Supplementary information.

Materials and Methods.

DNA constructs.

Summaries of the DLS mutants (Fig. S1) and the primers to introduce mutations (Table S1) was shown. Three series of the mutants were constructed, pDDNd2 derivatives, pNLAI43 derivatives, and pNLAINh derivatives. pNLAI43 was constructed by the exchange of 2.7kbp SalI-HpaI region of pNLAINh to corresponding region of pNL4-3.

pDDNd2 derivatives were constructed as described below. First, d2 fragment containing plasmid (pGEMLp4Δ2sub (26))was applied as a template of PCR. For the construction of the mutants containing mutation(s) at 3' end of d2 fragment, the sense primer A3F+N and the antisense primer to introduce mutation(s) were applied for PCR. The amplified fragment was isolated, purified and ligated in pGEM-Teasy (Promega, Madison, WI). For the construction of the mutants containing mutation(s) at 5' end of d2 fragment, the sense primer to introduce mutation the antisense primer SL4Rs were applied for PCR. The amplified fragments were isolated, purified and ligated in pGEM-Teasy. For the construction of the mutants containing mutations both at 3' and 5' end of d2 fragment, the sense and antisense primers to introduce mutations were applied for PCR. The amplified fragments were isolated, purified and ligated in pGEM-Teasy. For the construction of the mutants containing mutation(s) at the center of d2 fragment, two step construction of the fragment by PCR were applied. First two pairs of primers were used for amplification of 5' half and 3' half of the fragment. Antisense primer for 5' half and sense primer for 3' half contained mutation at the same position. After isolation, two fragments were mixed and applied as templates of PCR by primer pairs which annealed to 5' and 3' ends of d2 fragment. The amplified fragments were isolated, purified and ligated in pGEM-Teasy. Digestion of these plasmids with EcoRI, blunt-ended by T4 DNA polymerase (T4p), resulted in the isolation of approximately 0.15kbp fragments, including the 5’ leader region of HIV-1. These fragments were then ligated into the T4p-treated NheI site of pNL4-3 to construct pDDNd2 derivatives.

pNLAINh derivatives were constructed as described below. The plasmid pNLdSNdSN, which contains one 5'LTR and first 0.7kbp of gag gene, was used for template of PCR. pNL4-3 was digested with SpeI and NcoI, blunt ended with T4p and resealed to construct pNLdSN. pNLdSN was digested with StuI and NdeI, blunt ended with T4p and resealed to construct pNLdSNdSN. For the construction of the mutants, S4m1, S4m2 or S4m3, two-step construction of the fragment by PCR was applied. First, two pairs of primers were used for

amplification of 5' half and 3' half of the fragment, and pNLdSNdSN was used as template for both PCR reactions. The sense primer SL1F and antisense primers (SL4RsGCCTC, SL4RsGCTTC, SL4RsGCGAG) for 5' half, and sense primer (SL4FlongGCCTCG, SL4FlongGCTTCG, SL4FlongGCGAGT) and antisense primer MASgrAIR were used for 3' half. After first PCR, two fragments were mixed and applied as templates of second PCR, and SL1F and MASgrAIR, which annealed to 5' and 3' ends of two first-PCR fragments, were used for primers. The amplified fragments were isolated, purified and ligated in pGEM-Teasy (Promega, Madison, WI) and the sequences of the PCR fragments were confirmed by sequencing. Digestion of these plasmids with BssHII and SgrAI resulted in the isolation of 0.1kbp fragments, including the 5'LTR and 5'-small portion of gag, carried one mutation, respectively. These fragments were then ligated into the BssHII and SgrAI digested-pNLAINh to construct pNLAIS4m1Nh, pNLAIS4m2Nh or pNLAIS4m3Nh, respectively.

pNLAI43 derivatives were constructed as described below. For the construction of the mutants, C236G or C236A, two-step construction of the fragment by PCR was applied. First, two pairs of primers were used for amplification of 5' half and 3' half of the fragment, and pNLdSNdSN was used as template for both PCR reactions. The sense primer TAR-F and antisense primers (U5L-SL1-F-GCcTCG or U5L-SL1-F-GCtTCG) for 5' half, and sense primer (U5L-SL1-F-GCcTCG or U5L-SL1-F-GCtTCG) and antisense primer SL4Rs were used for 3' half. After first PCR, two fragments were mixed and applied as templates of second PCR, and TAR-F and SL4Rs, which annealed to 5' and 3' ends of d2 fragment, were used for primers. The amplified fragments were isolated, purified and ligated in pGEM-Teasy (Promega, Madison, WI). Digestion of these plasmids with HindIII and BssHII resulted in the isolation of approximately 0.18kbp fragments, including the 5' leader region of HIV-1 and carrying one mutation, respectively. These fragments were then ligated into the HindIII and BssHII digested pNLdSNdSN to construct pNLdSNdSN-C236G or pNLdSNdSN-C236A. pNLdSNdSN-C236G or pNLdSNdSN-C236A were digested with AatII and BssHII, and the approximately 1.6kbp fragments were ligated into the pNLAI-43 or pNLAIS4m1 or pNLAIS4m2, resulted to pNLAI-S4m1, -S4m2, -S4m3, -C236G, -C236A, -S4m1PK, -S4m2PK, respectively.

SHAPE analysis.

<Transfection and preparation of genomic RNAs out of HIV-1 viruses>

Viral particles of HIV-1 NLAINh or its mutant were produced by the transfection of 293T cells with the pNLAINh, pNLAIC236G, pNLAIS4m1PK or pNLAIS4m1, with or without pCAG-VSV-G plasmids by means of the calcium

phosphate precipitation method (29). $7x10^6$ 293T cells per a dish of 150mm in diameter. 6-12hr after incubation, 30 plates of seeded at a density of, were transfected with plasmids, 6hr later. Cultures were then incubated for 3 days and supernatants were harvested, and treated with DNaseI as described elsewhere (56). Clarified culture medium with centrifugation were ultra-centrifuged at 30000rpm at4C for 1hr to collect viral particle. Genomic RNAs of HIV-1 particle were prepared as described previously, and were quantified by one Step SYBR PrimeScript PLUS RT-PCR kit (RT-PCR; TAKARA Bio).

<RNA folding and NMIA treatment>

0.5pmoles of genomic RNAs were precipitated with ethanol, and the pelleted RNAs were resuspended with 1x TE, and precipitated with ethanol to remove misplaced materials, such as SDS, and so on. After centrifuge, washed pellet with 70 % ethanol was dissolved with 14.4μl of 0.5x TE and divided into two 0.2ml tubes. 1.8μl of folding buffer (250mM Hepes [pH 8.0], 1M potassium acetate [pH8.0], 25mM $MgCl_2$) were added to each tubes and incubated at 37°C for 10 min. 1μl of 32mM NMIA in DMSO or DMSO was added to each tube. After incubation at 37°C for 30 min or 60 min, RNAs were precipitated with ethanol. The pelleted RNAs were suspended with 10μl of 1x TE, and then heated at 95°C for 3 min and chilled at 4°C for 2 min.

<Primer Extension>

1.5μl of 0.3μM of NED-labeled primer (905R UNI; CTGCTTGCCCATACTA, Applied Biosystems) was added to each tubes and incubated at 65°C for 5 min and then at 35°C for 10min to anneal the primers. 2μl of 5x First-Strand Synthesis Buffer (Invitrogen), 0.5μl of 0.1M of DTT (Invitrogen) and 0.65μl of dNTP-mix (7.7 mM each of dATP, dCTP, dTTP(PCR Grade; Roche), and 7-Deaza-dGTP (Jena Bioscience). After incubation at 35°C for 1min, 0.5μl of SuperScript III Reverse Transcriptase (Invitrogen; 200 U/μl) was added to each tubes, and incubated at 45°C for 1min, at 52°C for 25min, and 65°C for 5min, and then kept on ice for 2 min. 0.5μl of 200mM EDTA (pH 8.0) was added to stop the synthesis of DNAs and the precipitated with ethanol. Washed pelleted DNAs with 70 % ethanol was dissolved with 10μl of HiDi formamide (Applied Biosystems) supplemented with 0.5μl of GeneScan600LIZ Size Standard (Applied Biosystems). cDNA samples were denatured at 72°C for 2min, chilled at 4°C for 2min and then separated on 3130xl Genetic Analyzer (Applied Biosystems) by using GeneMapper module.

<Sequencing reactions>

Sequencing reactions to identify peaks in the (+) or (-) reagent experiments were performed by Thermo Sequenase Fluorescent Labelled Primer Cycle Sequenasing Kit with 7-Deaza-dGTP (GE Healthcare). pNLAINh was used for a template DNA with NED-labeled 905R-UNI primer. After PCR, each reaction was precipitated with ethanol separately, washed with 70% ethanol, and then dissolved with 10ul of HiDi formamide. 0.5μl of GeneScan600LIZ Size Standard was added to each sample, denatured at 72°C for 2min and chilled at 4°C for 2min then separated on 3130XL Genetic Analyzer by using GeneMapper module.

<Data processing> The obtained fragment analysis data (FSA) were exported to text file format data by DataFileConverter software (Applied Biosystems). The peaks of LIZ fluorescence were identified with PeakScanner software (Applied Biosystems) and utilized for normalization of each data. The NED fluorescence data of NMIA(+) and (-) were aligned and assembled with the sequence data with ShapeFinder (57). SHAPE value were calculated by subtracting the signal intensity of (-) from that of (+) at each base position, in succession to dividing the count by the average count of most reactive peaks. The standard of the reactive peak selection was based on the previous report (58).

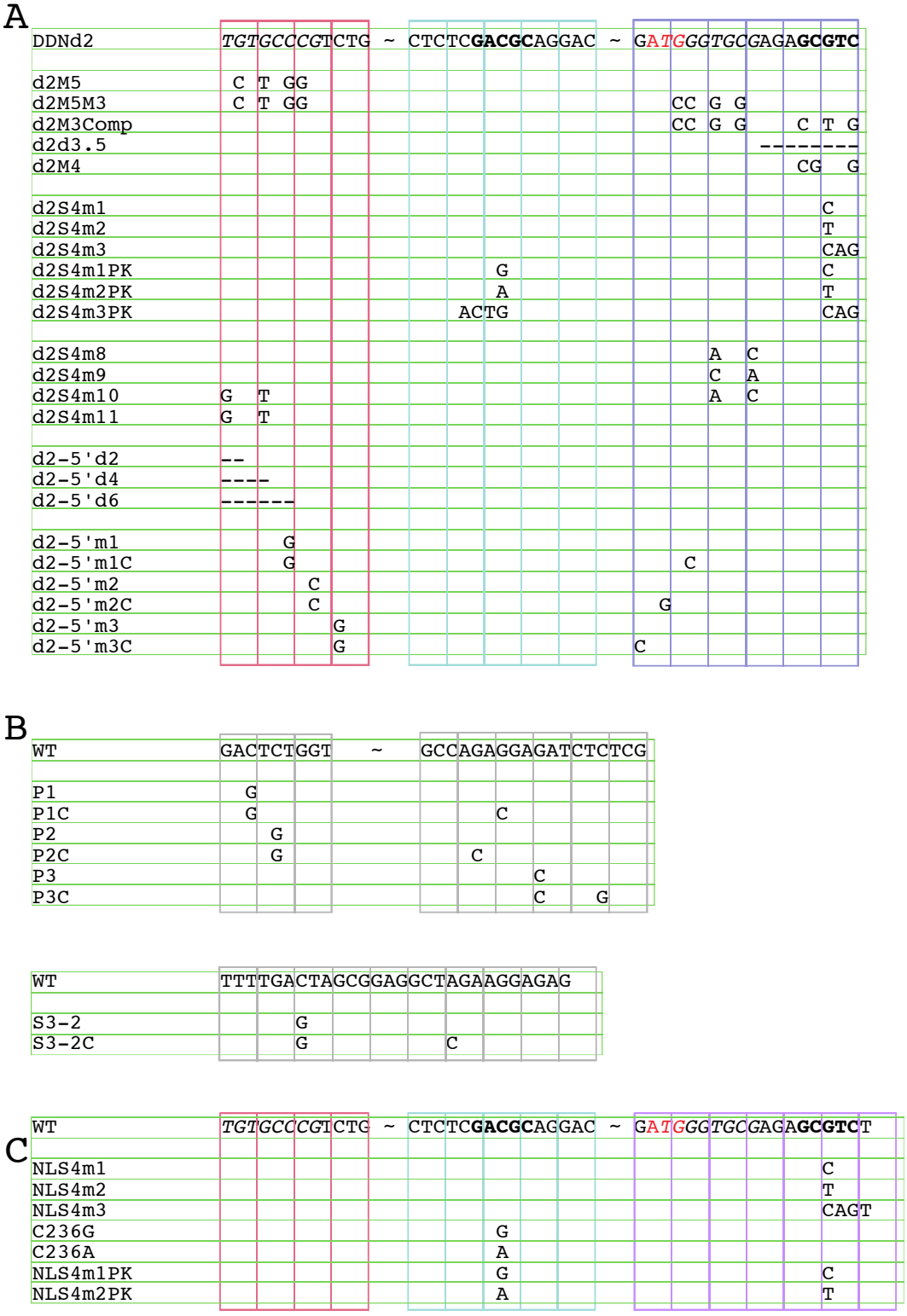
Supplementary Figure legends

Fig.S1 Summaries of mutants constructed. A) DDNd2 derivatives with 5' and/or 3' mutations and the pseudoknot mutants. B) DDNd2 derivatives with PBS stem and SL3 stem mutations. C) NLAINh and NLAI43 derivatives.

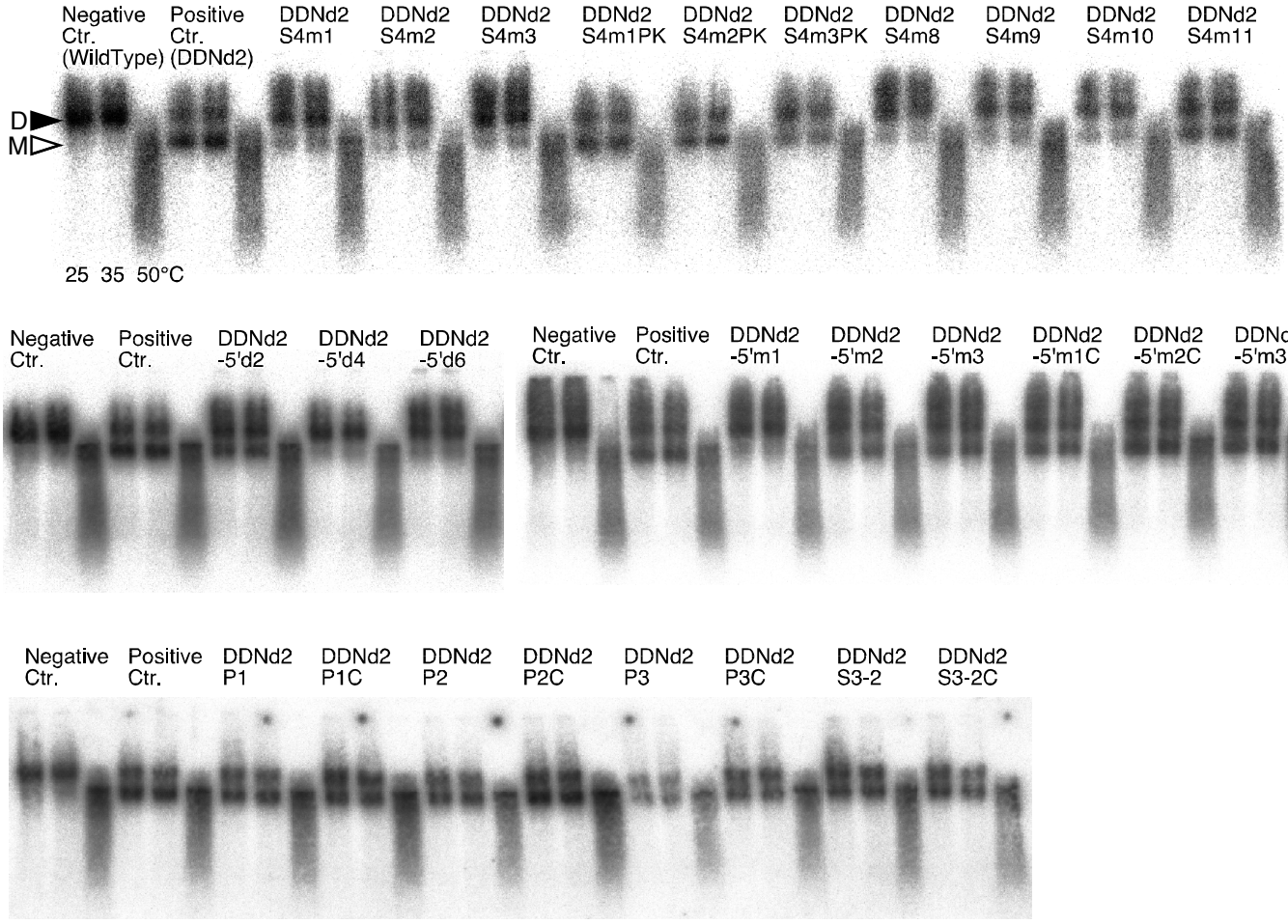
Fig.S2. Northern blot profiles of DDNd2 derivative mutants' viral RNAs from virion. Incubation temperatures of RNA are shown at the bottom of the blot. The positions of viral RNA dimers and monomers on the blots are shown as black and white arrowheads. Northern blots were performed as described in Materials and Methods.

Fig.S3 SHAPE reactivity of the GACGC-GCGUC duplex. A, B) SHAPE value of each base position. The values are averages of four to five independent experiments. Error bars represent SEM. Y axis of right graph of A) was enlarged for clear presentation. C) Native viral RNA stabilities of the wild-type and mutant virions. Incubation temperatures of RNA were shown at the bottom of the blot. The positions of viral RNA dimers and monomers on the blots are shown are shown as black and white arrowheads. Northern blots were performed as described in Materials and Methods. Incubation temperatures of RNA were shown at the bottom of the blot.

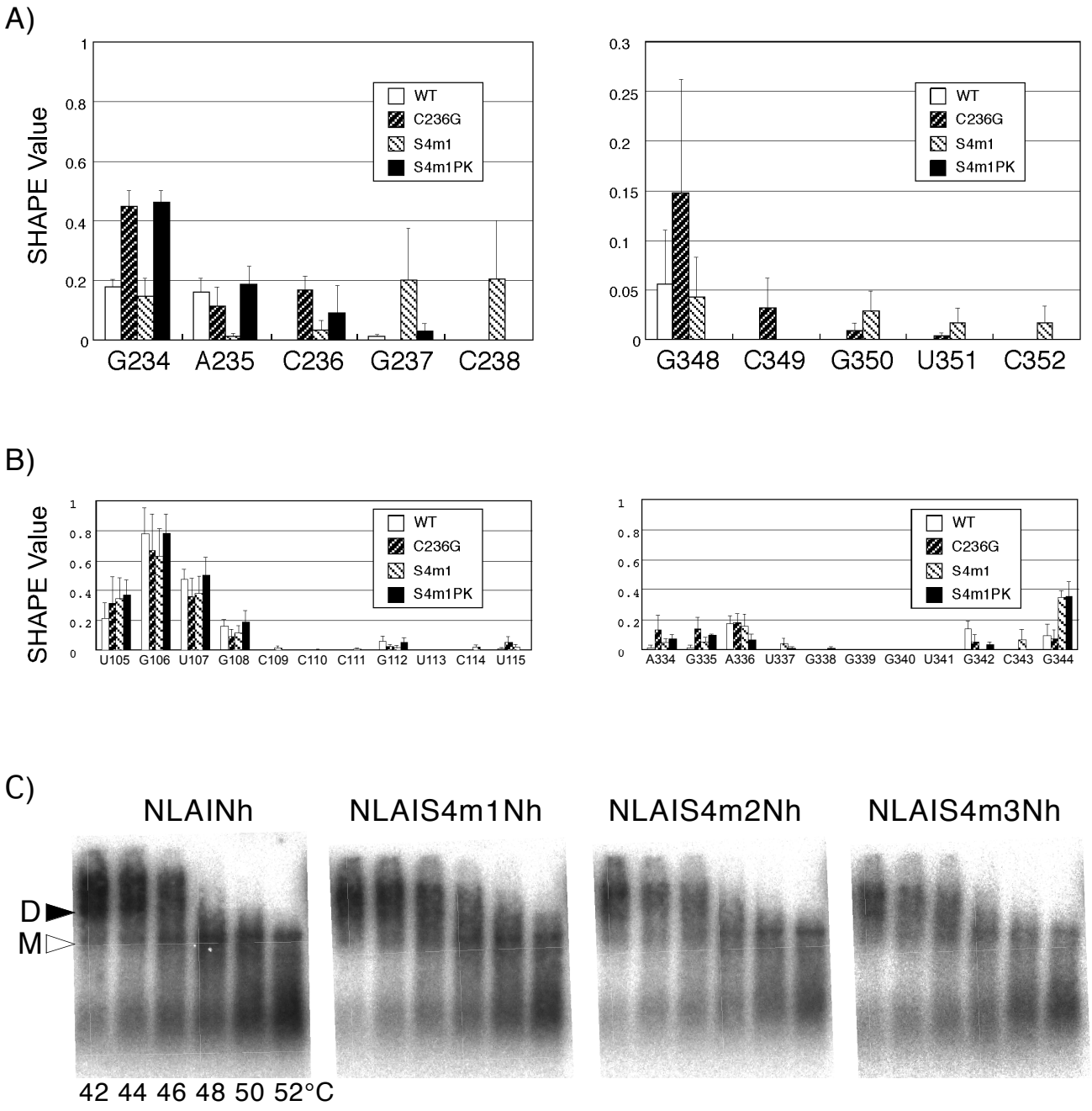
Figure S4. Predicted secondary structure of the DLS model used for constructing its 3D structure. The likely base pairs predicted by CentroidFold and those additionally predicted by MC-Fold with constraints of the most likely base pairs predicted by CentroidFold were highlighted in red and green lines, respectively. Among the predicted pairs, we only considered authentic base pairs, A-U, G-C, or G-U pairs, at the ends of each stem to minimize instability of the structure. In addition, for the construction of DLS dimmer models, we ignored intra-molecular base pairs at the dimerization site in SL1 to preserve inter-molecular base pairs critical for the RNA dimerization. The ignored base pairs in the 3D modeling were shown in gray lines. The base pairs predicted from experimental results in this study were also highlighted with yellow background.



Supplementary Fig.S1



Supplementary Fig.S2



Supplementary Fig.S3

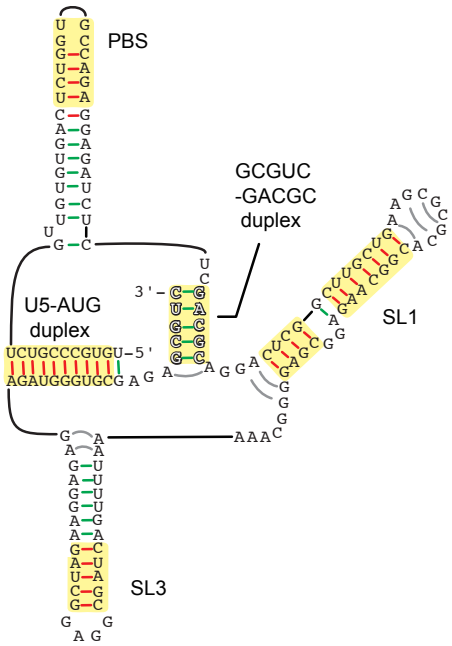

PBS
GCGUC
-GACGC
duplex
U5-AUG
duplex
SL1
SL3
5'
3'

Table S1. Primers used for construction of mutants.

name of primer	sequence	constructs
5'MutF	TCTTCGGGTCTGTTGTGTG	d2M5, d2M5M3
MutR	GACGCTCTCCCCCGGATCTCTCTC	d2M5M3
3'MutCompR	CAAGGTCTCCCCCGGATCTCTCTC	d2M3Comp
dS4half	CGCACCCATCTCTCTCCTTC	d2d3.5
mS4R	CACCGTCTCGCACCCATCTCTC	d2M4
SL4FlongGCCTCG	GGGTGCGAGAGCCTCGGTATTA	NLAIS4m1, DDNd2S4m1
SL4FlongGCTTCG	GGGTGCGAGAGCTTCGGTATTA	NLAIS4m2, DDNd2S4m2
SL4FlongGCCAGT	GGGTGCGAGAGCCAGTGTATTA	NLAIS4m3, DDNd2S4m3
SL4RsGCCTC	GAGGCTCTCGCACCCATCTC	NLAIS4m1, DDNd2S4m1, d2S4m1PK
SL4RsGCTTC	GAAGCTCTCGCACCCATCTC	NLAIS4m2, DDNd2S4m2, d2S4m2PK
SL4RsGCCAG	CTGGCTCTCGCACCCATCTC	NLAIS4m3, DDNd2S4m3, d2S4m3PK
U5L-SL1FGCCTCG	GATCTCTCGAGGCAGGACTC	C236G, d2S4m1PK, NLAIS4m1PK
U5L-SL1RGCCTCG	GAGTCCTGCCTCGAGAGATC	C236G, d2S4m1PK, NLAIS4m1PK
U5L-SL1FGCTTCG	GATCTCTCGAAGCAGGACTC	C236A, d2S4m2PK, NLAIS4m2PK
U5L-SL1RGCTTCG	GAGTCCTGCTTCGAGAGATC	C236A, d2S4m2PK, NLAIS4m2PK
U5L-SL1FGCCAGT	GATCTCTACTGGCAGGACTC	d2S4m3PK
U5L-SL1RGCCAGT	GAGTCCTGCCAGTAGAGATC	d2S4m3PK
GGCGCA-SL4R	GACGCTCTTGCGCCCATC	d2S4m9
GGAGCC-SL4R	GACGCTCTGGCTCCCATC	d2S4m8, d2S4m10
GGCGCA-SL4F	GATGGGCGCAAGAGCGTC	d2S4m9
GGAGCC-SL4F	GATGGGAGCCAGAGCGTC	d2S4m8, d2S4m10
LA3F-GGAGCCcomp	GGTTCCCGTCTGTTGTGTGACTCTGG	d2S4m10, d2S4m11
L4H5'd2ntF	TGCCCGTCTGTTGTGTGAC	d2-5'd2
L4H5'd4ntF	CCCGTCTGTTGTGTGACTC	d2-5'd4
L4H5'd6nt	CGTCTGTTGTGTGACTCTG	d2-5'd6
L4H5'm1	TGTGCGCGTCTGTTGTGTG	d2-5'm1, d2-5'm1C
L4H5'm2	TGTGCCCCTCTGTTGTGTG	d2-5'm2, d2-5'm2C
L4H5'm3	TGTGCCCGTGTGTTGTGTG	d2-5'm3, d2-5'm3C
SL4R-m1	GACGCTCTCGCACGCATCTCTCTCC	d2-5'm1C
SL4R-m2	GACGCTCTCGCACCCCTCTCTCTCC	d2-5'm2C
SL4R-m3	GACGCTCTCGCACCCATGTCTCTCC	d2-5'm3C
PBSstem1F	GTGAGTCTGGTGCCAGAGGAG	P1
PBSstem1Fc	GTGAGTCTGGTGCCAGACGAG	P1C
PBSstem2F	GTGACTGTGGTGCCAGAGGAG	P2
PBSstem2Fc	GTGACTGTGGTGCCACAGGAG	P2C
PBSstem3F	CCACAGGACATCTCTCGACGC	P3
PBSstem3Fc	CCACAGGACATCTGTCGACGC	P3C
PBSstem1R	CTCCTCTGGCACCAGACTCAC	P1
PBSstem1Rc	CTCGTCTGGCACCAGACTCAC	P1C
PBSstem2R	CTCCTCTGGCACCACAGTCAC	P2
PBSstem2Rc	CTCCTGTGGCACCACAGTCAC	P2C
PBSstem3R	GCGTCGAGAGATGTCCTGTGG	S3-1C
PBSstem3Rc	GCGTCGACAGATGTCCTGTGG	S3-2C
SL3stem1F	GCAAAAATATTGACTAGCGG	S3-1, S3-1C
SL3stem2F	GCAAAAATTTTGAGTAGCGG	S3-2, S3-2C
SL3stem1Fc	GCGGAGGCTAGAAGGTGAG	S3-1C
SL3stem2Fc	GCGGAGGCTACAAGGAGAG	S3-2C
SL3stem1R	CCGCTAGTCAATATTTTTGC	S3-1, S3-1C
SL3stem2R	CCGCTACTCAAAATTTTTGC	S3-2, S3-2C
SL3stem1Rc	CTCACCTTCTAGCCTCCGC	S3-1C
SL3stem2Rc	CTCTCCTTGTAGCCTCCGC	S3-2C
Lp4F	TGTGCCCGTCTGTTGTGTG	in general use
TarF	GGTCTCTCTGGTTAGACCAG	in general use
SL4Rs	GACGCTCTCGCACCCATC	in general use
R/U5F	CACTGCTTAAGCCTCAACGATCG	in general use
U5/LF	TCTGTTGTGTGACTCTGGTAAC	in general use
A3F+N	TGTGCCCGTCTGTTGTGTGACTC	in general use
SL1F	ACTCGGCTTGCTGAAGCGCGCAC	in general use
MASgrA1R	AATTCTCCGCCGGTGAGTACCGACGCTCTCGCACCC	pNLAINh, pNLAI-43