecipitated complexes. (**C**) HPV16-immortalized human keratinocyte cell line 3310 cells were treated with 2.4 mM $CaCl_2$ to induce differentiation, lysed in RIPA buffer and subjected to immunoprecipitation with IgG or anti-hnRNP G antibody followed by extraction of the immunoprecipitated RNA and RT-PCR by HPV16 specific primers 773s and E2asxba. Input represents RT-PCR with the same primers on RNA extracted from 5% of the input in the immunoprecipitation incubation. (**D**) RT-PCR with primers specific for mRNAs encoding the cell differentiation marker involucrin performed on the cDNA samples used in (C). (**E**) Densitometric quantitation of the RT-PCR bands in (C) that represent HPV16 E2 mRNAs spliced from SD880 to SA2709. The quantitations are displayed as relative levels of 880∧2709 cDNAs obtained from $CaCl_2$-treated cells over untreated cells. (**F**) Cells from the HPV16-positive tonsillar cancer cell line HN26 were lysed in RIPA buffer and subjected to immunoprecipitation with IgG or anti-hnRNP G antibody followed by extraction of the immunoprecipitated RNA and RT-PCR by HPV16 specific primers 773s and E2asxba. Input represents RT-PCR with the same primers on RNA extracted from 5% of the input in the immunoprecipitation incubation. M, molecular size marker; *, primer dimers. (**G**) Densitometric quantitation of the RT-PCR bands in (F) that represent HPV16 E2 mRNAs spliced from SD880 to SA2709. The quantitations are displayed as relative levels of 880∧2709 cDNAs obtained from anti-hnRNP G immunoprecipitated RNA-protein complexes over IgG immunoprecipitated complexes. (**H**) HN26 cells were treated with DMSO (D) or melphalan (M) for 21 h, lysed in RIPA buffer and subjected to immunoprecipitation with IgG or anti-hnRNP G antibody followed by extraction of the immunoprecipitated RNA and RT-PCR with HPV16-specific primers 773s and E2asxba. Input represents RT-PCR with the same primers on RNA extracted from 5% of the input in the immunoprecipitation incubation. cDNAs representing HPV16 mRNAs spliced between SD880 and SA2709 are indicated to the right. *, primer dimers. (**I**) Immunoprecipitation of proteins in cell extracts from DMSO (D) or melphalan (M) treated HN26 cells using either IgG or anti-U2AF65 antibody followed by western blotting with antibody to hnRNP G. Input samples represent 10% of the sample volume used for immunoprecipitation. (**J**) Linearized HPV16 genome (numbers refer to the HPV16 reference strain GeneBank: K02718.1). Early and late genes are indicated. P97: HPV16 early promoter. P670: HPV16 late promoter. Black oval: 5′-splice site/splice donor. White oval: 3′-splice site/splice acceptor. Splice acceptors SA2582 and SA2709 located upstream of the E2 open reading frame are indicated in red. pAE: HPV16 early polyadenylation site. pAL: HPV16 late polyadenylation site. LCR: long control region. A subset of alternatively spliced HPV16 E2 mRNAs all utilizing HPV16 3′-splice site SA2709. Arrows indicate RT-PCR primers used to detect the HPV16 E2 mRNAs in HPV16 immortalized 3310 cells or tonsillar cancer cell line HN26.

human keratinocytes and that this interaction increased in response to cell differentiation.

Melphalan treatment of apoptosis-prone HN26 cells reduced E2 mRNA levels, reduced hnRNP G levels and reduced association of hnRNP G with U2AF65

We have previously shown that melphalan treatment of the p53-negative C33A2 cell line activated DDR but failed to induce apoptosis (46). In contrast, melphalan treatment of the tonsillar cancer cell line HN26 resulted in a block of HPV16 transcription, degradation of the HPV16 early mRNAs and liberation of the wild-type p53 protein (43) immediately followed by p53-mediated apoptosis (49). We therefore wished to investigate the effect of melphalan on the association of hnRNP G with HPV16 E2 mRNAs also in HN26 cells. An RNA-immunoprecipitation experiment demonstrated that HPV16 E2 mRNAs were associated with hnRNP G in the HN26 cells (Figure 10F and G). Treatment of HN26 cells with melphalan resulted in reduced association of hnRNP G with the HPV16 E2 mRNAs (Figure 10H) and reduced the association of hnRNP G with U2AF65 (Figure 10I). U2AF65 did not interact with hnRNP A1 in HN26 cells either (data not shown). We concluded that treatment of the HPV16-positive tonsillar cancer cell line HN26 with the cancer drug melphalan, reduced the association of hnRNP G with HPV16 E2 mRNAs and with U2AF65, which may contribute to the previously described, melphalan-induced reduction of HPV16 early mRNAs including the E2 mRNAs in HN26 cells (49).

DISCUSSION

The AG-rich and U-less HPV16 splicing enhancer ACGAGGACGAGGACAAGG is composed of 42% A, 42% G and 16% C and harbors three AC(G/A)AGG repeats. Repeats 1 and 3 were more important to splicing, whereas repeat 2 appeared to contribute less to the enhancer activity. However, when all three repeats were mutated, splicing to SA2709 was significantly impaired, indicating that this 18nt RNA sequence constitutes the core part of the splicing enhancer. The splicing enhancer is strictly conserved among all HPV16 lineages and sublineages while bigger variation was seen among other HPV-species. It will be of interest to determine how well conserved this splicing enhancer is among a greater number of papillomaviruses since all known papillomaviruses presumably produce E2 protein from spliced mRNAs.

The HPV16 E2 splicing enhancer interacted specifically with multiple RNA binding proteins. It remains to be determined if they bind RNA directly or if they acquire specificity for the enhancer via protein-protein interactions. hnRNP G is a multifunctional protein that interacted with the HPV16 E2 enhancer. Overexpression and knock-down experiments demonstrated a role for hnRNP G in the control of HPV16 mRNA splicing. hnRNP G may bind RNA directly and has affinity for purine-rich sequences for example GGAAA, but also for shorter sequence motifs such as AAGU or AAN (50) (51) (52) (53). Since the HPV16 E2 splicing enhancer at SA2709 is AG-rich, one may speculate that hnRNP G binds directly to the HPV16 splicing enhancer, but the exact mode of interaction between hnRNP G and the HPV16 splicing enhancer remains to be determined. Furthermore, sequences with alternative nucleotide composition have been reported as binding sites for hnRNP G as well, including CCC and CCA (53), and they are less similar to the HPV16 enhancer. hnRNP G has a classical RRM in its N-terminus and it has been shown that hnRNP G binds RNA via the RRM, as does the hnRNP G paralogous protein RBMY (52,53). However, some reports suggest that other regions of hnRNP G or its paralogous bind directly to RNA. It was reported that hnRNP G has a second RNA binding domain in its C-terminus named C-terminal RNA binding domain (C-RBD) (50). The RRM and the C-RBD of hnRNP G appear to have distinct RNA-binding specificities. While RRM seems to prefer CCC and CCA, AAGU, or AAN, C-RBD has affinity for longer, purine-rich GGAAA-sequences (53) (52). Furthermore, the C-RBD also binds directly to m6A-modified RNA (54,55). We found that the C-RBD of hnRNP G was dispensable for activation of HPV16 E2-mRNA splice site SA2709 and for inhibition of E6/E7 mRNA splicing (GM2, Figure 6A, C and D), indicating that C-RBD does not play a major role in the control of HPV16 early mRNA splicing, at least not under the experimental conditions used here. Our results also showed that deletion of the RRM of hnRNP G impaired the ability of hnRNP G to promote E2 mRNA splicing, but in no means did the absence of RRM abolish the splicing-promoting function of hnRNP G (GM3, Figure 6A and C). hnRNP G lacking the RRM also inhibited production of the HPV16 E7 mRNAs spliced 226∧409, but rather than efficiently promoting production of intron-containing E6-mRNAs, it primarily redirected splicing from the 226∧409-mRNA to production of the alternatively spliced 226∧742 mRNA, and therefore displayed a phenotype that was different from the wild type hnRNP G phenotype. Thus, hnRNP G RRM is not totally required for the enhancing effect of the HPV16 E2 mRNA splicing, nor for the splicing inhibitory effect of hnRNP G on the HPV16 E6/E7 mRNAs, indicating that interactions of hnRNP G with HPV16 RNA via the RRM is not an absolute necessity for control of HPV16 mRNA splicing. However, loss of RNA-binding via the hnRNP G RRM may be compensated for by the C-terminal RNA binding domain of hnRNP G, C-RBD. Alternatively, hnRNP G acquires specificity for HPV16 mRNAs through protein-protein interactions.

We show that hnRNP G exerted diametrically different effects on splicing of the HPV16 mRNAs initiated at nucleotide position 97 (which is the transcription start of the HPV16 early promoter P97) or in the E7 coding region, similar to the position of the HPV16 late promoter P670. We speculate that when the HPV16 early promoter is active and the HPV16 pre-mRNAs all contain the E6 and E7 coding regions, hnRNP G has a splicing inhibitory effect that inhibits production of E7 mRNAs (Figure 11). When transcription of the HPV16 genome switches from the early P97 promoter to the late P670 promoter, hnRNP G has an entirely positive effect on E2 mRNA splicing through interactions with the E2 mRNA splicing enhancer downstream of HPV16 3′-splice site SA2709 (Figure 11). Intriguingly, we were able to separate the splicing-enhancing and splicing-inhibitory functions of the 391-amino acid, hnRNP G protein. The splicing-enhancing function of hnRNP G mapped

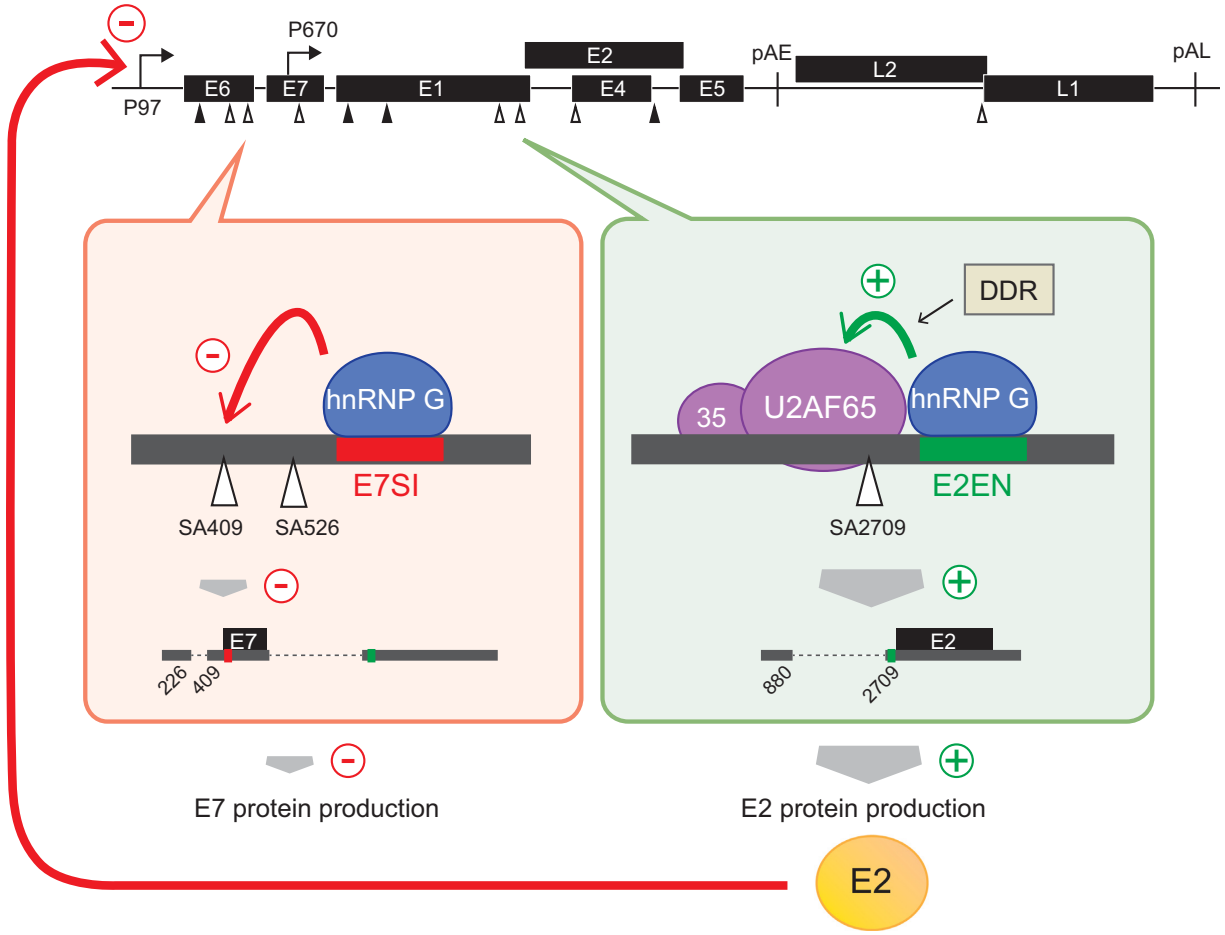


Figure 11. Schematic representation of the HPV16 genome. The early P97 promoter and the late P670 promoter are indicated. HPV16 splice sites and early and late polyadenylation signals pAE and pAL, respectively, are indicated. hnRNP G interacts with HPV16 sequences in the E6/E7 coding region that are present on HPV16 mRNAs produced from the HPV16 early promoter P97 and promotes retention of the E6-coding intron, at the same time inhibiting splicing and inhibiting production of the E7 mRNAs and E7 protein. On HPV16 mRNAs lacking E6/E7 sequences, i.e. mRNAs produced from the HPV16 late promoter P670, hnRNP G enhances splicing to E2 mRNA-specific 3′-splice site SA2709 which results in increased E2 mRNA production. Increased E2 protein levels would contribute to a transcriptional shut-down of the early HPV16 P97 promoter further reducing production of E7 mRNAs and E7 protein thereby paving the way for cell differentiation and induction of late gene expression.

to a region between hnRNP G amino acids 236 and 286, whereas the splicing-inhibitory function mapped to a region between hnRNP G amino acids 127 and 186. The splicing inhibitory region of hnRNP G is located in between the RGG-domain (109–127) and the previously identified nascent RNA targeting domain (NTD) (186 and 236) in hnRNP G (50). The NTD domain of hnRNP G was defined based on its ability to associate with nascent RNA in Xenopus oocytes, apparently through protein-protein interactions rather than direct RNA binding (50). If hnRNP G inhibits HPV16 E6/E7 mRNA splicing via protein-protein interactions, the protein targets remain to be identified to fully understand the role of hnRNP G in HPV16 gene regulation. The splicing-enhancing region of hnRNP G overlaps with a sequence referred to as SRGY-domain originally described for the RNA-binding motif, Y chromosome (RBMY) protein, and so named for the presence of four serine-arginine-glycine-tyrosine-boxes (SRGY)-boxes in this region in RBMY (56). However, hnRNP G contains only one such motif and its significance is unknown. We found that the hnRNP G sequences between amino acids 236 and 286 were required for interactions with general splicing factor U2AF65, suggesting a model for enhancement if HPV16 E2 mRNA splicing by hnRNP G in which hnRNP G interacts with the HPV16 splicing enhancer, directly or indirectly, to recruit U2AF65 to HPV16 E2 splice site SA2709 (Figure 11). However, the exact amino acids within the 236–286 sequence that are required for recruitment of U2AF65 remain to be identified. We are currently attempting to identity amino acid positions in hnRNP G that are required for activation of HPV16 E2 mRNA splicing.

The hnRNP G protein is part of the DNA damage response (DDR) machinery (45). HPVs have been shown to induce DDR and utilize it for DNA replication (6,57–59) as well as for induction of HPV16 late gene expression (46,47). hnRNP G controls alternative splicing of the HPV16 late L1 mRNAs (41) and here we show that hnRNP G controls splicing of the HPV16 E2 mRNAs through the HPV16 E2 splice site SA2709 via a splicing enhancer situated immedi-

ately downstream of SA2709. Activation of hnRNP G by cell differentiation and the DDR may converge on the E2 splicing enhancer identified here to boost production of E2 mRNAs, thereby securing production of sufficient quantities of E2 protein for the increased HPV genomic replication observed in differentiated cells (Figure 11). It has previously been shown that high levels of HPV E2 inhibits the HPV early promoter (10,11). Thus, it is reasonable to speculate that in HPV16 infected cells that have replicated HPV16 genomic DNA to high levels and are set up to generate virus particles, the E2 protein shuts down the HPV16 early promoter and represses expression of the promitotic and anti-apoptotic E6 and E7 proteins, thereby paving the way for cell differentiation and induction of HPV16 late gene expression. We speculate that differentiation-dependent activation of hnRNP G inhibits E6/E7 mRNA splicing and thereby production of E7 mRNA and E7 protein, which enables the cell to leave its mitotic state (Figure 11). Reduced E7 protein levels would allow activation of the HPV16 late promoter that produces pre-mRNAs lacking E6 and E7 sequences thereby allowing stimulation of E2 mRNA splicing by hnRNP G. As a result, E2 protein levels would increase. High levels of E2 protein would bind to the HPV16 early promoter in competition with cellular transcription factors and shut down HPV16 early mRNA production. Thus, hnRNP G and E2 may join forces to complete the switch to the late stage of the HPV16 replication cycle (Figure 11). hnRNP G promotes production of pro-apoptotic E2 and prevents production of E7 protein by inhibiting splicing between SD226 and SA409 and indirectly inhibiting both E6 and E7 expression by promoting production of E2 that shuts down the early promoter P97 (Figure 11). Thus, high levels of E2 would counteract cell transformation and progression to cancer of HPV16-infected cells. One may therefore speculate that hnRNP G would have tumor suppressive properties. Although a few reports have suggested that under-expression of hnRNP G predicted a favorable outcome of head and neck cancer and T-cell lymphomas (60,61), the vast number of publications of the role of hnRNP G in cancer indicate that high expression of hnRNP G correlates significantly with a favorable outcome of cancer and suggest that hnRNP G has tumor suppressive effects (45,62–67). These latter reports are in line with the results presented here that hnRNP G promotes production of the pro-apoptotic HPV16 E2 protein while simultaneously shutting down production of the HPV16 E6 and E7 oncogenes.

DATA AVAILABILITY

Original data is available on request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Dr I. C. Eperon for generously providing hnRNP G plasmid and to Lijing Gong, Shirin Shoja Chaghervand, Júlia Ortís Sunyer, Neele Bergemann and Kersti Nilsson for contributing to some experiments and discussions.

FUNDING

Swedish Research Council-Medicine [VR2019-01210 to S.S.]; Swedish Cancer Society [CAN2018/702 to S.S.]; China Scholarship Council [201706170061 to C.H., 201606525004 to Y.Z., 201809120016 to X.C.]. Funding for open access charge: Swedish Research Council [VR2019-01210 to S.S.].
Conflict of interest statement. None declared.

REFERENCES

1. Walboomers,J.M., Jacobs,M.V., Manos,M.M., Bosch,F.X., Kummer,J.A., Shah,K.V., Snijders,P.J., Peto,J., Meijer,C.J. and Munoz,N. (1999) Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J. Pathol.*, **189**, 12–19.
2. Bouvard,V., Baan,R., Straif,K., Grosse,Y., Secretan,B., El Ghissassi,F., Benbrahim-Tallaa,L., Guha,N., Freeman,C., Galichet,L. *et al.* (2009) A review of human carcinogens–Part B: biological agents. *Lancet Oncol.*, **10**, 321–322.
3. Kajitani,N., Satsuka,A., Kawate,A. and Sakai,H. (2012) Productive lifecycle of human papillomaviruses that depends upon squamous epithelial differentiation. *Front. Microbiol.*, **3**, 152.
4. Chow,L.T., Broker,T.R. and Steinberg,B.M. (2010) The natural history of human papillomavirus infections of the mucosal epithelia. *APMIS*, **118**, 422–449.
5. Schiffman,M., Doorbar,J., Wentzensen,N., de Sanjose,S., Fakhry,C., Monk,B.J., Stanley,M.A. and Franceschi,S. (2016) Carcinogenic human papillomavirus infection. *Nat. Rev. Dis. Primers*, **2**, 16086.
6. Hong,S. and Laimins,L.A. (2013) Regulation of the life cycle of HPVs by differentiation and the DNA damage response. *Fut. Microbiol.*, **8**, 1547–1557.
7. Roman,A. and Munger,K. (2013) The papillomavirus E7 proteins. *Virology*, **445**, 138–168.
8. Vande Pol,S.B. and Klingelhutz,A.J. (2013) Papillomavirus E6 oncoproteins. *Virology*, **445**, 115–137.
9. Bergvall,M., Melendy,T. and Archambault,J. (2013) The E1 proteins. *Virology*, **445**, 35–56.
10. Thierry,F. (2009) Transcriptional regulation of the papillomavirus oncogenes by cellular and viral transcription factors in cervical carcinoma. *Virology*, **384**, 375–379.
11. McBride,A.A. (2013) The papillomavirus E2 proteins. *Virology*, **445**, 57–79.
12. Kadaja,M., Silla,T., Ustav,E. and Ustav,M. (2009) Papillomavirus DNA replication - from initiation to genomic instability. *Virology*, **384**, 360–368.
13. Johansson,C. and Schwartz,S. (2013) Regulation of human papillomavirus gene expression by splicing and polyadenylation. *Nat. Rev. Microbiol.*, **11**, 239–251.
14. Johansson,C., Somberg,M., Li,X., Backström Winquist,E., Fay,J., Ryan,F., Pim,D., Banks,L. and Schwartz,S. (2012) HPV-16 E2 contributes to induction of HPV-16 late gene expression by inhibiting early polyadenylation. *EMBO J.*, **13**, 3212–3227.
15. Wu,S.Y., Nin,D.S., Lee,A.Y., Simanski,S., Kodadek,T. and Chiang,C.M. (2016) BRD4 phosphorylation regulates HPV E2-Mediated viral transcription, origin replication, and cellular MMP-9 expression. *Cell Rep.*, **16**, 1733–1748.
16. Hwang,E.S., Riese,D.J., 2nd,Settleman, Nilson,J., Honig,L.A., Flynn,J. and DiMaio,D. (1993) Inhibition of cervical carcinoma cell line proliferation by the introduction of a bovine papillomavirus regulatory gene. *J. Virol.*, **67**, 3720–3729.
17. Iftner,T., Haedicke-Jarboui,J., Wu,S.Y. and Chiang,C.M. (2017) Involvement of BRD4 in different steps of the papillomavirus life cycle. *Virus Res.*, **231**, 76–82.
18. Nishimura,A., Ono,T., Ishimoto,A., Dowhanick,J.J., Frizzell,M.A., Howley,P.M. and Sakai,H. (2000) Mechanisms of human

papillomavirus E2-mediated repression of viral oncogene expression and cervical cancer cell growth inhibition. *J. Virol.*, **74**, 3752–3760.
19. Demeret,C., Garcia-Carranca,A. and Thierry,F. (2003) Transcription-independent triggering of the extrinsic pathway of apoptosis by human papillomavirus 18 E2 protein. *Oncogene*, **22**, 168–175.
20. Parish,J.L., Kowalczyk,A., Chen,H.T., Roeder,G.E., Sessions,R., Buckle,M. and Gaston,K. (2006) E2 proteins from high- and low-risk human papillomavirus types differ in their ability to bind p53 and induce apoptotic cell death. *J. Virol.*, **80**, 4580–4590.
21. Groves,I.J. and Coleman,N. (2018) Human papillomavirus genome integration in squamous carcinogenesis: what have next-generation sequencing studies taught us? *J. Pathol.*, **245**, 9–18.
22. Lai,M.C., Teh,B.H. and Tarn,W.Y. (1999) A human papillomavirus E2 transcriptional activator. The interactions with cellular splicing factors and potential function in pre-mRNA processing. *J. Biol. Chem.*, **274**, 11832–11841.
23. Bodaghi,S., Jia,R. and Zheng,Z.M. (2009) Human papillomavirus type 16 E2 and E6 are RNA-binding proteins and inhibit in vitro splicing of pre-mRNAs with suboptimal splice sites. *Virology*, **386**, 32–43.
24. Warburton,A., Della Fera,A.N. and McBride,A.A. (2021) Dangerous liaisons: long-term replication with an extrachromosomal HPV genome. *Viruses*, **13**, 1846.
25. Van Doorslaer,K., Tan,Q., Xirasagar,S., Bandaru,S., Gopalan,V., Mohamoud,Y., Huyen,Y. and McBride,A.A. (2013) The papillomavirus episteme: a central resource for papillomavirus sequence data and analysis. *Nucleic Acids Res.*, **41**, D571–D578.
26. Sherman,L. and Alloul,N. (1992) Human papillomavirus type 16 expresses a variety of alternatively spliced mRNAs putatively encoding the E2 protein. *Virology*, **191**, 953–959.
27. Alloul,N. and Sherman,L. (1999) The E2 protein of human papillomavirus type 16 is translated from a variety of differentially spliced polycistronic mRNAs. *J. Gen. Virol.*, **80**, 29–37.
28. Zheng,Y., Cui,X., Nilsson,K., Yu,H., Gong,L., Wu,C. and Schwartz,S. (2020) Efficient production of HPV16 E2 protein from HPV16 late mRNAs spliced from SD880 to SA2709. *Virus Res.*, **285**, 198004.
29. Jia,R. and Zheng,Z.M. (2009) Regulation of bovine papillomavirus type 1 gene expression by RNA processing. *Front. Biosci.*, **14**, 1270–1282.
30. Graham,S.V. and Faizo,A.A. (2017) Control of human papillomavirus gene expression by alternative splicing. *Virus Res.*, **231**, 83–95.
31. Kajitani,N. and Schwartz,S. (2020) Role of viral ribonucleoproteins in human papillomavirus type 16 gene expression. *Viruses*, **12**, 1110.
32. Kajitani,N., Glahder,J., Wu,C., Yu,H., Nilsson,K. and Schwartz,S. (2017) hnRNP L controls HPV16 RNA polyadenylation and splicing in an akt kinase-dependent manner. *Nucleic Acids Res.*, **45**, 9654–9678.
33. Zheng,Y., Jonsson,J., Hao,C., Shoja Chaghervand,S., Cui,X., Kajitani,N., Gong,L., Wu,C. and Schwartz,S. (2020) Heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) and hnRNP A2 inhibit splicing to human papillomavirus 16 splice site SA409 through a UAG-Containing sequence in the E7 coding region. *J. Virol.*, **94**,e01509-20.
34. Rosenberger,S., De-Castro Arce,J., Langbein,L., Steenbergen,R.D. and Rosl,F. (2010) Alternative splicing of human papillomavirus type-16 E6/E6* early mRNA is coupled to EGF signaling via erk1/2 activation. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 7006–7011.
35. Ajiro,M., Tang,S., Doorbar,J. and Zheng,Z.M. (2016) Serine/Arginine-Rich splicing factor 3 and heterogeneous nuclear ribonucleoprotein A1 regulate alternative RNA splicing and gene expression of human papillomavirus 18 through two functionally distinguishable cis elements. *J. Virol.*, **90**, 9138–9152.
36. Cheunim,T., Zhang,J., Milligan,S.G., McPhillips,M.G. and Graham,S.V. (2008) The alternative splicing factor hnRNP A1 is up-regulated during virus-infected epithelial cell differentiation and binds the human papillomavirus type 16 late regulatory element. *Virus Res.*, **131**, 189–198.
37. Wu,C., Kajitani,N. and Schwartz,S. (2017) Splicing and polyadenylation of human papillomavirus type 16 mRNAs. *Int. J. Mol. Sci.*, **18**, 366.
38. Zhao,X., Rush,M. and Schwartz,S. (2004) Identification of an hnRNP A1 dependent splicing silencer in the HPV-16 L1 coding region that prevents premature expression of the late L1 gene. *J. Virol.*, **78**, 10888–10905.
39. Li,X., Johansson,C., Cardoso-Palacios,C., Mossberg,A., Dhanjal,S., Bergvall,M. and Schwartz,S. (2013) Eight nucleotide substitutions inhibit splicing to HPV-16 3′-splice site SA3358 and reduce the efficiency by which HPV-16 increases the life span of primary human keratinocytes. *PLoS One*, **8**, e72776.
40. Li,X., Johansson,C., Glahder,J., Mossberg,A.K. and Schwartz,S. (2013) Suppression of HPV-16 late L1 5′-splice site SD3632 by binding of hnRNP D proteins and hnRNP A2/B1 to upstream AUAGUA RNA motifs. *Nucleic Acids Res.*, **22**, 10488–10508.
41. Yu,H., Gong,L., Wu,C., Nilsson,K., Li-Wang,X. and Schwartz,S. (2018) hnRNP G prevents inclusion on the HPV16 L1 mRNAs of the central exon between splice sites SA3358 and SD3632. *J. Gen. Virol.*, **99**, 328–343.
42. Collier,B., Oberg,D., Zhao,X. and 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.
43. Forslund,O., Sugiyama,N., Wu,C., Ravi,N., Jin,Y., Swoboda,S., Andersson,F., Bzhalava,D., Hultin,E., Paulsson,K. *et al.* (2019) A novel human in vitro papillomavirus type 16 positive tonsil cancer cell line with high sensitivity to radiation and cisplatin. *BMC Cancer*, **19**, 265.
44. Johansson,C., Jamal Fattah,T., Yu,H., Nygren,J., Mossberg,A.K. and Schwartz,S. (2015) Acetylation of intragenic histones on HPV16 correlates with enhanced HPV16 gene expression. *Virology*, **482**, 244–259.
45. Adamson,B., Smogorzewska,A., Sigoillot,F.D., King,R.W. and Elledge,S.J. (2012) A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. *Nat. Cell Biol.*, **14**, 318–328.
46. Nilsson,K., Wu,C., Kajitani,N., Yu,H., Tsimtsirakis,E., Gong,L., Winquist,E.B., Glahder,J., Ekblad,L., Wennerberg,J. *et al.* (2018) The DNA damage response activates HPV16 late gene expression at the level of RNA processing. *Nucleic Acids Res.*, **46**, 5029–5049.
47. Nilsson,K., Wu,C. and Schwartz,S. (2018) Role of the DNA damage response in human papillomavirus RNA splicing and polyadenylation. *Int. J. Mol. Sci.*, **19**, 1735.
48. Doorbar,J., Quint,W., Banks,L., Bravo,I.G., Stoler,M., Broker,T.R. and Stanley,M.A. (2012) The biology and life-cycle of human papillomaviruses. *Vaccine*, **30**, F55–F70.
49. Wu,C., Nilsson,K., Zheng,Y., Ekenstierna,C., Sugiyama,N., Forslund,O., Kajitani,N., Yu,H., Wennerberg,J., Ekblad,L. *et al.* (2019) Short half-life of HPV16 E6 and E7 mRNAs sensitizes HPV16-positive tonsillar cancer cell line HN26 to DNA-damaging drugs. *Int. J. Cancer*, **144**, 297–310.
50. Kanhoush,R., Beenders,B., Perrin,C., Moreau,J., Bellini,M. and Penrad-Mobayed,M. (2010) Novel domains in the hnRNP G/RBMX protein with distinct roles in RNA binding and targeting nascent transcripts. *Nucleus*, **1**, 109–122.
51. Nasim,M.T., Chernova,T.K., Chowdhury,H.M., Yue,B.G. and Eperon,I.C. (2003) HnRNP G and tra2beta: opposite effects on splicing matched by antagonism in RNA binding. *Hum. Mol. Genet.*, **12**, 1337–1348.
52. Moursy,A., Allain,F.H. and Clery,A. (2014) Characterization of the RNA recognition mode of hnRNP G extends its role in SMN2 splicing regulation. *Nucleic Acids Res.*, **42**, 6659–6672.
53. Heinrich,B., Zhang,Z., Raitskin,O., Hiller,M., Benderska,N., Hartmann,A.M., Bracco,L., Elliott,D., Ben-Ari,S., Soreq,H. *et al.* (2009) Heterogeneous nuclear ribonucleoprotein G regulates splice site selection by binding to CC(A/C)-rich regions in pre-mRNA. *J. Biol. Chem.*, **284**, 14303–14315.
54. Liu,N., Zhou,K.I., Parisien,M., Dai,Q., Diatchenko,L. and Pan,T. (2017) N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res.*, **45**, 6051–6063.
55. Zhou,K.I., Shi,H., Lyu,R., Wylder,A.C., Matuszek,Z., Pan,J.N., He,C., Parisien,M. and Pan,T. (2019) Regulation of Co-transcriptional Pre-mRNA splicing by m(6)A through the low-complexity protein hnRNPG. *Mol. Cell*, **76**, 70–81.

56. Chai,N.N., Zhou,H., Hernandez,J., Najmabadi,H., Bhasin,S. and Yen,P.H. (1998) Structure and organization of the RBMY genes on the human y chromosome: transposition and amplification of an ancestral autosomal hnRNPG gene. *Genomics*, **49**, 283–289.
57. Bristol,M.L., Das,D. and Morgan,I.M. (2017) Why human papillomaviruses activate the DNA damage response (DDR) and how cellular and viral replication persists in the presence of DDR signaling. *Viruses*, **9**, 268.
58. Reinson,T., Toots,M., Kadaja,M., Pipitch,R., Allik,M., Ustav,E. and Ustav,M. (2013) Engagement of the ATR-dependent DNA damage response at the human papillomavirus 18 replication centers during the initial amplification. *J. Virol.*, **87**, 951–964.
59. Anacker,D.C. and Moody,C.A. (2017) Modulation of the DNA damage response during the life cycle of human papillomaviruses. *Virus Res.*, **231**, 41–49.
60. Guo,J., Wang,X., Jia,J. and Jia,R. (2020) Underexpression of SRSF3 and its target gene RBMX predicts good prognosis in patients with head and neck cancer. *J. Oral. Sci.*, **62**, 175–179.
61. Schumann,F.L., Bauer,M., Gross,E., Terziev,D., Wienke,A., Wickenhauser,C., Binder,M. and Weber,T. (2021) RBMX protein expression in T-Cell lymphomas predicts chemotherapy response and prognosis. *Cancers (Basel)*, **13**, 4788.
62. Hirschfeld,M., Ouyang,Y.Q., Jaeger,M., Erbes,T., Orlowska-Volk,M., Zur Hausen,A. and Stickeler,E. (2015) HNRNP G and HTRA2-BETA1 regulate estrogen receptor alpha expression with potential impact on endometrial cancer. *BMC Cancer*, **15**, 86.
63. Shin,K.H., Kang,M.K., Kim,R.H., Christensen,R. and Park,N.H. (2006) Heterogeneous nuclear ribonucleoprotein G shows tumor suppressive effect against oral squamous cell carcinoma cells. *Clin. Cancer Res.*, **12**, 3222–3228.
64. Shin,K.H., Kim,R.H., Yu,B., Kang,M.K., Elashoff,D., Christensen,R., Pucar,A. and Park,N.H. (2011) Expression and mutation analysis of heterogeneous nuclear ribonucleoprotein G in human oral cancer. *Oral. Oncol.*, **47**, 1011–1016.
65. Ouyang,Y.Q., zur Hausen,A., Orlowska-Volk,M., Jager,M., Bettendorf,H., Hirschfeld,M., Tong,X.W. and Stickeler,E. (2011) Expression levels of hnRNP G and hTra2-beta1 correlate with opposite outcomes in endometrial cancer biology. *Int. J. Cancer*, **128**, 2010–2019.
66. Zheng,T., Zhou,H., Li,X., Peng,D., Yang,Y., Zeng,Y., Liu,H., Ren,J. and Zhao,Y. (2020) RBMX is required for activation of ATR on repetitive DNAs to maintain genome stability. *Cell Death Differ.*, **27**, 3162–3176.
67. Yan,Q., Zeng,P., Zhou,X., Zhao,X., Chen,R., Qiao,J., Feng,L., Zhu,Z., Zhang,G. and Chen,C. (2021) RBMX suppresses tumorigenicity and progression of bladder cancer by interacting with the hnRNP A1 protein to regulate PKM alternative splicing. *Oncogene*, **40**, 2635–2650.