

## Supporting Information

# Controlling Mammalian Gene Expression by Allosteric Hepatitis Delta Virus Ribozymes

Yoko Nomura, Linlin Zhou, Anh Miu, and Yohei Yokobayashi\*

## Plasmid Sequences

### pEGFP-wtHDVRz

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCGCGGTT  
ACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCGCCATTGACGTCAA  
TAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTA  
TTTACGGTAAACTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATT  
GACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCAGTACATGACCTTATGGGACTTTC  
CMV CTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTA  
promoter CATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATGACGTC  
AATGGGAGTTTTGTTTTGGCACAAAATCAACGGGACTTTCCAAAATGTCGTAACAACCTCCGCC  
CATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTTTTAGT  
GAACCGTCAGATCCGCTAGCGCTACCGGACTCAGATCTCGAGCTCAAGCTTCGAATTCTGCAGT  
CGACGGTACCGCGGGCCCGGATCCACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTT  
CACCGGGGTGGTGGCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTG  
TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCG  
GCAAGCTGCCCGTGCCCTGGCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAG  
CCGCTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTC  
CAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCC  
EGFP AGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACAT  
CCTGGGGCACAAGCTGGAGTACAACACAACAGCCACAACGTCTATATCATGGCCGACAAGCAG  
AAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCG  
CCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCGTGCTGCTGCCCGACAACCACTA  
CCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTG

GAGTTCGTGACCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGCCGCG  
 wtHDV AC**TCTAGA**acaaacaaagaagac<sup>+1</sup>ggccggcatggtcccagcctcctcgctggcgccggtgg  
 ribozyme gcaacattccgaggggaccgtcccctcggaatg<sup>+75</sup>ggaatgggacgcacaaatctctctagctt  
 (-2 to +152) cccagagagaagcgagagaaaagtggctctcccttgccatccgagtgg**TCCGGA**GCCATACCA  
 CATTGTAGAGGTTTACTTGCTTTAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAA  
 AATGAATGCAATTGTTGTTGTTAACTTGTATTATTGCAGCTTATAATGGTTACAAATAAAGCAAT  
 AGCATCACAAATTTACAAATAAAGCATTTTTTTCTACTGCATTCTAGTTGTGGTTTGTCCAAAC  
 TCATCAATGTATCTTAAGGCGTAAATTGTAAGCGTTAATATTTTGTAAAATTCGCGTTAAATT  
 TTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAA  
 AGAATAGACCGAGATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGAGTCCACTATTAAAGAAC  
 fl ori GTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCACTACGTGAACCAT  
 CACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG  
 CCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCG  
 AAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACACCCG  
 CCGCGCTTAATGCGCCGCTACAGGGCGCGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAAC  
 CCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGA  
 TAAATGCTTCAATAATATTGAAAAAGGAAGAGTCCCTGAGGCGGAAAGAACCAGCTGTGGAATGT  
 GTGTCAGTTAGGGTGTGGAAGTCCCAGGCTCCCAGCAGGCAGAAGTATGCAAAGCATGCAT  
 CTCAATTAGTCAGCAACCAGGTGTGGAAGTCCCAGGCTCCCAGCAGGCAGAAGTATGCAAA  
 GCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAAC  
 TCCGCCCAGTTCCGCCCATTTCTCCGCCCATGGCTGACTAATTTTTTTTTATTTATGCAGAGGCC  
 GAGGCCGCCTCGGCCTCTGAGCTATTCAGAAGTAGTGAGGAGGCTTTTTTTGGAGGCCTAGGCT  
 TTTGCAAAGATCGATCAAGAGACAGGATGAGGATCGTTTTCGCATGATTGAACAAGATGGATTGC  
 ACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAAT  
 CGGCTGCTCTGATGCCCGCTGTTCCGGCTGTCAGCGCAGGGGCGCCCGTCTTTTTGTCAAG  
 ACCGACCTGTCCGGTGCCCTGAATGAACTGCAAGACGAGGCAGCGCGGCTATCGTGGCTGGCCA  
 CGACGGGCGTTCCCTTGCGCAGCTGTGCTCGACGTTGTCCTGAAGCGGGAAGGGACTGGCTGCT  
 ATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCC  
 ATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACC  
 AAGCGAAACATCGCATCGAGCGAGCACGTAICTGGATGGAAGCCGGTCTTGTGATCAGGATGA  
 TCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCCGAGGCTCAAGGCGAGCATG  
 CCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAA  
 ATGGCCGCTTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACAT  
 AGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCGTG

CTTTACGGTATCGCCGCTCCCGATTTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCT  
TCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCAACCTGCCATCACGAGAT  
TTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCT  
GGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCACCCTAGGGGAGGCTAAC  
TGAAACACGGAAGGAGACAATACCGGAAGGAACCCGCGCTATGACGGCAATAAAAAGACAGAAT  
AAAACGCACGGTGTGGGTCGTTTGTTCATAAACGCGGGTTCGGTCCCAGGGCTGGCACTCTG  
TCGATACCCACCGAGACCCCATTTGGGGCAATACGCCCGCGTTTCTTCCTTTTCCCCACCCCA  
CCCCCAAGTTCGGGTGAAGGCCAGGGCTCGCAGCCAACGTCGGGGCGGCAGGCCCTGCCATA  
GCCTCAGGTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAAGGATCT  
AGGTGAAGATCCTTTTTGATAATCTCATGACCAAATCCCTAACGTGAGTTTTCGTTCCACTG  
AGCGTCAGACCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC  
TGCTGCTTGCAAACAAAAAACACCGCTACCAGCGGTGGTTTGTGGCCGGATCAAGAGCTAC  
CAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAATACTGTCCTTCTAGT  
GTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTA  
ATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGAC  
GATAGTTACCGGATAAGGCGCAGCGGTCTGGGCTGAACGGGGGTTCTGTCACACAGCCCAGCTT  
GGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTT  
CCCGAAGGGAGAAAGGCGGACAGGTATCCGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGA  
GGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCCTGTGGGTTTCGCCACCTCTGACT  
TGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACCGC  
GCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTGCGTTATCCC  
CTGATTCTGTGGATAACCGTATTACCGCCATGCAT (4899 bp)

pUC ori

**pEGFP-wtHDVRz (inactive)**

Same as pEGFP-wtHDVRz above except for a point mutation in the HDV ribozyme sequence C(+75) → T.

**pEGFP-HDVRz (aptazymes)**

All EGFP-aptazyme constructs shown in Figures 2 and 3 are identical to pEGFP-wtHDVRz except for: 1) removal of “ACAAACAAAGAAGAC” just upstream of the HDV ribozyme for cloning purposes (indicated by double-underline in the above sequence); and 2) replacement of P4-L4 in wtHDVRz (boxed in Fig. 1 and underlined in the above sequence) with the aptamer-connector sequences depicted in Fig. 2a and Fig. 3a.

### **pEGFP-Theo6HDVRz-GuaM8HDV**

Same as pEGFP-wtHDVRz except for replacing the sequence shown in lower case above with the sequence shown below:

```
at+1ggccggcatggtcccagcctcctcgctggcgccggctgggcaaccacataaccagccgaaagg  
cccttggcaggtgggccaatgggacgcacaaatctctctagcttcccagagagaagcgagagaa  
aagtggctctcattacataccaaaacaacacacactacacatacatatacataaacaatcct  
tggccatccgagtggtccggtagaat+1ggccggcatggtcccagcctcctcgctggcgccggctg  
ggcaatgctataatcgcgtaggatggcacgcaagtcttctaccgggcaccgtaaatgtccgact  
agtagcgaatgggacgcacaaatctctctag
```

Red: Theo6HDV; blue: GuaM8HDV; yellow highlight: aptamers

### **pSEAP-GuaM8HDVRz**

Identical to pEGFP-GuaM8HDVRz except for the replacement of the EGFP coding sequence with the SEAP coding sequence shown below.

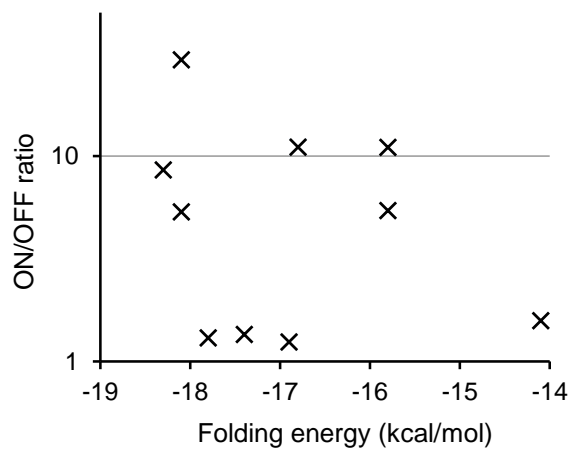
```
ATGCTGCTGCTGCTGCTGCTGCTGGGCCTGAGGCTACAGCTCTCCCTGGGCATCATCCCAGTTG  
AGGAGGAGAACCCGGACTTCTGGAACCGCGAGGCAGCCGAGGCCCTGGGTGCCGCCAAGAAGCT  
GCAGCCTGCACAGACAGCCGCCAAGAACCTCATCATCTTCCTGGGCGATGGGATGGGGGTGTCT  
ACGGTGACAGCTGCCAGGATCCTAAAAGGGCAGAAGAAGGACAAACTGGGGCCTGAGATACCC  
TGGCCATGGACCGCTTCCCATATGTGGCTCTGTCCAAGACATAACAATGTAGACAAACATGTGCC  
AGACAGTGGAGCCACAGCCACGGCCTACCTGTGCGGGGTCAAGGGCAACTTCCAGACCATTGGC  
TTGAGTGCAGCCCGCCGCTTTAACCAGTGCAACACGACACGCGGCAACGAGGTCATCTCCGTGA  
TGAATCGGGCCAAGAAAGCAGGGAAGTCAGTGGGAGTGGTAACCACCACACGAGTGCAGCACGC  
CTCGCCAGCCGGCACCTACGCCACACGGTGAACCGCAACTGGTACTCGGACGCCGACGTGCCT  
GCCTCGGCCCGCCAGGAGGGGTGCCAGGACATCGCTACGCAGCTCATCTCCAACATGGACATTG  
ACGTGATCCTAGGTGGAGGCCGAAAGTACATGTTTCGCATGGGAACCCAGACCCTGAGTACCC  
AGATGACTACAGCCAAGGTGGGACCAGGCTGGACGGGAAGAATCTGGTGCAGGAATGGCTGGCG  
AAGCGCCAGGGTGCCCGGTATGTGTGGAACCGCACTGAGCTCATGCAGGCTTCCCTGGACCCGT  
CTGTGACCCATCTCATGGGTCTCTTTGAGCCTGGAGACATGAAATACGAGATCCACCGAGACTC  
CACACTGGACCCCTCCCTGATGGAGATGACAGAGGCTGCCCTGCGCCTGCTGAGCAGGAACCC  
CGCGGCTTCTTCCCTCTTCGTGGAGGGTGGTCGCATCGACCATGGTCATCATGAAAGCAGGGCTT  
ACCGGGCACTGACTGAGACGATCATGTTTCGACGACGCCATTGAGAGGGCGGGCCAGCTCACCAG  
CGAGGAGGACACGCTGAGCCTCGTCACTGCCGACCACTCCCACGTCTTCTCCTTCGGAGGCTAC  
CCCCTGCGAGGGAGCTCCATCTTCGGGCTGGCCCCCTGGCAAGGCCCGGGACAGGAAGGCCTACA  
CGGTCCTCCTATACGAAACGGTCCAGGCTATGTGCTCAAGGACGGCGCCCGGCCGGATGTTAC
```

CGAGAGCGAGAGCGGGAGCCCCGAGTATCGGCAGCAGTCAGCAGTGCCCCCTGGACGAAGAGACC  
CACGCAGGCGAGGACGTGGCGGTGTTCGCGCGCGGCCCGCAGGCGCACCTGGTTCACGGCGTGC  
AGGAGCAGACCTTCATAGCGCACGTCATGGCCTTCGCCGCCTGCCTGGAGCCCTACACCGCCTG  
CGACCTGGCGCCCCCGCCGGCACCCGACGCCGCGCACCCGGGTTACTCTAGAGTCGGGGCG  
GCCGGCCGCTTCGAGCAGACATGA

**pEGFP-N1-BspEI ("Empty vector")**

Same as pEGFP-wtHDVRz above except for replacing the sequence between the two unique restriction sites XbaI and BspEI (flanking the wt HDV ribozyme sequence and shown in bold) with TTTTTT. This vector was used as a host to clone the above constructs using these restriction sites.

## Supporting Figure



**Figure S1.** Guanine-responsive switch performance (ON/OFF ratio in log scale) and the mfold-calculated folding energies of the connector-aptamer sequences. 5' A-(5'connector)-guanine aptamer-(3'connector)-G 3' was used to calculate the folding energy using the mfold web server <http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>.