Cell Host & Microbe, ***Volume 4***

Supplemental Data

HIV Evades RNA Interference Directed at TAR by an Indirect Compensatory Mechanism

Joshua N. Leonard, Priya S. Shah, John C. Burnett, and David V. Schaffer

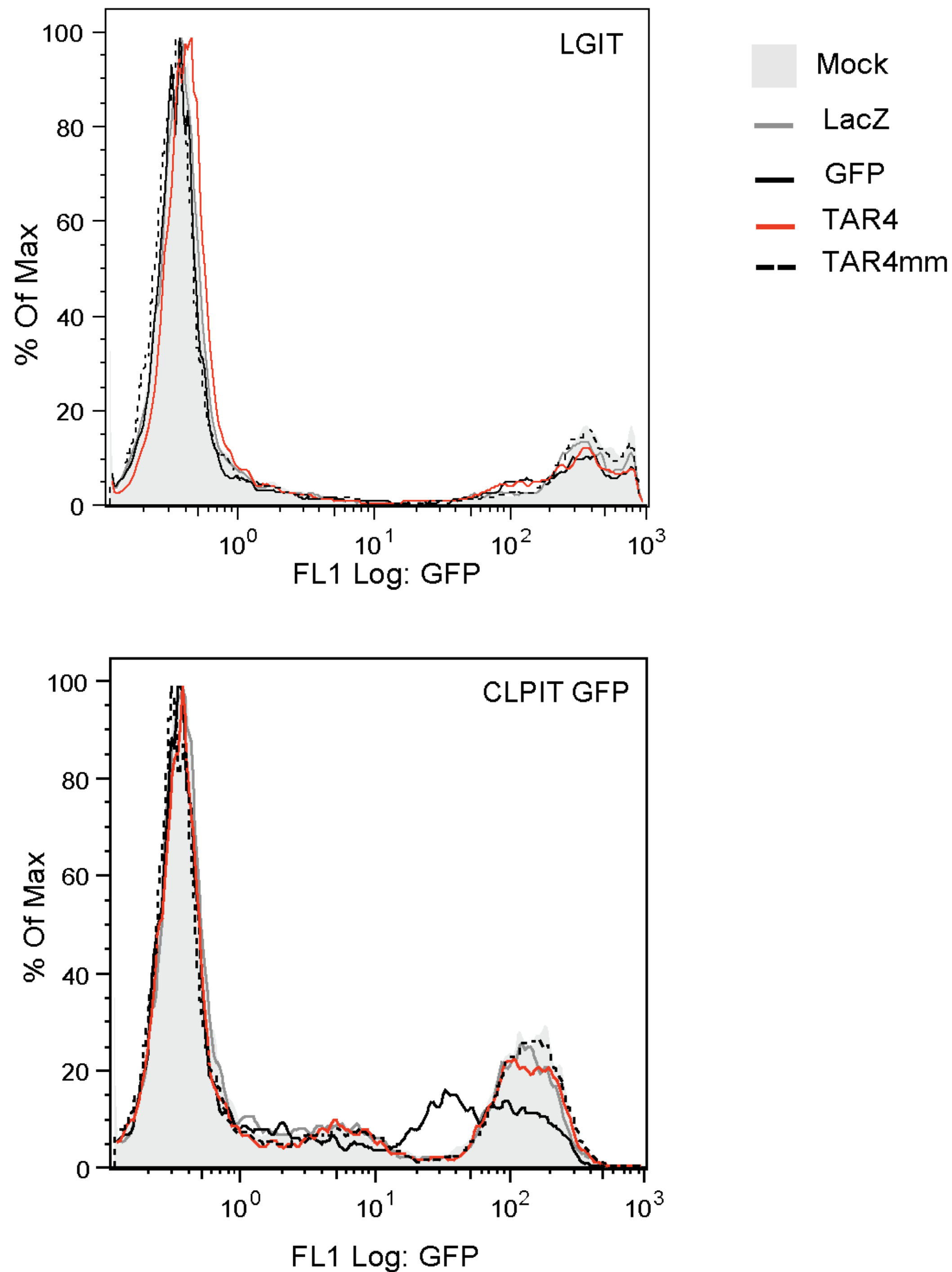


Figure S1: Transient GFP Knockdown

1×10^5 HEK 293T cells infected with LGIT or CLPIT GFP were transiently transfected with 1.5 μg of shRNA-expressing plasmid. Fluorescence was measured 3 days post-transfection. Sample flow data for each cell line is shown for mock (grey silhouette), LacZ control (solid grey line), GFP control (solid black line), TAR4 (solid red line) and the TAR4 mismatch (TAR4mm) control (dashed black line).

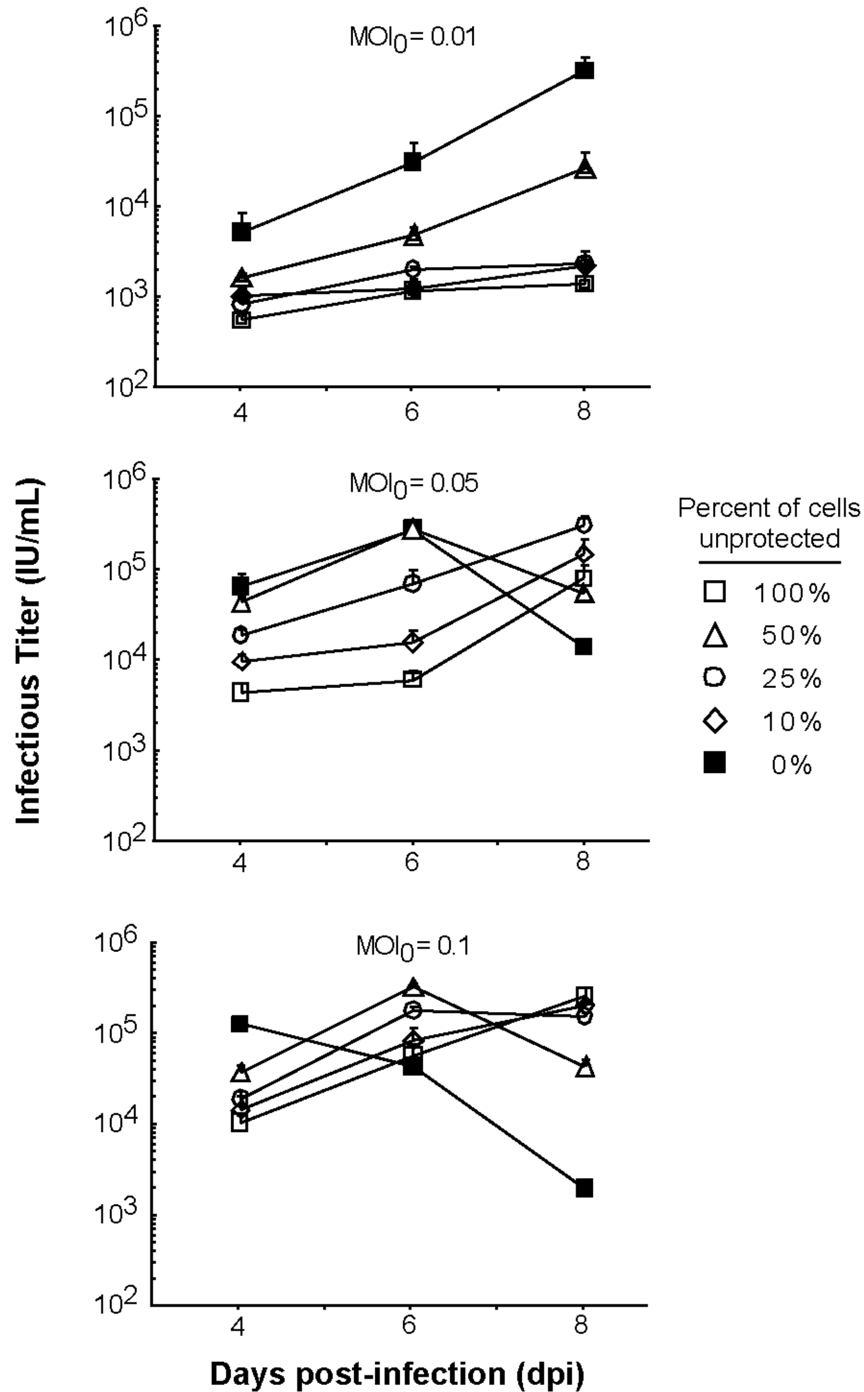


Figure S2: HIV Replication on Mixed Cell Populations

Unprotected SupT1s cells and TAR4-protected cells were mixed at the indicated ratios and challenged with WT virus at an intial MOI (MOI_0) of 0.01, 0.05 or 0.1. Most viral titers at 2 dpi were near or below the detection threshold and are not shown. Experiments were performed in biological triplicate, and error bars indicate one standard deviation.

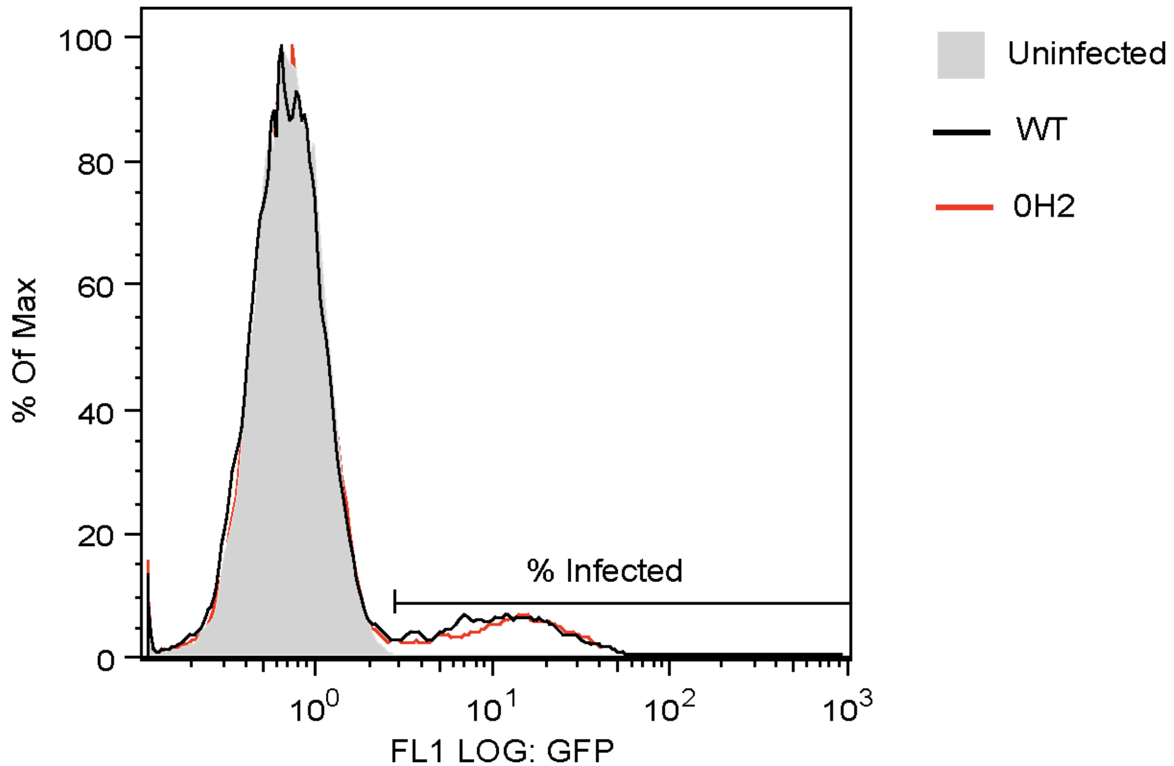
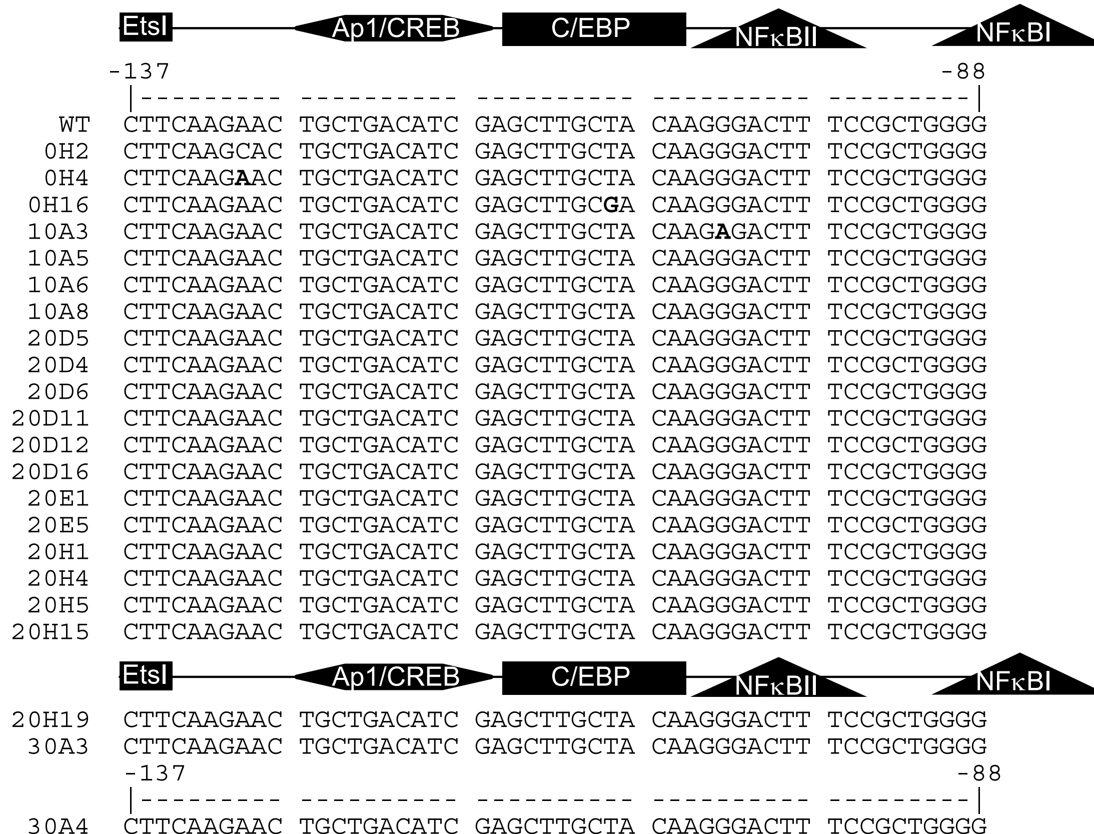


Figure S3: CEM GFP Indicator Cell Line

$1x10^5$ CEM GFP cells were infected with 100 μL of viral supernatant from the indicated HIV culture. Cells were fixed in 2% paraformaldehyde 3 dpi and analyzed using flow cytometry. Thresholds were set such that 0.05% of uninfected cells under similar culture conditions were gated in the "infected" region. This threshold was robust in response to slight differences in CEM GFP sensitivity to mutants compared to WT virus such that altering the threshold between 0.01% and 0.1% did not result in different outcomes for mutant or WT virus.

```
                                                              EtsI
       -187                                              -138
        |--------- ---------- ---------- ---------- ---------|
   WT  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
  0H2  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGACATCCC ATCCGGAGTA
  0H4  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 0H16  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 10A3  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 10A5  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 10A6  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 10A8  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 20D5  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGGGTA
 20D4  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 20D6  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
20D11  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
20D12  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
20D16  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 20E1  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 20E5  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 20H1  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 20H4  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 20H5  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
20H15  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
20H19  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 30A3  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 30A4  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 30A5  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 30A6  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 30G3  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 30G4  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 30G5  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
 30G8  TGGAAGGGCT AATTCACTCC CAAAGAAGAC AAGATATCCC ATCCGGAGTA
```




```
       -137                                              -88
        |--------- ---------- ---------- ---------- ---------|
   WT  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
  0H2  CTTCAAGCAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
  0H4  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 0H16  CTTCAAGAAC TGCTGACATC GAGCTTGCGA CAAGGGACTT TCCGCTGGGG
 10A3  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGAGACTT TCCGCTGGGG
 10A5  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 10A6  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 10A8  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 20D5  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 20D4  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 20D6  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
20D11  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
20D12  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
20D16  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 20E1  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 20E5  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 20H1  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 20H4  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 20H5  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
20H15  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG

20H19  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 30A3  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
        -137                                              -88
        |--------- ---------- ---------- ---------- ---------|
 30A4  CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
```

```
 30A5   CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACCT TCCGCTGGGG
 30A6   CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 30G3   CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 30G4   CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 30G5   CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG
 30G8   CTTCAAGAAC TGCTGACATC GAGCTTGCTA CAAGGGACTT TCCGCTGGGG

          NFκBI         Sp1 III        Sp1 II

       -87                                                ---
        |--------- ---------- ---------- ---------- ---------|
   WT   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
  0H2   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
  0H4   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 0H16   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 10A3   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 10A5   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 10A6   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 10A8   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 20D5   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 20D4   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 20D6   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
20D11   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
20D12   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
20D16   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 20E1   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 20E5   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 20H1   ACTTTCCAGG GAGGCGTGGC CTGGGTGGGA CTGGGGAGTG GCGAGCCCTC
 20H4   ACTTTCCAGG GAGGCGTGGC CTGGGCGGGA CTGGGGAGTG GCGAGCCCTC
 20H5   ACTTTCCAGG GAGGCGTGGC CTGGGCGGGA CTGGGGAGTG GCGAGCCCTC
20H15   ACTTTCCAGG GAGGCGTGGC CTGGGCGGGA CTGGGGAGTG GCGAGCCCTC
20H19   ACTTTCCAGG GAGGCGTGGC CTGGGCGGGA CTGGGGAGTG GCGAGCCCTC
 30A3   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 30A4   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 30A5   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 30A6   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 30G3   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 30G4   ACTTTCCAGG GAGGCGTGAC CTGGG..... .......... ..........
 30G5   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........
 30G8   ACTTTCCAGG GAGGCGTGGC CTGGG..... .......... ..........

                                           Sp1 II      Sp1 I
        ---                                               -50
        |--------- ---------- ---------- ---------- ---------|
   WT   .......... .......... .......... .......CGG GACTGGGGAG
  0H2   .......... .......... .......... .......CGG GACTGGGGAG
  0H4   .......... .......... .......... .......CGG GACTGGGGAG
 0H16   .......... .......... .......... .......CGG GACTGGGGAG
 10A3   .......... .......... .......... .......CGG GACTGGGGAG
 10A5   .......... .......... .......... .......CGG GACTGGGGAG
 10A6   .......... .......... .......... .......CGG GACTGGGGAG
 10A8   .......... .......... .......... .......CGG GACTGGGGAG
 20D5   .......... .......... .......... .......CGG GACTGGGGAG
 20D4   .......... .......... .......... .......CGG GACTGGGGAG
 20D6   .......... .......... .......... .......CGG GACTGGGGAG

                                           Sp1 II      Sp1 I
        ---                                               -50
        |--------- ---------- ---------- ---------- ---------|
20D11   .......... .......... .......... .......CGG GACTGGGGAG
20D12   .......... .......... .......... .......CGG GACTGGGGAG
20D16   .......... .......... .......... .......CGG GACTGGGGAG
 20E1   .......... .......... .......... .......CGG GACTGGGGAG
```

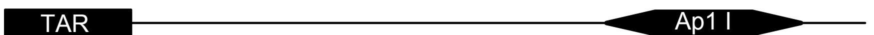
```
 20E5   .......... .......... .......... .......CGG GACTGGGGAG
 20H1   CCTGGG.... .......... .......... .......CGG GACTGGGGAG
 20H4   CCTGGG.... .......... .......... .......CGG GACTGGGGAG
 20H5   CCTGGG.... .......... .......... .......CGG GACTGGGGAG
20H15   CCTGGGCGGG ACTGGGGAGT GGCGAGCCCT CCCTGGGCGG GACTGGGGAG
20H19   CCTGGG.... .......... .......... .......CGG GACTGGGGAG
 30A3   .......... .......... .......... .......CGG GACTGGGGAG
 30A4   .......... .......... .......... .......CGG GACTGGGGAG
 30A5   .......... .......... .......... .......CGG GACTGGGGAG
 30A6   .......... .......... .......... .......CGG GACGGGGGAG
 30G3   .......... .......... .......... .......CGG AACTGGGGAG
 30G4   .......... .......... .......... .......CGG GACTGGGGAG
 30G5   .......... .......... .......... .......CGG GACTGGGGAG
 30G8   .......... .......... .......... .......CGG GACTGGGGAG
```

```
        -49                                                +1
        |--------- ---------- ---------- ---------- ---------|
   WT   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
  0H2   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
  0H4   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 0H16   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 10A3   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 10A5   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTCT GCCTGTACTG
 10A6   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 10A8   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTA
 20D5   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 20D4   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 20D6   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
20D11   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
20D12   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
20D16   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTCTTT GCCTGTACTG
 20E1   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 20E5   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 20H1   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 20H4   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACGG
 20H5   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
20H15   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
20H19   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 30A3   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 30A4   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 30A5   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 30A6   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 30G3   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 30G4   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 30G5   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCTTTTT GCCTGTACTG
 30G8   TGGCGAGCCC TCAGATGCTG CATATAAGCA GCTGCT.TTT GCCTGTACTG
```

```
        +2                                                +51
        |--------- ---------- ---------- ---------- ---------|
   WT   GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
  0H2   GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
  0H4   GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
 0H16   GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
 10A3   GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
 10A5   GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
 10A6   GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
 10A8   GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
 20D5   GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
```

20D4 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
20D6 GGTCTCTCTG GTTAGACCA**A** ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
20D11 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
20D12 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
20D16 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
20E1 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
20E5 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
20H1 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
20H4 **C**GTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
20H5 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
20H15 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
20H19 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
30A3 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
30A4 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
30A5 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
30A6 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
30G3 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
30G4 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA
30G5 GGTCTCTCTG GTTAGACCAG AT**T**TGAGCCT GGGAGCTCTC TGGCTAACTA
30G8 GGTCTCTCTG GTTAGACCAG ATCTGAGCCT GGGAGCTCTC TGGCTAACTA

+52 +101
|--------- ---------- ---------- ---------- ---------|
WT GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
0H2 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
0H4 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGC**T**TTGAG TGCTCAAAGT
0H16 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
10A3 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
10A5 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
10A6 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGC**C**CAAAGT
10A8 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
20D5 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
20D4 GGGAACCCAC TGCTTAAG**T**C TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
20D6 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
20D11 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
20D12 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
20D16 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
20E1 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
20E5 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
20H1 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
20H4 GGGAACCCAC TGCTTAAGCC TC**G**ATAAAGC TTGCCTTGAG TGCTCAAAGT
20H5 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
20H15 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT



20H19 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
+52 +101
|--------- ---------- ---------- ---------- ---------|
30A3 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
30A4 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TT**A**CCTTGAG TGCTCAAAGT
30A5 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
30A6 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
30G3 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
30G4 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
30G5 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT
30G8 GGGAACCCAC TGCTTAAGCC TCAATAAAGC TTGCCTTGAG TGCTCAAAGT

+102 +151
|--------- ---------- ---------- ---------- ---------|
WT AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
0H2 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
0H4 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
0H16 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
10A3 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC

10A5 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
10A6 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
10A8 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
20D5 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
20D4 AGTGTGTGCC C**A**TCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
20D6 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
20D11 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
20D12 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
20D16 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
20E1 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA **C**CCCTCAGAC
20E5 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAG**C**C
20H1 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
20H4 AGTGTGTGCC CGTCTGTTGT GTGA**T**TCTGG TAACTAGAGA TCCCTCAGAC
20H5 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
20H15 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
20H19 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
30A3 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
30A4 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
30A5 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
30A6 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
30G3 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
30G4 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
30G5 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC
30G8 AGTGTGTGCC CGTCTGTTGT GTGACTCTGG TAACTAGAGA TCCCTCAGAC

Ap1 III NFAT/AP3

+152 +181
|--------- ---------- ---------- |
WT CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
0H2 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
0H4 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
0H16 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
10A3 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
10A5 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
10A6 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
10A8 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
20D5 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
20D4 CCTTTT.AGT CAGTGTGGAA AATCTCTAG**T** A
20D6 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
20D11 CCTTTT.AGT CAGTGTGGAA AATCTC**C**AGC A

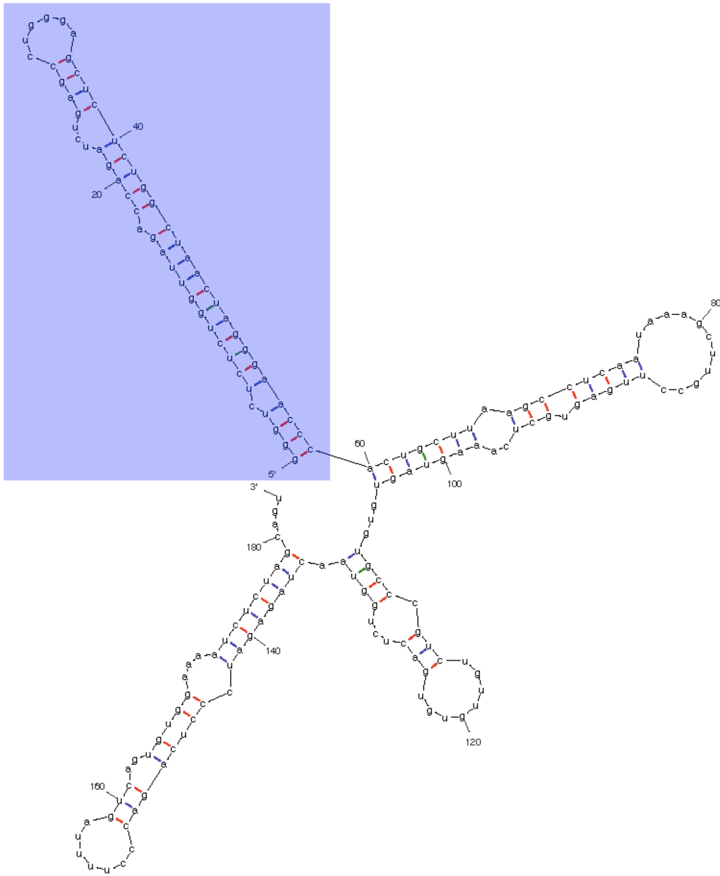
Ap1 III NFAT/AP3

+152 +181
|--------- ---------- ---------- |
20D12 CCTTTT.AGT CAGTGTGGAA AATCTCTAG**A** A
20D16 CCTTTT.AGT **A**AGTGTGGAA AATCTCTAGC A
20E1 C**T**TTTT**T**A**A**T CAGTGTGGAA AATCTCTAGC A
20E5 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
20H1 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
20H4 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
20H5 CCTTTT.AGT CAGTGCGGAA AATCTCTAGC A
20H15 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
20H19 CCTT**G**T.AGT CAGTGTGGAA AATCTCTAGC A
30A3 CCTTTT.AGT CAGTGTGGAA AATCTCTAG**A** A
30A4 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
30A5 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
30A6 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
30G3 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
30G4 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
30G5 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A
30G8 CCTTTT.AGT CAGTGTGGAA AATCTCTAGC A

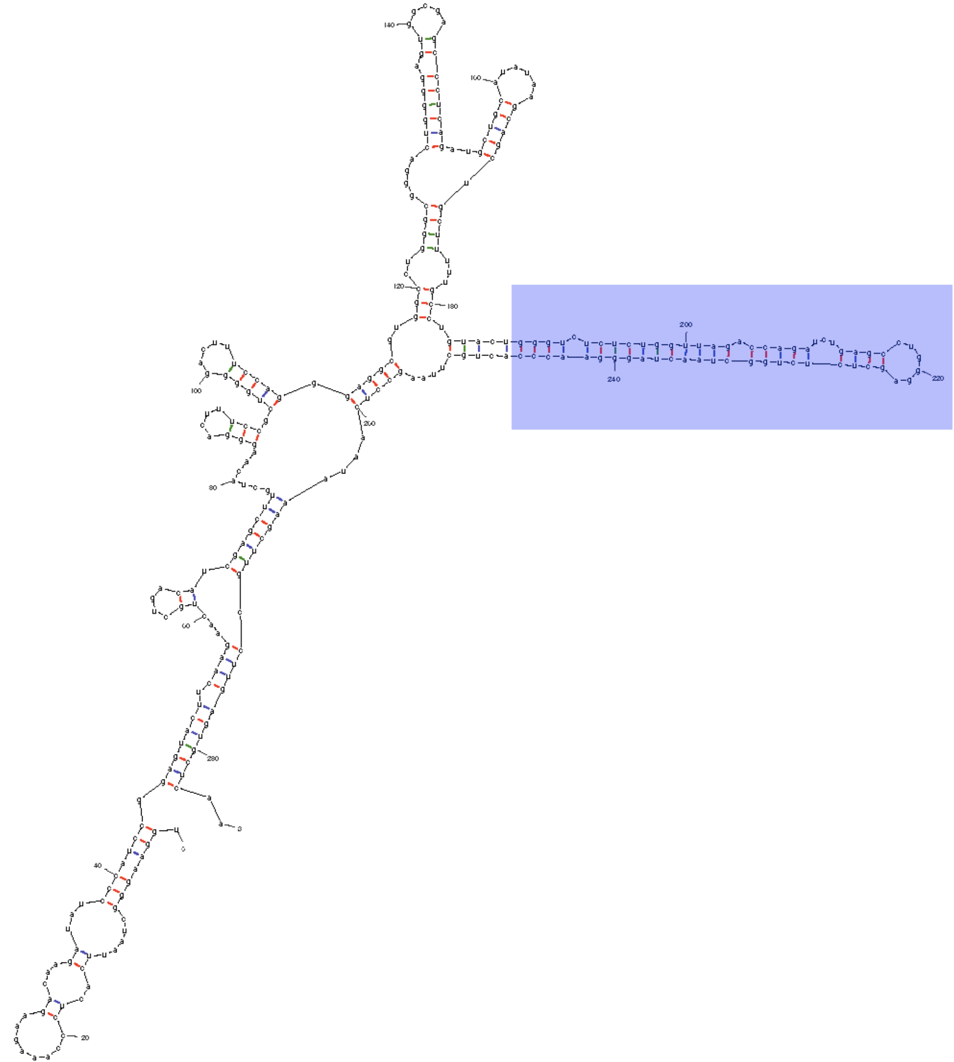
Figure S4: Complete Mutant LTR Sequences

Mutants were aligned using Vector NTI AlignX. Point mutations are indicated in bold face. Exact sequence repeats of 20H mutants, indicated in red type, suggest that Sp1 site multimers were generated through recombination. Transcription factor binding sites are indicated directly above the sequence

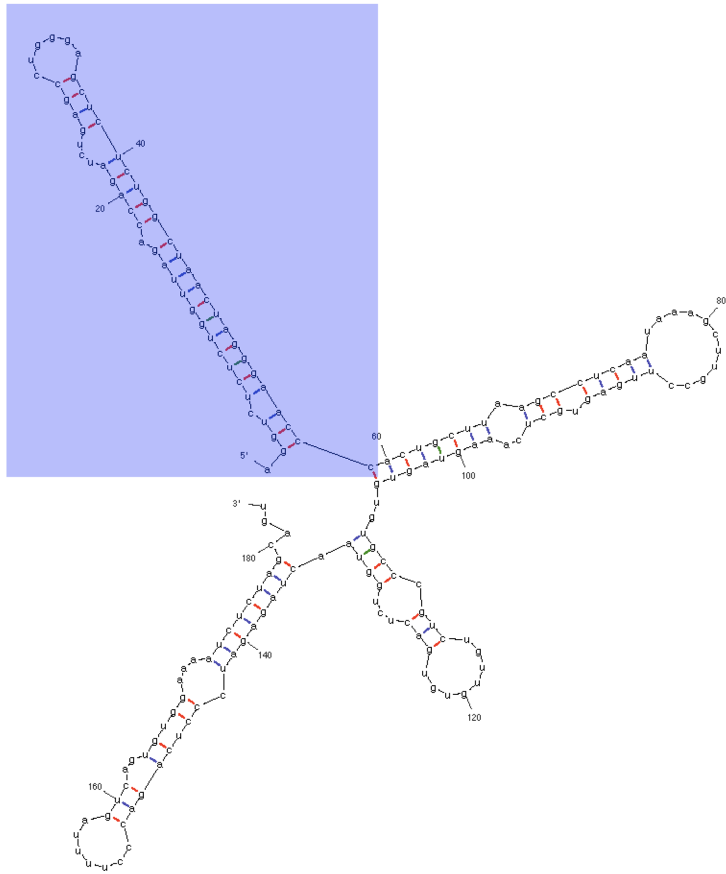
WT 5' (R-U5)



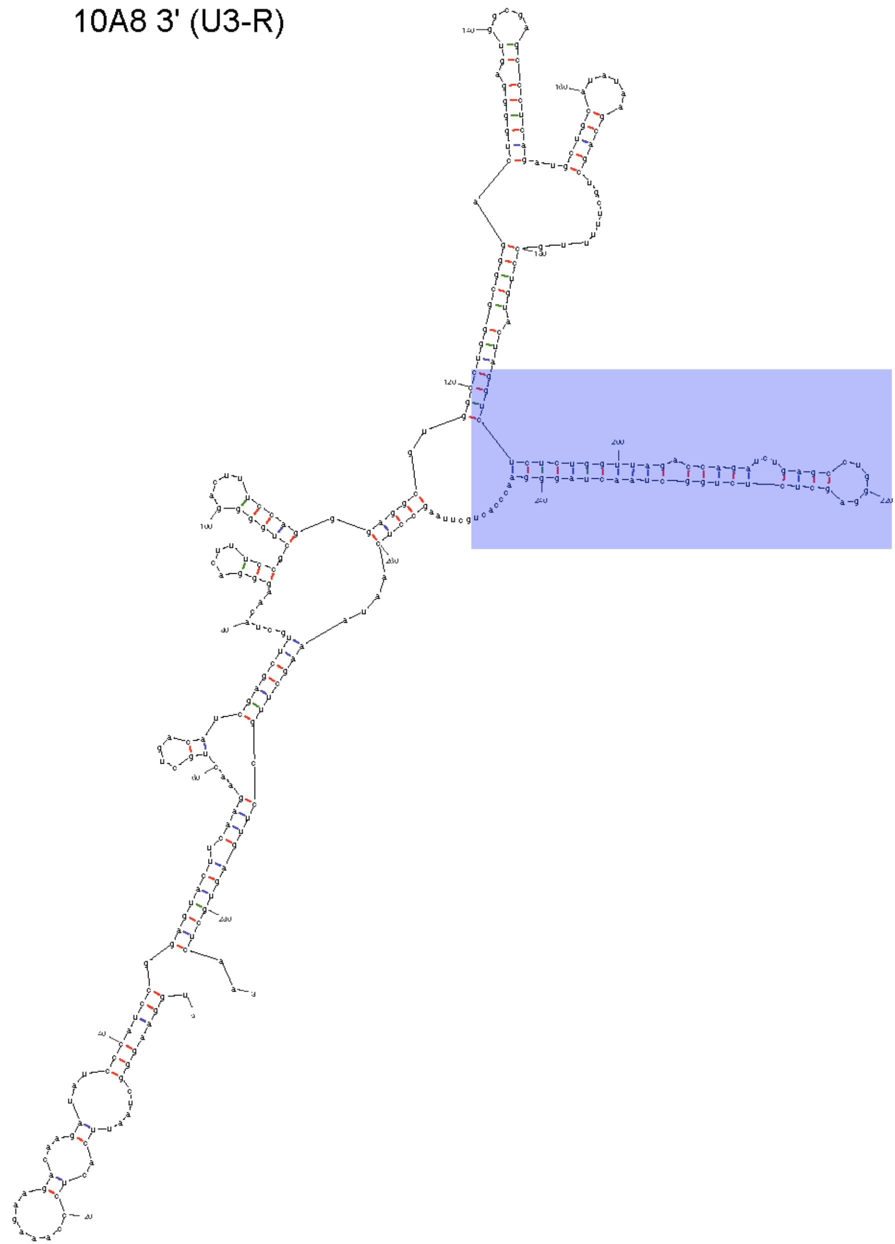
WT 3' (U3-R)



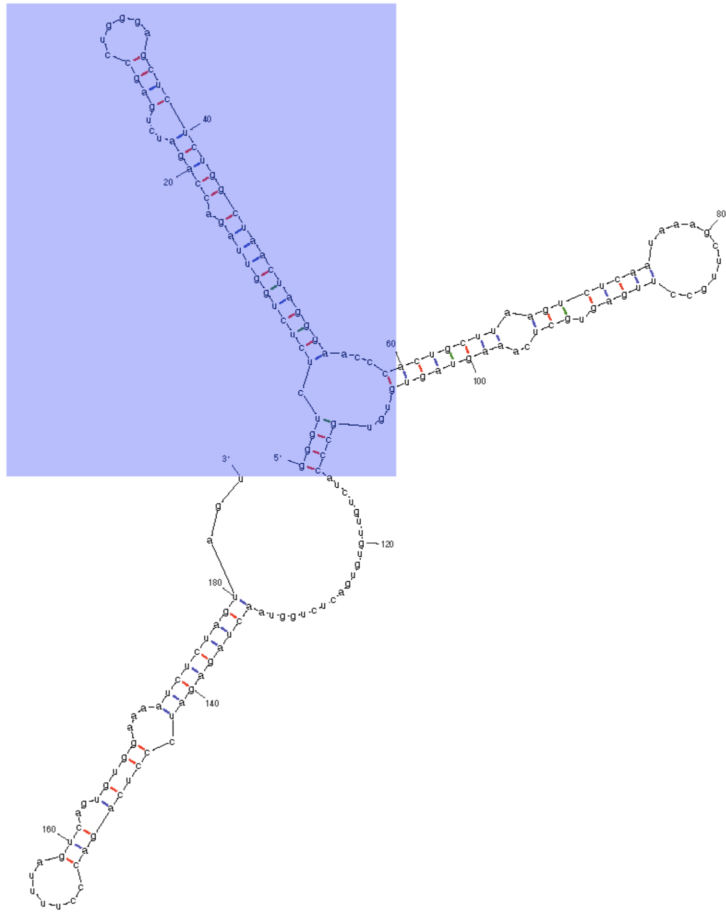
10A8 5' (R-U5)



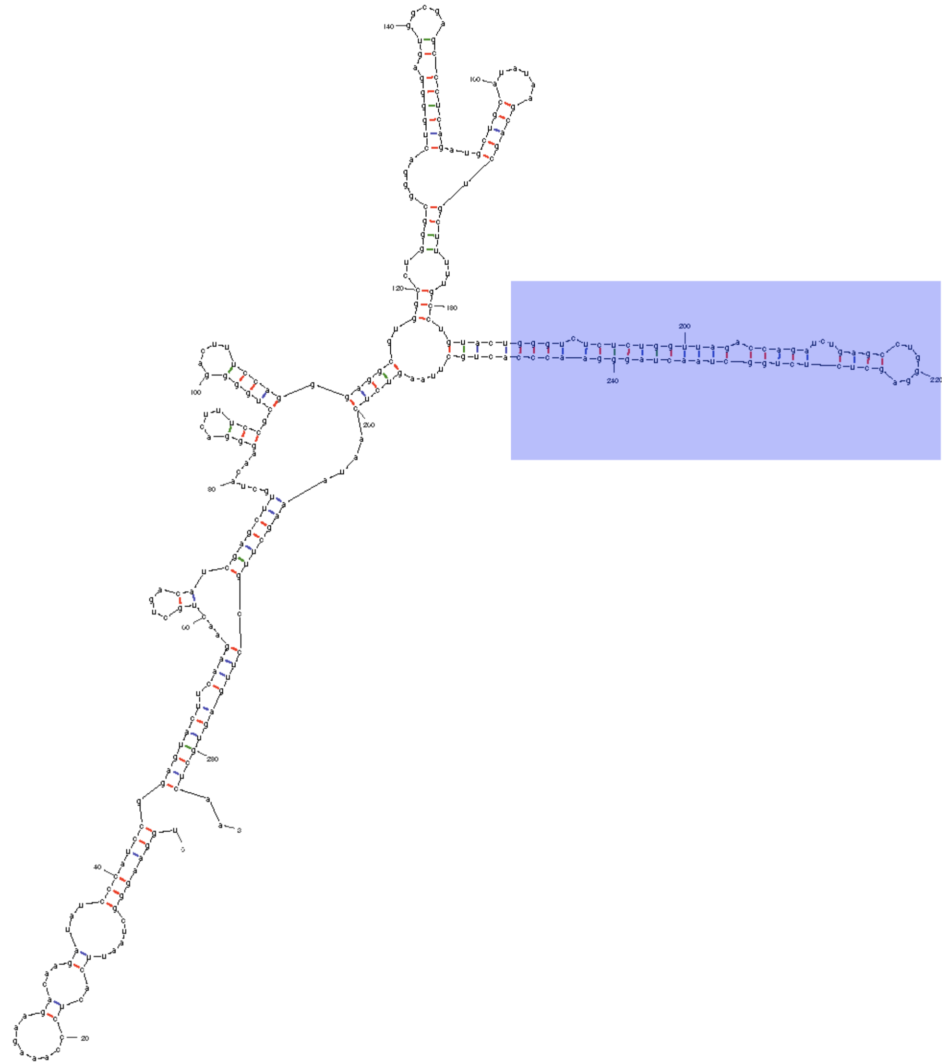
10A8 3' (U3-R)



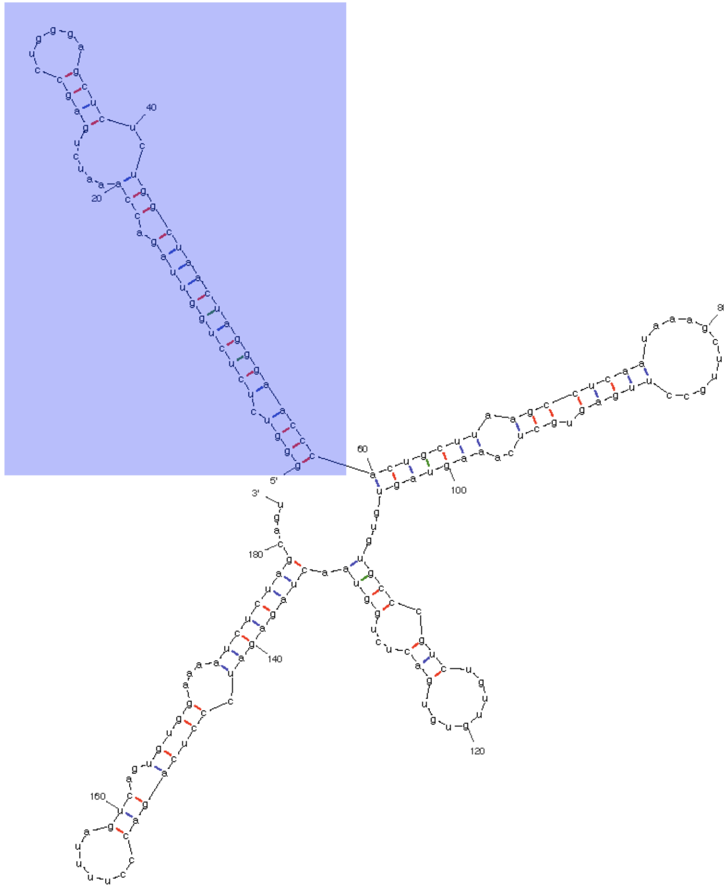
20D4 5' (R-U5)



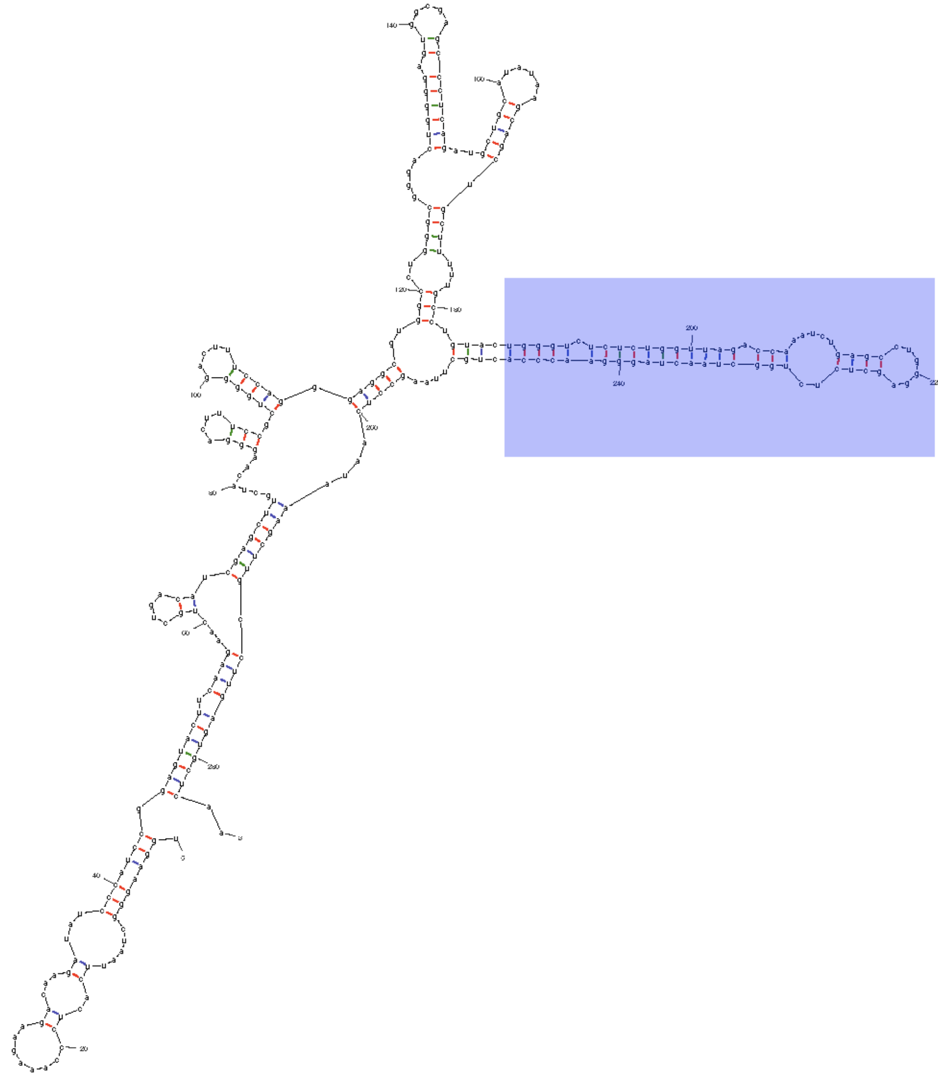
20D4 3' (U3-R)



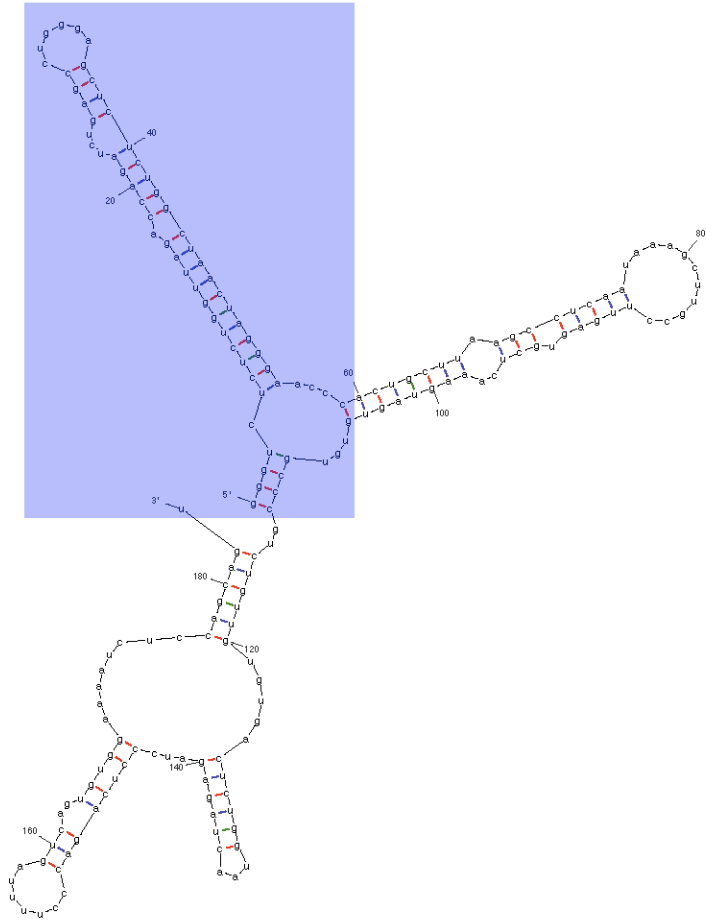
20D6 5' (R-U5)



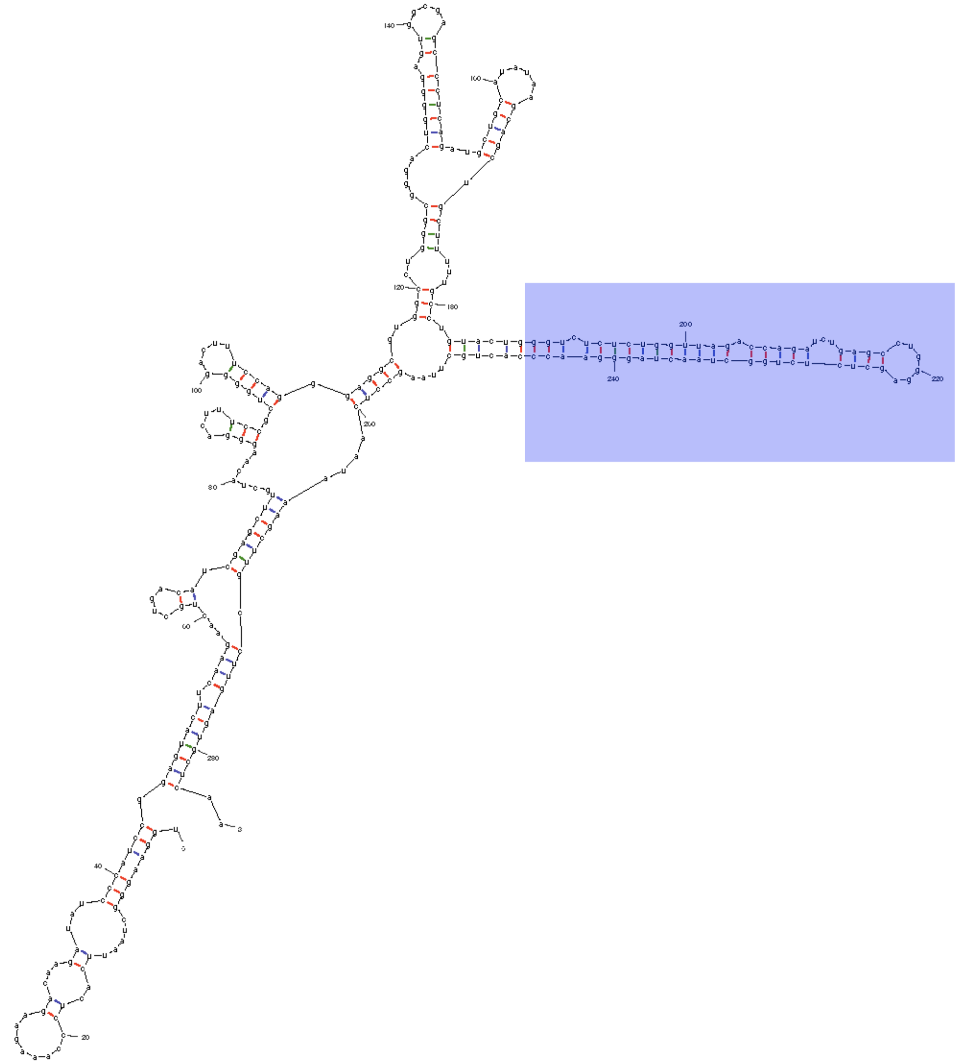
20D6 3' (U3-R)



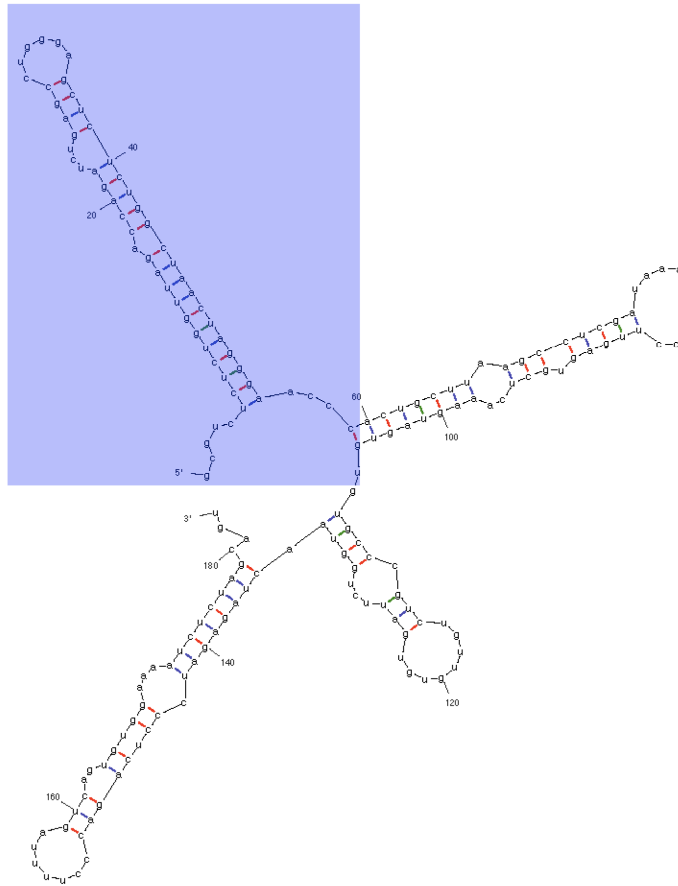
20D11 5' (R-U5)



20D11 3' (U3-R)



20H4 5' (R-U5)



20H4 3' (U3-R)

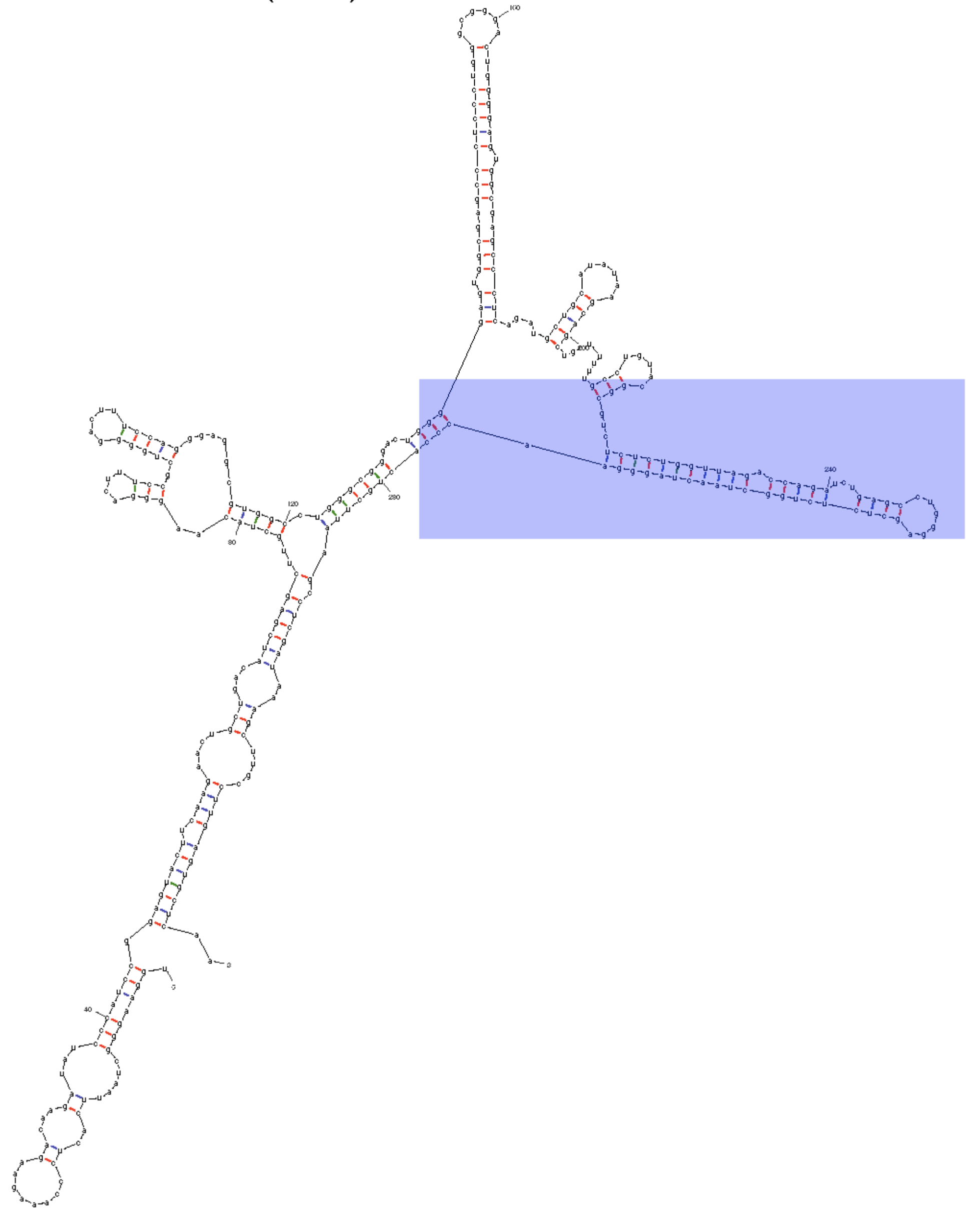


Figure S5: RNA Secondary Structure

Secondary structure of the 5' (left) and 3' (right) UTRs were predicted for WT and mutant sequences in mFold using the R and U5 regions and the U3 and R regions, respectively. The TAR hairpin is highlighted in blue. Mutants with TAR structures that vary from WT structure are shown, however both 5' and 3' structures may not vary simultaneously. Of all mutants analyzed, only mutants 10A8, 20D4, 20D11 and 20H4 had altered structures at the base of the 5' hairpin. Only mutant 10A8 had an altered structure at the base of the 3' hairpin and only 20D6 had any alterations adjacent to the RNAi-targeted region.

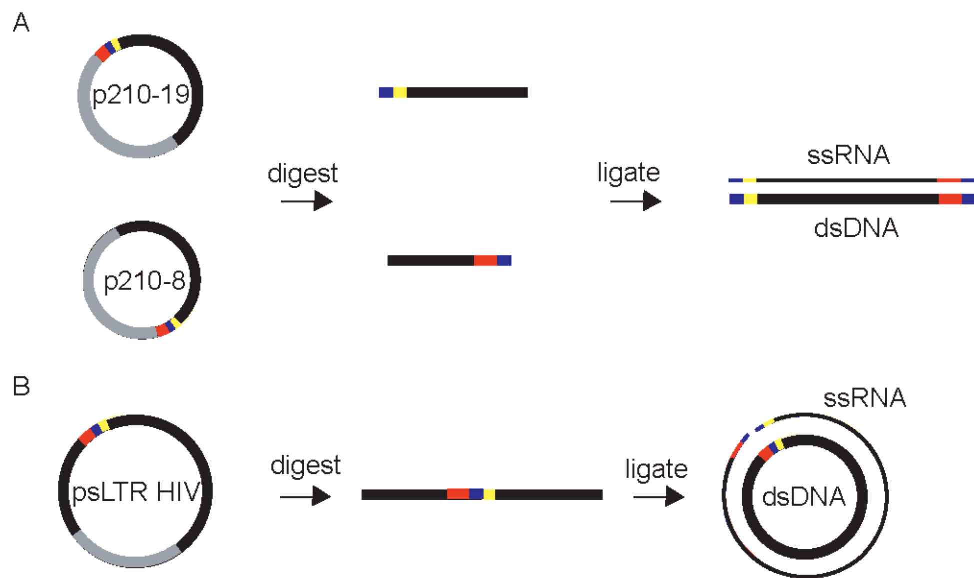


Figure S6: Generation of HIV Using a Single-LTR Platform (psLTR HIV)
(A) In the hemigenomic method, the p210-19 and p210-8 plasmids are linearized and rejoined to form linear full length HIV (black). The linear DNA is introduced into HEK 293Ts and transcribed for viral production. (B) In the psLTR HIV method the HIV genome (black) is interrupted by bacterial sequence (grey). The 3' half of the HIV genome and 3' LTR from p210-8 were fused immediately upstream of the 5' half of the genome, plasmid origin of replication and ampicillin selection gene from p210-19 (replacing the 5' LTR). Each psLTR HIV variant was digested with *Eco* RI (New England Biolabs), to excise the non-HIV sequences, and recircularized with T4 DNA ligase (New England Biolabs) to produce full-length HIV with a single LTR (U3, red; R, blue; U5, yellow). The LTR is transcribed once at the beginning and once at the end of each transcript in HEK 293Ts. GFP was also inserted into the *nef* open reading frame to track HIV replication in cell culture.

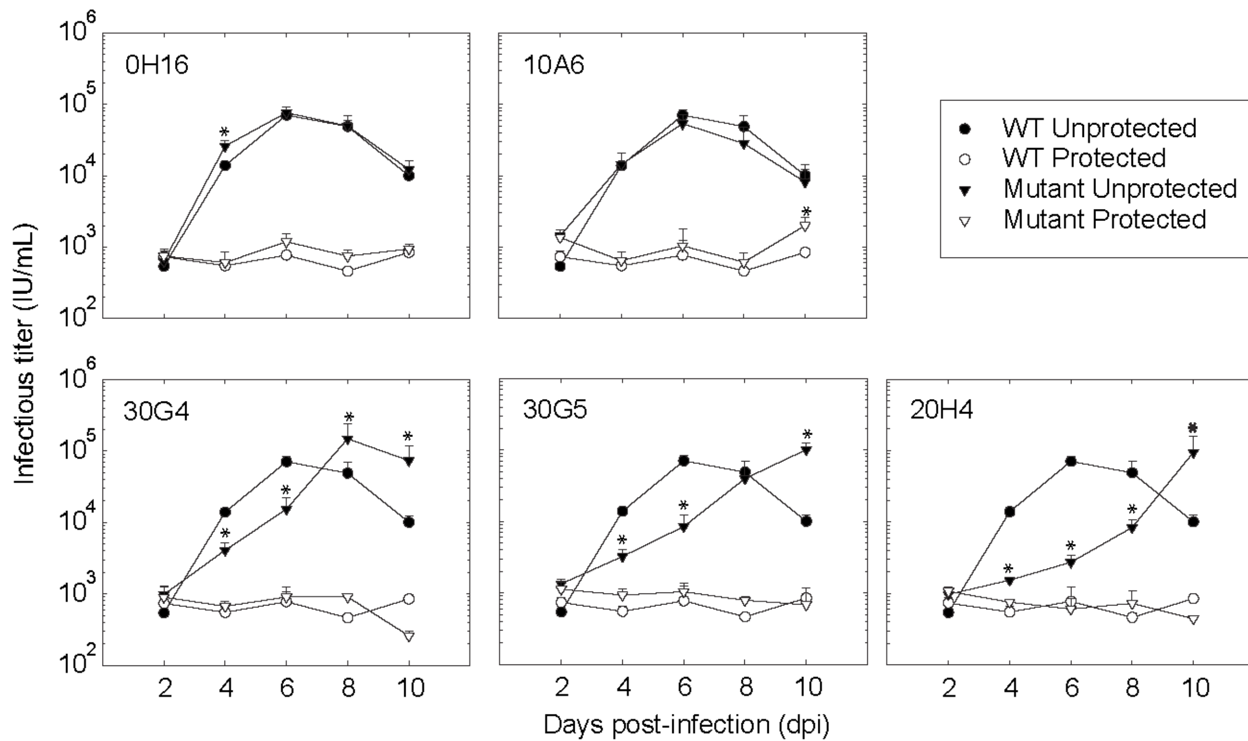


Figure S7: Replication of Nonevasion Mutants

TAR4-protected or unprotected SupT1 cells were challenged with either WT virus or a homogeneous preparation of each potential evasion mutant at a MOI of 0.015. Infections were performed in biological triplicate, error bars represent one standard deviation, and (*) indicates a significantly different titer as compared to WT virus replicating in the same cell type ($p<0.05$).

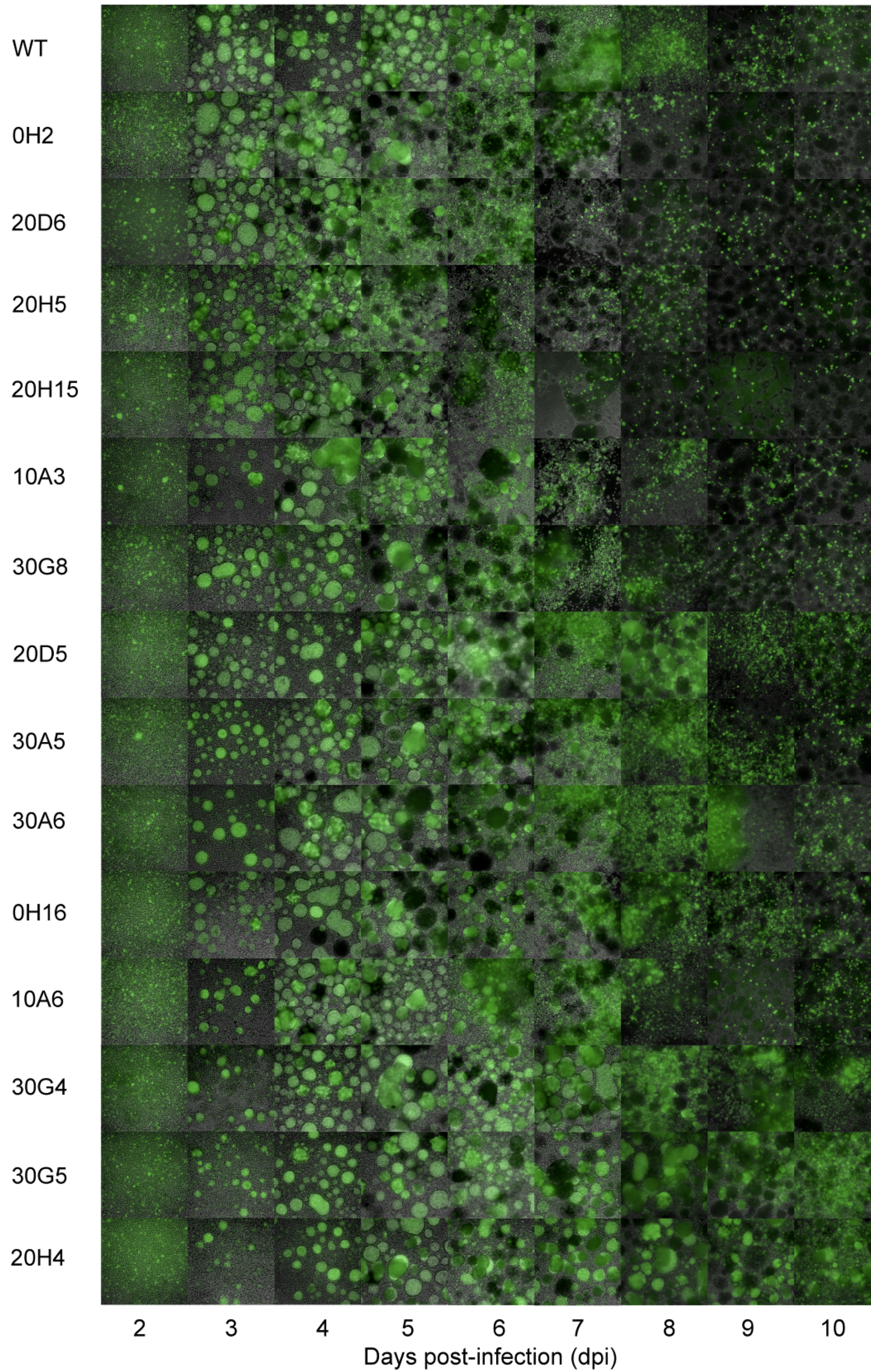


Figure S8: Viral Replication-Mediated Cell Death
Overlays of GFP and bright field images of HIV replication in unprotected cells.

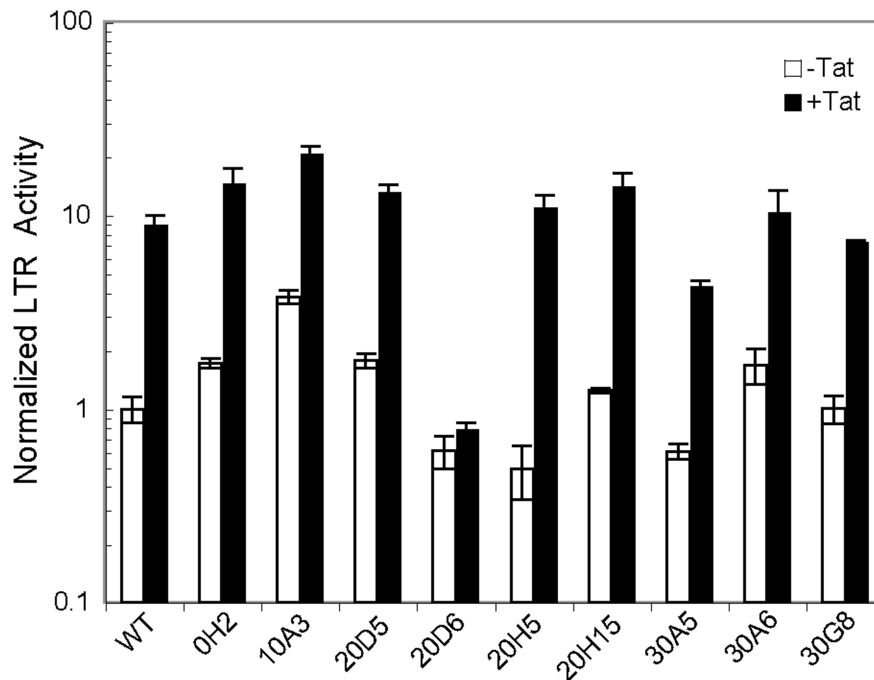


Figure S9: Transcriptional Activities Determined by Luciferase Assay
Basal and transactivated transcription was measured by luciferase assay in naive or Tat-expressing SupT1s transduced with the indicated LLIG variant. Luminescence was normalized by infection efficiency (determined by QPCR) and protein content of cell lysates. All values were then normalized to the WT luminescence in the absence of Tat (759 RLU). Experiments were performed in technical triplicate and error bars represent one standard deviation.

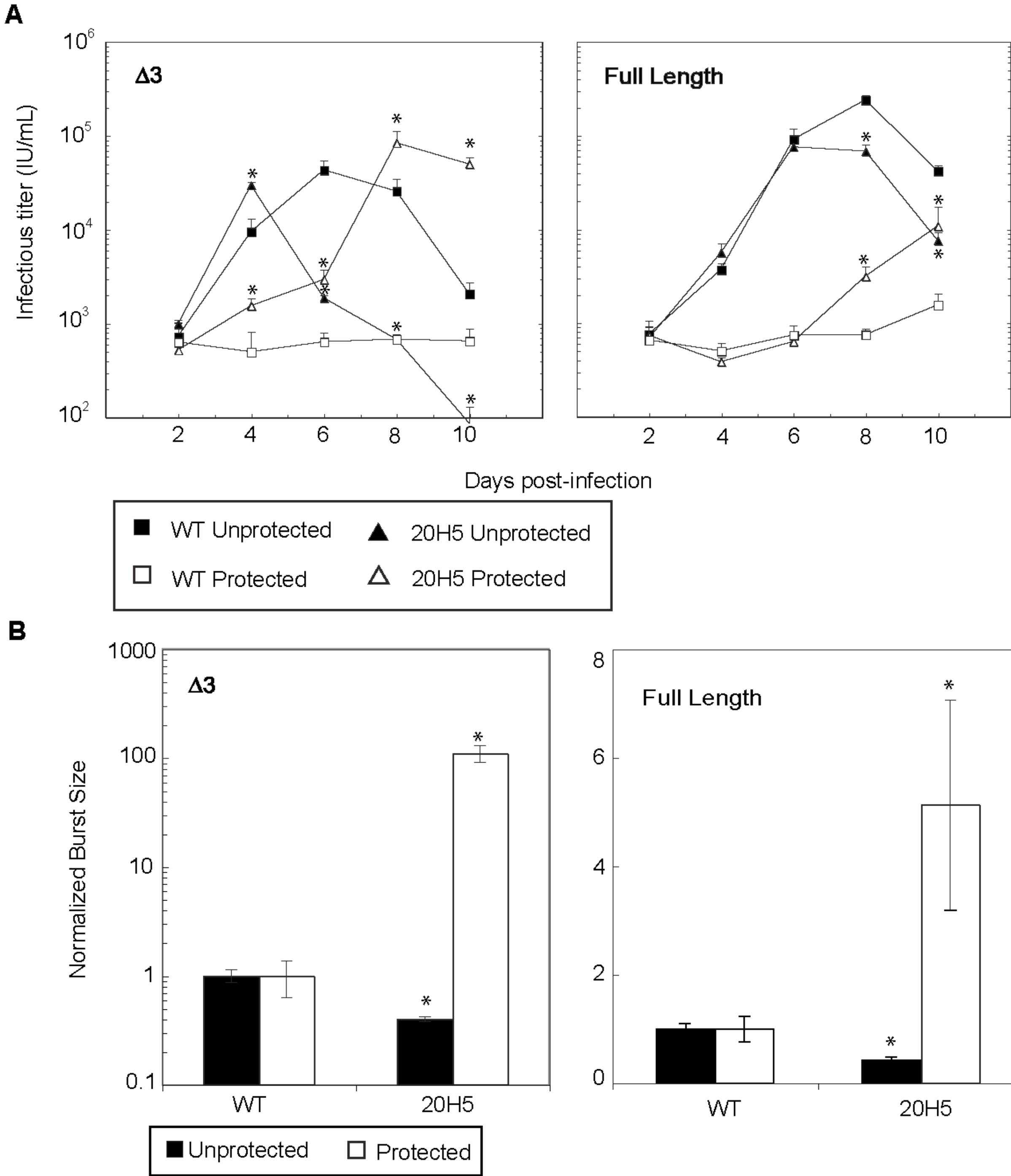


Figure S10: Enhanced Replication of Sp1 Duplication Mutants in Full-Length and Triple-Deletion (Δ3) Virus

(A) Unprotected and protected cells were infected with the indicated virus at a MOI of 0.015 and 0.002 for Δ3 and full length virus, respectively. (B) Total viral yield (burst size) was calculated by integrating the titers from (A) and normalizing to WT virus (in identical cells). Experiments were performed in biological triplicate and error bars indicate one standard deviation, and (*) indicates a significantly different titer (or burst size) as compared to WT virus replicating in the same cell type ($p<0.05$).

Primer name	Oligonucleotide sequence (5' to 3')
Universal (+)	ATAAGAATGCGGCCGCCCCGGGGATCCAAGGTCGGG
LacZ (-)	GCTCTAGAAAAAGTGACCAGCGAATACCTGTTCTCTTGAAAC AGGTATTCGCTGGTCACGGTGTTTCGTCCTTTCCAC
GFP (-)	GCTCTAGAAAAAGAAGAAGATGGTGCGCTCCTCTCTTGAAGG AGCGCACCATCTTCTTCGGTGTTTCGTCCTTTCCAC
TAR1 (-)	GCTCTAGAAAAAGACCAGATCTGAGCCTGGGTCTCTTGAACC CAGGCTCAGATCTGGTCGGTGTTTCGTCCTTTCCAC
TAR2 (-)	GCTCTAGAAAAAGATCTGAGCCTGGGAGCTCTCTCTTGAAGA GCTCCCAGGCTCAGATCGGTGTTTCGTCCTTTCCAC
TAR3 (-)	GCTCTAGAAAAAGACCAGATCTGAGCCTGGTGACAGGAAGC CAGGCTCAGATCTGGTCTCGGTGTTTCGTCCTTTCCAC
TAR4 (-)	GCTCTAGAAAAAGCCTGGGAGCTCTCTGGCTATCTCTTGAAT AGCCAGAGAGCTCCCAGGCGGTGTTTCGTCCTTTCCAC
TAR4mm (-)	GCTCTAGAAAAAGCCTGGGACGTCTCTGGCTATCTCTTGAAT AGCCAGAGACGTCCCAGGCGGTGTTTCGTCCTTTCCAC

Table S1: Primers for shRNA Construction
Each PCR reaction used the pTZU6+1 template, the Universal (+) forward primer, and a unique reverse (-) primer. Each PCR product contains the U6 promoter, the shRNA encoding sequence (sense-loop-antisense), and a TTTTT RNA Polymerase III transcription termination signal.

Mutant	Frequency	Mutant	Frequency
0H2	1/22	20E5	1/6
0H4	1/22	20H1	1/9
0H16	1/22	20H4	1/9
10A3	1/14	20H5	5/9
10A5	1/14	20H15	1/9
10A6	1/14	20H19	1/9
10A8	1/14	30A3	1/10
20D4	1/16	30A4	1/10
20D5	1/16	30A5	1/10
20D6	1/16	30A6	1/10
20D11	1/16	30G3	1/7
20D12	1/16	30G4	1/7
20D16	1/16	30G5	1/7
20E1	1/6	30G8	1/7

Table S2: Mutation Frequency per Culture

The frequency of each mutation as listed in Figure 3 is given as a fraction of the total number of samples sequenced. The remaining fraction of each culture contained no mutations within the LTR.