the FEBS Journal

www.febsjournal.org

A structured RNA in hepatitis B virus post-transcriptional regulatory element represses alternative splicing in a sequence-independent and position-dependent manner

Chen Huang, Mao-Hua Xie, Wei Liu, Bo Yang, Fan Yang, Jingang Huang, Jie Huang, Qijia Wu, Xiang-Dong Fu and Yi Zhang

DOI: 10.1111/j.1742-4658.2010.08077.x

Supplementary data

Construction of the HBV genome library for screening the potential ESS and ISS elements

HBV genomic DNA was released from the pCH-9/3091 (1) plasmid with KpnI and SalI restriction enzymes and purified with a gel purification kit (Omega), which was then fragmented with a cocktail of AluI, BamHI, BglII, DraI, HaeIII, HpaII, XhoI, DpnI and TaqI restriction enzymes, producing either 5' overhang or blunt terminal DNA fragments. The 50-250 bp fractions were then purified by gel electrophoresis and dAMPs were added to the 3' termini of the DNA fragments by Taq polymerase. The A-tailed DNA fragments were later inserted into the T-tailed pZW8MT and pZW8-SMN1C vectors to screen for ESSs and ISSs, respectively, using T4 DNA ligase. The ligation products were transformed into competent DH5α cells. Kanamycin resistant colonies were selected and the plasmid DNA was extracted for transfection. This procedure is depicted schematically in Supplementary Fig. S5A.

Plasmid pZW8 (a kind gift from Dr. Christopher B. Burge, MIT) contains a *SIRT1* alternative splicing cassette interrupting the coding region of a green fluorescent protein (GFP) (2), while pZW8M was derived from pZW8 into which an EcoRV site was introduced between HindIII and KpnI sites of the *SIRT1* exon (see Supplementary Fig. S5A). Plasmid pZW8M was digested by EcoRV and the T-tailed pZW8MT was constructed as previously described (3).

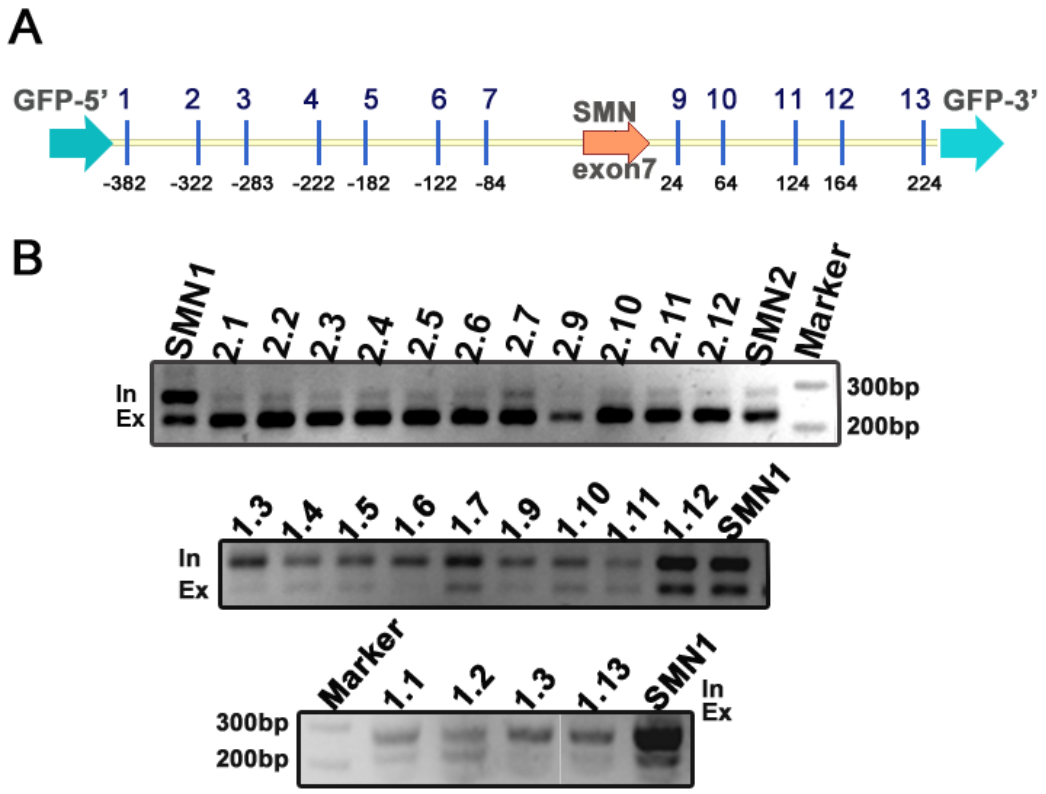
The pZW8-SMN1C (1.6) reporter contains a split *GFP* gene interrupted by *SMN1* splicing cassette (Intron6-Exon7-Intron7) was modified to carry a C insertion upstream of the UAA stop codon in exon7 to avoid the non-sense mediated decay (NMD) of the spliced GFP mRNA including *SMN1* exon7. The T-A cloning strategy as shown in Fig. S5A was used to clone HBV DNA fragments with tailed A nucleotide to produce HBV genomic library in pZW8-SMN1C. The insertion site was at position 6 (See Supplementary Fig. S1A) between the engineered SacI and SalI restriction sites.

Supplementary Table S1. The sequences of the unique ESSs identified from HBV genome to repress the inclusion of the alternative *SIRT1* exon.

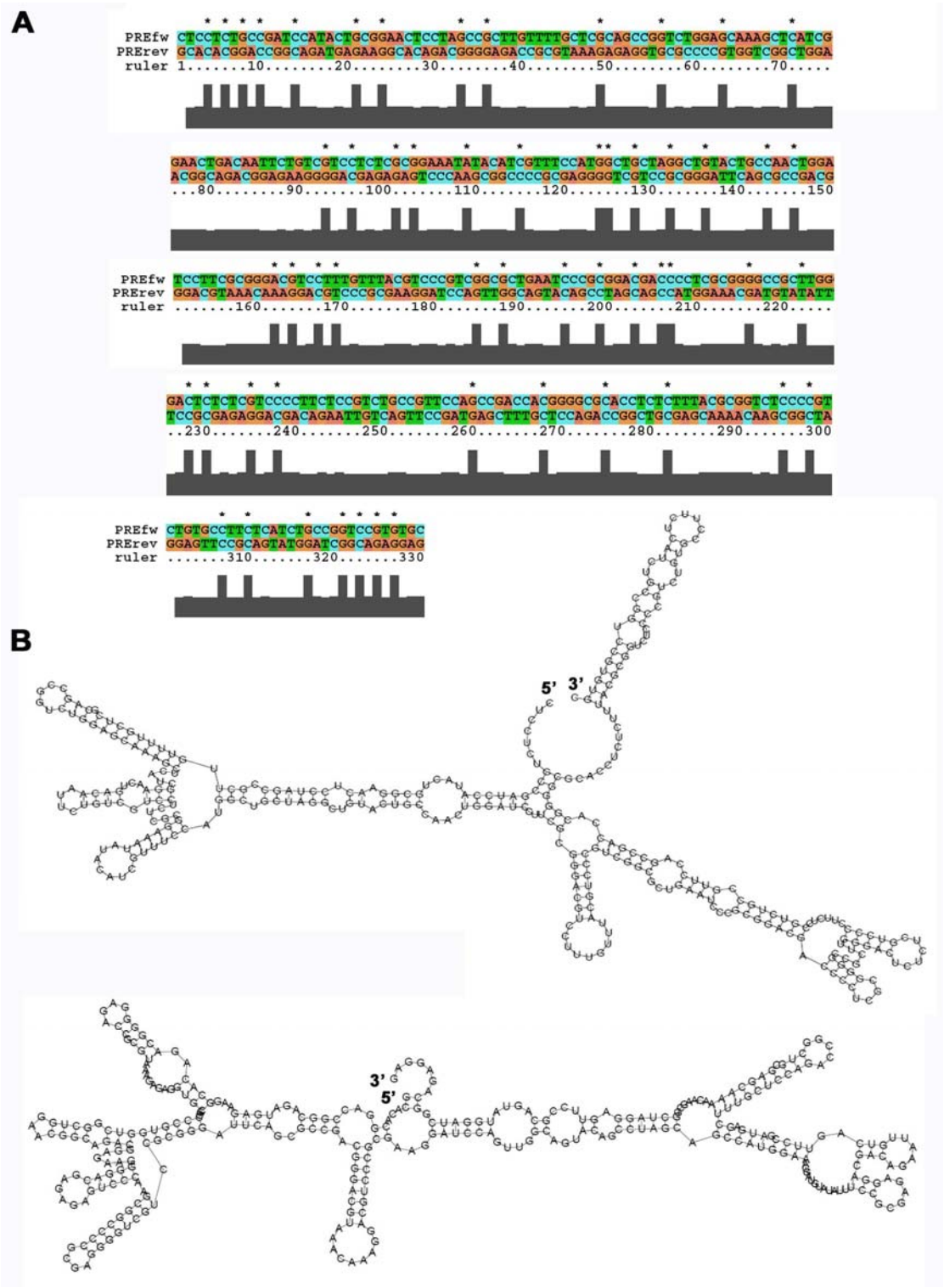
ESS_7	aaagactgggaggagttgggggaggagattaggttaaaggtctttgtactaggaggctgtaggcataaattggtctgcgcaccagcaccatgcaacttttcacctctgcctaatcatctcttgttcatgtcctactgttcaagcctccaagctgtgccttgggtggctttggggcatggacatcg	direct
ESS_14,20,34	ccaagtctgtacagcatcttgagtccctttttaccgctgttaccaattttcttttgtctttgggtatacattt	direct
ESS_33,132	aaaccctaacaaaacaaagagatggggttactctctaaattttatgggttatgtcattggatgttatgggtccttgccacaagaacacatcatacaaaaaatcaaagaatgttttagaaaacttcctattaacagg	direct
ESS_98	cctacaaactgttcacatttttgataatgtcttggtgtaaatgtatattaggaaaagatggtgttttccaatgaggattaaagacaggtacagtagaagaataaagcccagtaaagttccccaccttatgagtccaaggaatactaacattgagattcccgagattgagatc	reverse
ESS_35	cctaaagttcaggcaactcttgtggtttcacatttcttgtctcacttttggaagagaaacagttatagagtatttggtgtctttcggagtgtggattcgcactcctccagcttatagaccaccaaatgcccctatcctatcaacacttccg	direct
ESS_143	ctacctgggtgggtgttaatttggaagatccagcgtctagagacctagtagtcagttatgtcaacactaatatgggacctaaagttcaggcaactcttgtggtttcacatttcttgtctcacttttggaagagaaacagttatagagtatttggtgtctttcggagtgtggattcgcactcctccag	direct
ESS_147	tcttgttcccaagaatatggtgacccacaaaatgaggcgctatgtgttgtttctctcttatataatatacccgccttccatagagtgtgtaaatagtgtctagtttggaagtaatgattaactagatgttctggataataaggtttaataccctttatccaatggtaaatatttggtaacctttggataaaacctggcaggcataatcaattgcaatcttcttttctcattaactgtgagtggg	reverse
ESS_156	cctgttaataggaagttttctaaaacattctttga**ttt**tttgtatgatgtgttcttgtg**gc**aaggacccataacatccaatgacataacccataaaatttagagagtaaccccatctc**t**ttg**t**tttgttagggttt	reverse
ESS_183	cctcagcccgtttctcctggctcagtttactagtgccatttgttcagtggttcgtagggctttcccccactgtttggctttcagttatatggatgatgtggtattggggg	direct
ESS_185	cctattgattggaaagtatgtcaacgaattgtgggtcttttgggttttgctgccccttttacacaatgtggttatcctgcgttgatgcctttgtatgcatgtattcaatctaagcaggctttcactttctcgccaacttacaagg	direct
ESS_69	cgatgtccatgccccaaagccacccaaggcacagcttggaggcttgaacagtaggacatgaacaagagatgattaggcagaggtgaaaaagttgcatggtgctggtgcgcagaccaatttatgcctacagcctcctagtacaaagacctttaacctaatctcctcccccaactcctcccagtcttt	reverse
ESS_155	gatcctgcgcgggacgtcctttgtttacgtcccgtcggcgctgaatcctgcggacgacccttctcgggggtcgcttgggactctctcgtccccttctccgtctgccgttccgaccgaccacggggcgcacctctctttacgcggactccccgtctgtgccttctcatctgccg	direct

Supplementary Table S2. The sequences of the unique ISSs identified from HBV genome to repress the inclusion of the *SMN1* exon 7.

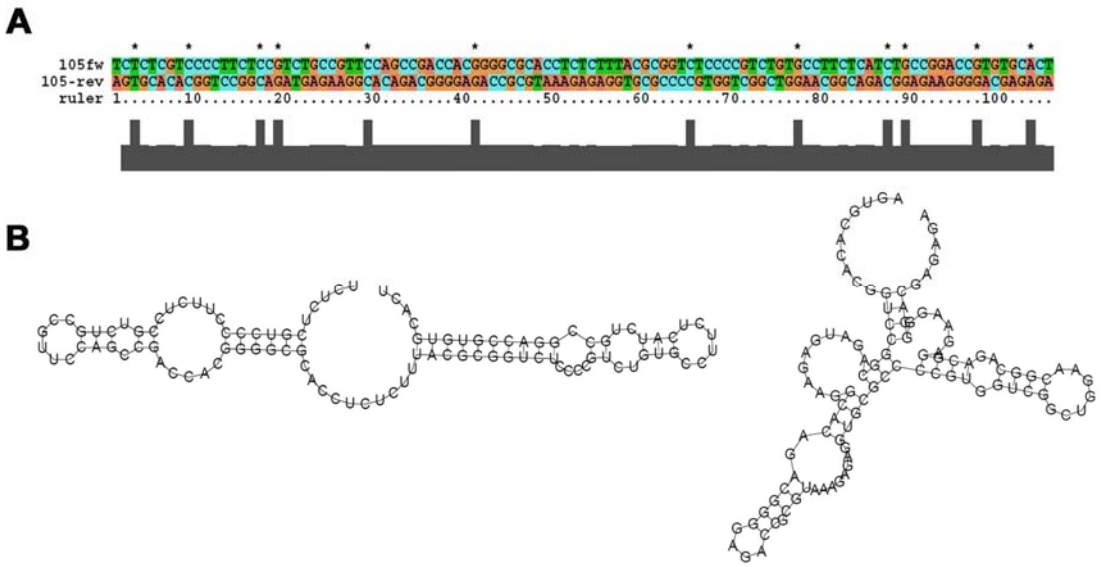
ISS 97,57	gatccagccttcagagcaaccaccgcaaatccagattgggacttcactcccaccaaggacacctgg	direct
ISS 2	gatcttctgcgacgcggcgattgagaccttcgtctgcgaggcgagggagttcttcttctaggggacctgcc tcgtcgtctaacaacagtagtctccga	reverse
ISS 96,117	ccaggtgtccttggtgggagtgaagtcccaatctggatttgcggtggttgctctgaaggctggatc	reverse
ISS 26	aaaccctaacaaaacaaagagatggggttactctctaaattttatgggttatgtcattggatgttatgggt ccttgccacaagaacacatcatacaaaaaatcaaagaatgttttagaaaacttcctattaacagg	direct
ISS 94,101, 38,16	cggaccgtgtgcacttcgcttcacctctgcacgtcgcatggagaccaccgtgaacgcccaccaaatattg cccaaggtcttacataagaggactcttggactctcagcaatgtcaacgaccgaccttgaggcatacttca aagactgtttgttt	direct
ISS 103	ggcagatgagaagggacagacggggagtccgcggaaagagaggtgcgccccgtggtcggtcggaacg gcagacggagaaggggacgagagagtcccaagcgaccccgagaagggtcgtccgcaggattcagcgc cgacgggacgtatacaaaggacgtcccgcgcaggatc	reverse
ISS 95	cctgttaataggaagttttctaaaacattctttgatttttgtatgatgtgttcttgtggcaaggacccataac atccaatgacataacccataaaatttagagagtaaccccatctctttgttttgttagggtttaaatgtatacc caaagacaaaagaaaattggtaacagcggtaaaaagggactcaagatgctgtacagacttgg	reverse
ISS 118	gatcctgcgcgggacgtcctttgtatacgtcccgtcggcgctgaatcctgcggacgacccttctcggggtc gcttgggactctctcgtccccttctccgtctgccgttccgaccgaccacggggcgcacctctctttccgcgg actccccgtctgtcccttctcatctgccg	direct
ISS 54	ccaaaattcgcagtccccaacctccaatcactcaccaacctcttgtcctccaacttgtcctggttatcgctg gatgtgtctgcggcgttttatcatcttcctcttcatcctgctgctatgcctcatcttcttgttggttcttctggac tatcaaggtatgttgcccgtttgtcctctaattccaggatc	direct
ISS 65	cggagactactgttgttagacgacgaggcaggtcccctagaagaagaactccctcgcctcgcagacgaa ggtctcaatcgccgcgtcgcagaagatc	direct



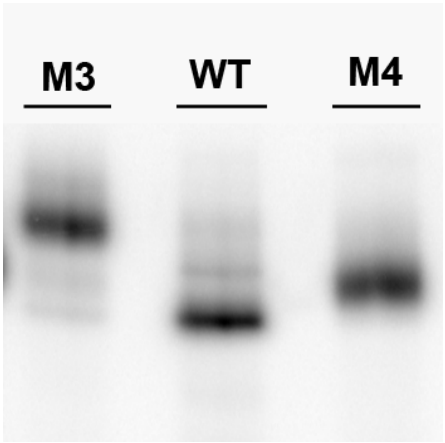
Supplementary Fig. S1. Splicing of *SMN1* and *SMN2* splicing reporter controls. **(A)** Illustration of the pZW8-*SMN1* and pZW8-*SMN2* reporter series with a control sequence [ACCTCAGGCG (4)] inserted at various locations of the upstream and downstream introns of *SMN* exon7. Each insertion site and the corresponding nucleotide position from the 3' or 5' splice site of exon 7 are indicated. **(B)** RT-PCR analysis of the exon7 splicing from the pre-mRNA corresponding to each of the reporter plasmid. The top panel shows the effect of insertions on *SMN2* splicing and the middle and bottom panels show the effect of insertions on *SMN1* splicing. The top and middle panels were results from the same batch of experiments, while the bottom one from a different batch of experiment. The results demonstrate the lack of effect of the control sequence on splicing of either *SMN1* or *SMN2*.



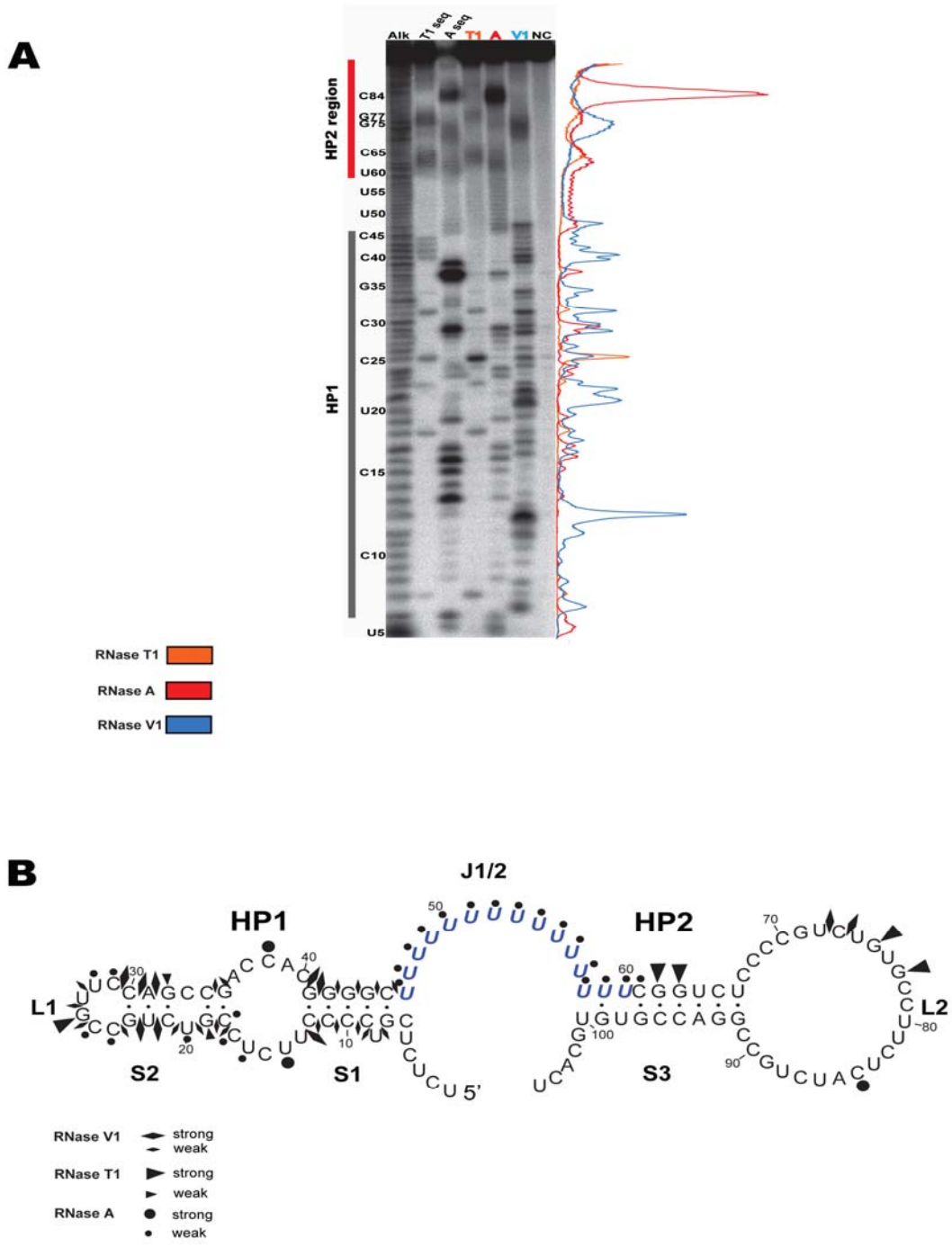
Supplementary Fig. S2. (A) Sequence alignment of the sense (PREfw) and antisense (PRErev) HBV-PRE$_{1253\text{-}1582}$ by ClustalX. **(B)** Secondary structures of the sense (upper) and antisense (lower) HBV-PRE$_{1253\text{-}1582}$ were predicted by the minimum free energy algorithm (5) on the RNAfold web server (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi). The calculated minimum free energy (MFE) is -124.50 kcal/mol for the sense structure, and -132.80 kcal/mol for the antisense structure.



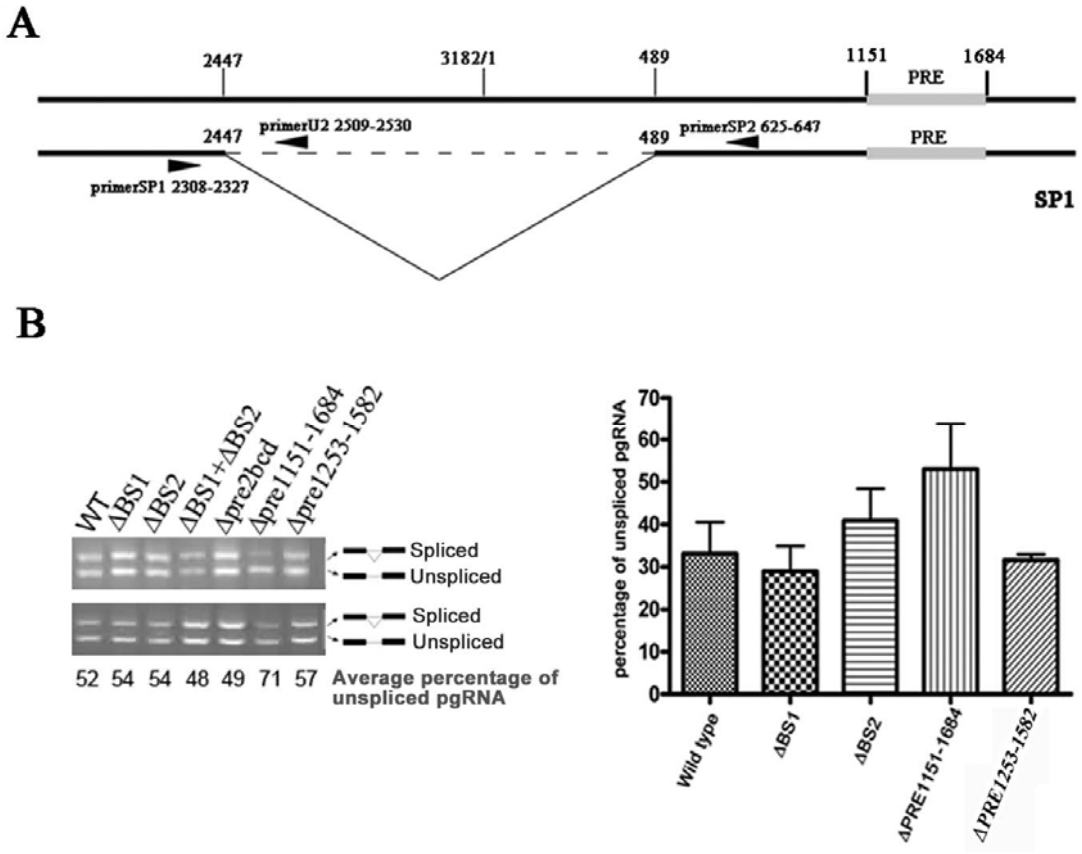
Supplementary Fig. S3. (A) Alignment of the sense and antisense sequences of PRE-ISS (105-nt, $PRE_{1481-1585}$) by ClustalX as in Supplementary Fig. S2. 105fw indicates the sequence of PRE-ISS and 105-rev indicates the antisense sequence. **(B)** Secondary structures of the sense (left) and antisense (right) PRE-ISS were predicted on the RNAfold web server as in Supplementary Fig. S2. The calculated minimum free energy (MFE) is -26.50 kcal/mol for the sense strand structure, and -31.10 kcal/mol for the antisense structure.



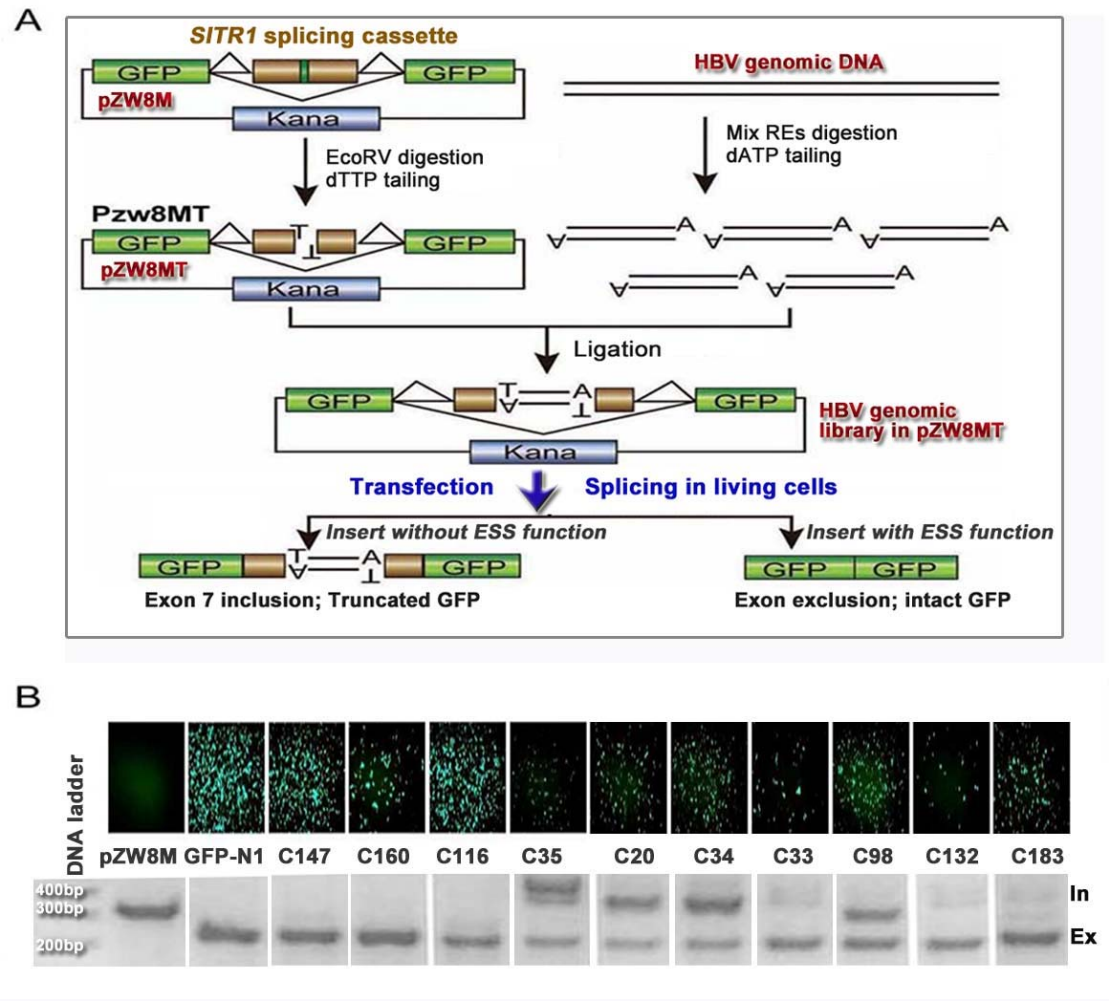
Supplementary Fig. S4. Gel shift of the pre-folded RNA of WT, M3 and M4 on a 5% native PAGE gel. All samples were 5'end-labeled and treated under the same condition as used for the footprinting assay described in Material and Methods.



Supplementary Fig. S5. (A) RNase footprinting analysis of ribonuclease accessibility of the PRE-ISS M4 mutant. The ribonuclease cleavage assay was carried out with RNase T1, V1 and A using end-labeled RNA samples. T1 seq and A seq represent RNase T1 and RNase A sequencing marker, respectively. NC is a negative control using water instead of a RNase. Colored curves represent normalized intensity of RNase cleavage signals. The region covering HP2 in wild type is indicated by a red bar. (**B**) Predicted secondary structure of M4. RNase T1, RNase V1 and RNase A signals are superimposed on the secondary structure of the wild type element. Triangles, diamonds and ellipses represent signals for RNase T1, RNase V1 and RNase A, respectively, and bigger ones represent stronger signals while smaller ones weaker signals.



Supplementary Figure S6. Splicing of the SP1 variant of HBV pgRNA. (A) Schematic diagram of the HBV splicing product SP1 and the primer pairs used to detect the spliced SP1 pgRNA (spliced, primers $SP1_{2308-2327}$ and $SP2_{625-647}$) and unspliced pgRNA (unspliced, primers $SP1_{2308-2327}$ and $U2_{2509-2530}$). **(B)** Analysis of SP1 splicing in different pgRNAs transcribed from the wild-type and mutant HBV genomes in Huh-7 (left) and HepG2-wh (right) cells. The transfected HBV genomes engineered in plasmid pCH9/3091 were indicated on the top (left) of the gel or the bottom (right) of the bar graph. Left, two sets of transfection were performed on Huh-7 cells, which were quantified by semi-quantitative RT-PCR. The average percentage of unspliced pgRNA in relative to the total pgRNA (a sum of the spliced and unspliced) for each corresponding HBV genome is indicated at the bottom. Right, three sets of transfection were performed on HepG2-wh cells, and radioactive RT-PCR was performed to quantify the spliced and unspliced products. The data were plotted using SigmaPlot 11.0 program. The standard deviation (SD) is shown.



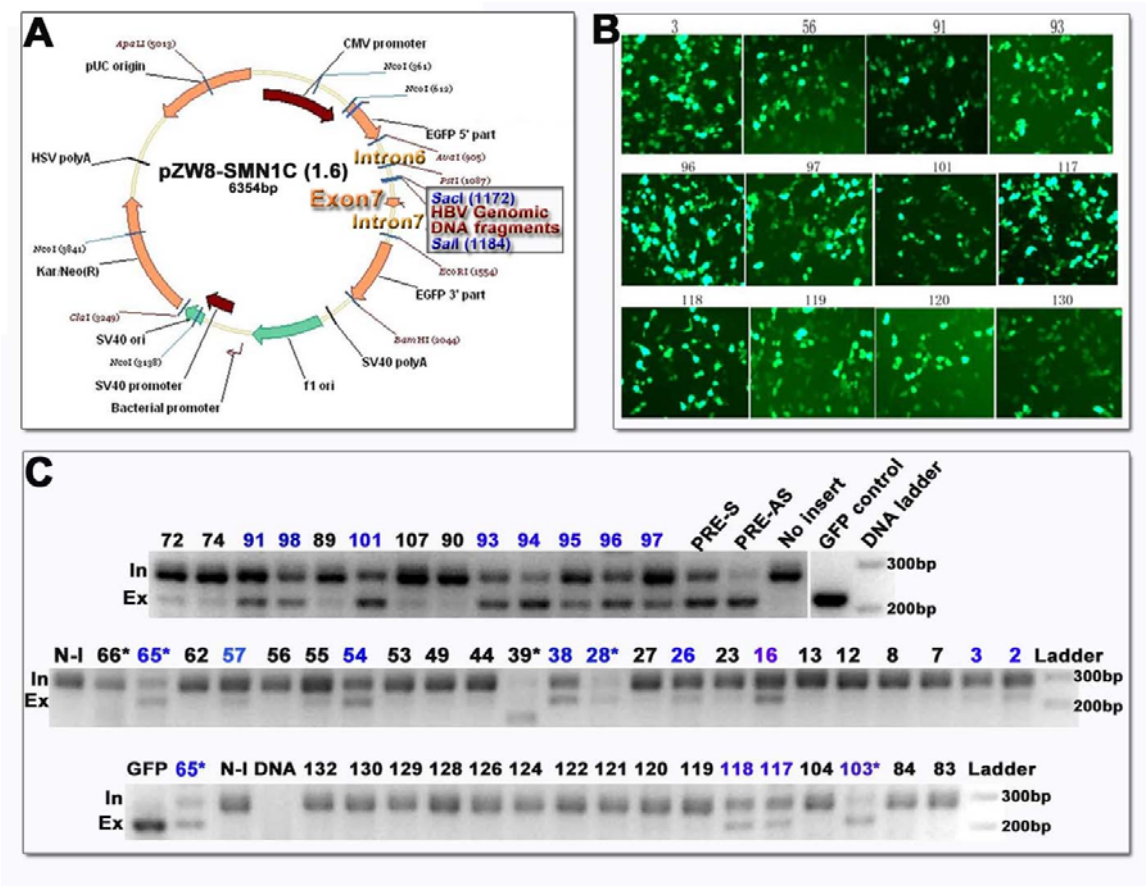
Supplementary Figure S7. The fluorescence-activated screening of exonic splicing silencers (ESSs). **(A)** Diagram of the strategy used to screen for ESSs. The pZW8 reporter containing a split *GFP* gene interrupted by *SIRT1* splicing cassette was modified to carry an EcoRV site in the *SIRT1* exon (pZW8M). pZW8MT contains a T nucleotide overhang. HBV genomic library was constructed by cloning HBV DNA fragments with tailed A nucleotide into pZW8T. **(B)** Representative results of the selection, with the upper panel showing the results for fluorescence microscopy and the lower panel for RT-PCR analysis of the splicing products. "Ex"-splicing product with *SIRT1* exon being excluded. "In"-splicing product with *SIRT1* exon inclusion. Plasmid DNA from a total of 150 single colonies that contain HBV genomic DNA inserts were isolated and transiently transfected into cultured Linx cells in 96-well plates. The insert free pZW8M plasmid was used as a negative control, which generates the *SIRT1* exon-included GFP mRNA that produced no functional GFP .

Selection procedure:

1. The GFP signals indicating an ESS activity of the HBV genomic insert in the *SIRT1* exon were examined under a fluorescence microscopy (Leica, Germany). The plasmid clones used to transfect those GFP signal-producing wells were

identified.

2. Those candidate ESE-containing plasmids were selected to transfect cultured Linx cells in 24-well plates, from which total RNA was extracted with the Trizol reagent (Invitrogen) according to the manufacturer's instructions. cDNA was generated by incubation of 3 μg of total RNA with random primers and M-MLV reverse transcriptase (Promega). PCR analysis of the alternative splicing of the reporter system was performed in a single reaction of 25 μl volume. GFP3R and GFP1F primers as described in the text were used to amplify the splicing products.
3. Panel B shows that all the plasmid clones with activated GFP signal also produce intact *GFP* mRNA with the insert-containing *SIRT1* exon successfully excluded, although the level of the intact *GFP* mRNA differed among those clones. Generally, the intensity of the GFP signal generated by each plasmid clone positively correlates with the level of the intact *GFP* mRNA.
4. Only the plasmid clones which resulted in both the functional GFP protein and the production of the intact *GFP* mRNA due to splicing repression were selected for sequencing analysis.
5. Only the plasmid clones which contain a single HBV genomic fragment were considered as the source of an ESS sequence, while those containing 2 or more ligated fragments were discarded. A total of 12 unique ESSs were identified (Supplementary Table S1) for further analysis.



Supplementary Figure S8. Fluorescence-based screening of intronic splicing silencers (ISSs). **(A)** Diagram of the strategy used to screen for ISSs. The pZW8-SMN1C (1.6) reporter contains a split *GFP* gene interrupted by the *SMN1* splicing cassette (Intron6-Exon7-Intron7). We inserted a C in position 6 in exon7 to eliminate potential NMD, which may interfere with the calculation of the splicing efficiency based on the ratio of spliced and unspliced mRNA. The T-A cloning strategy as shown in Fig. S5 was used to clone HBV DNA fragments with tailed A nucleotide to produce HBV genomic library in pZW8-SMN1C. **(B-C)** Plasmid DNA from a total of about 150 single colonies that contain HBV genomic DNA inserts were isolated and transiently transfected into cultured Linx cells in 96 well plates. Many plasmid clones resulted in activated expression of GFP protein. **(B)** Plasmids from over 50 such clones were then used to re-transfect the Linx cells, and total mRNA was prepared for RT-PCR analysis of the splicing products as in Fig. S5 **(C)** Expression of about half of the plasmids in Linx cells resulted in the expected intact GFP mRNA band (the clone number in blue), while the other half did not produce such band. Only the plasmids producing detectable exon-skipping band (intact GFP mRNA) in the RT-PCR analysis were considered containing ISS elements.

The following selection procedure was similar to that in Fig.5S, and a total of 10 unique ISSs (Supplementary Table S1) were identified and subjected to further analysis.

References

1. Heise, T., Sommer, G., Reumann, K., Meyer, I., Will, H. and Schaal, H. (2006) The hepatitis B virus PRE contains a splicing regulatory element. *Nucleic acids research*, **34**, 353-363.
2. Wang, Z., Rolish, M.E., Yeo, G., Tung, V., Mawson, M. and Burge, C.B. (2004) Systematic identification and analysis of exonic splicing silencers. *Cell*, **119**, 831-845.
3. Marchuk, D., Drumm, M., Saulino, A. and Collins, F.S. (1991) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic acids research*, **19**, 1154.
4. Fairbrother, W.G., Yeh, R.F., Sharp, P.A. and Burge, C.B. (2002) Predictive identification of exonic splicing enhancers in human genes. *Science*, **297**, 1007-1013.
5. Zuker, M. and Stiegler, P. (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res*, **9**, 133-148.