Supplementary Information to:

Spacer-length dependence of programmed -1 or -2 ribosomal frameshifting on a U_6A heptamer supports a role for mRNA tension in frameshifting

Zhaoru Lin, Robert J. C. Gilbert* and Ian Brierley[†].

Division of Virology, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, United Kingdom. *Division of Structural Biology, Henry Wellcome Building for Genomic Medicine, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, United Kingdom

Supplementary Figure S1

Further characterisation of AON-mediated frameshifting

(A). The two bases immediately 5' of the U_6A stretch within pFSHIV-AON stopAll were changed as indicated. Also shown are the sequences of four AONs complementary to the mRNA beginning three bases downstream of the U_6A stretch.

(B). *In vitro* translation of pFSHIV-AON stopAll variants. The wild-type *Nco-I*-derived mRNA (wt) and three sequence variants (pAON AG, CA, CG) were translated in RRL in the absence (H_20) or presence of 10μM 25MO, 25OMe, 15OMe or 15RNA. Products were analysed and quantified as in the legend to Figure 1. The frameshifting efficiency measured for each signal (to the nearest integer) is indicated below the relevant lanes (%FS) and takes

into account the number of methionines present in each product (nFS, 19; -1 FS, 10; -2 FS, 11).

Supplementary Figure S2

The effect of slippery sequence -pseudoknot (PK) spacing on -1 and -2 frameshifting on U_6A.

(A). The spacer of pFSHIV-SRV, which contains the stimulatory RNA pseudoknot of the SRV-1 *gag/pro* signal, was changed from three to nine nucleotides as indicated. A diagrammatic representation of potential translation products of pFSHIV-SRV mRNAs and predicted molecular masses is also shown. The sizes of the encoded frameshift products were normalised by appropriate deletion of bases downstream of the PK.

(B). Messenger RNAs derived from Nco I-cut pFSHIV-SRV spacer variants were translated in RRL and products analysed and quantified as in the legend to Figure 3. The numbers above each gel represent the spacer length. The frameshifting efficiency measured for each signal (to the nearest integer) is indicated below the relevant lanes (-1%FS; -2%FS) and takes into account the number of methionines present in each product (nFS, 10; -1 FS, 10; -2 FS, 11). The asterisk indicates that the line also marks the position of the -1 FS product of pFS cass 5 control mRNA (IBV PK).

Supplementary Figure S3

Mass spectrometry traces.

(A). MALDI mass fingerprint of tryptic digest of -2FS-eGFP-N2 fusion protein.

(B). Peptide coverage of -2FS-eGFP-N2 fusion protein.

(C). Spectrum of the observed γ- and β-ions identified by MS/MS analysis resulting from the fragmentation of the doubly charged 2030pk peptide (panel B).

(D). Table of the observed γ- and β-ions identified by the MS/MS analysis.

Figure S1

A

3' ACCCUAGGAAGUUGA 5' *15RNA*

3' ACCCUAGGAAGUUGA 5' *15OMe*

3' ACCCUAGGAAGUUGAAGGGACUCGA 5' *25OMe*

3' ACCCTAGGAAGTTGAAGGGACTCGA 5' *25MO*

pFSHIV-AON stopAll (wt) GAUCUGAAUUUUUUUACGGUGGGAUCCUUCAACUUCCCUGGGCUCGAAAA

pAON AG -------AGUUUUUUA----------

pAON CA -------CAUUUUUUA----------

pAON CG -------CGUUUUUUA----------

B

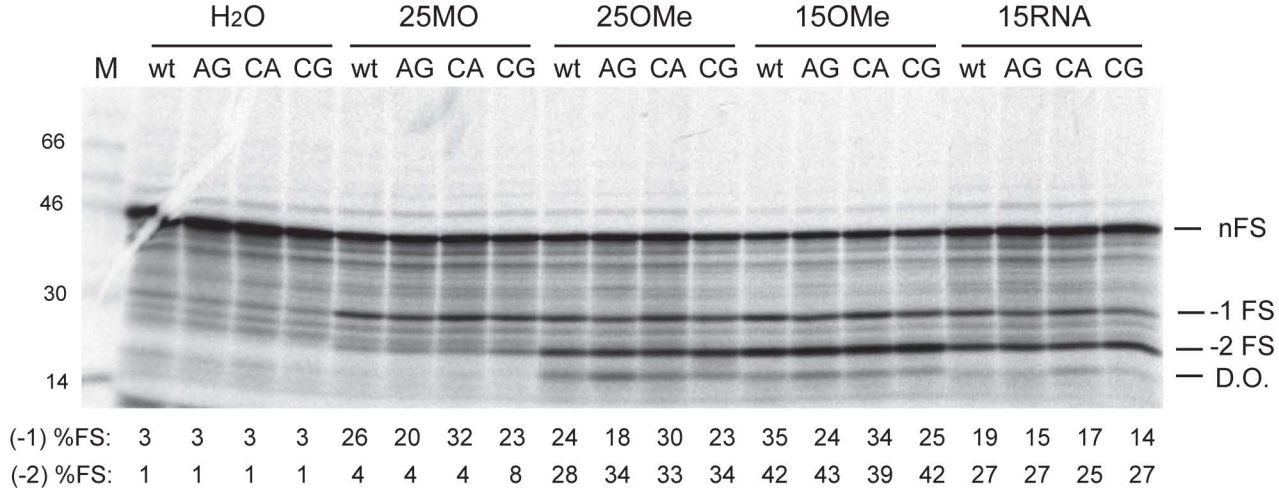


Figure S2

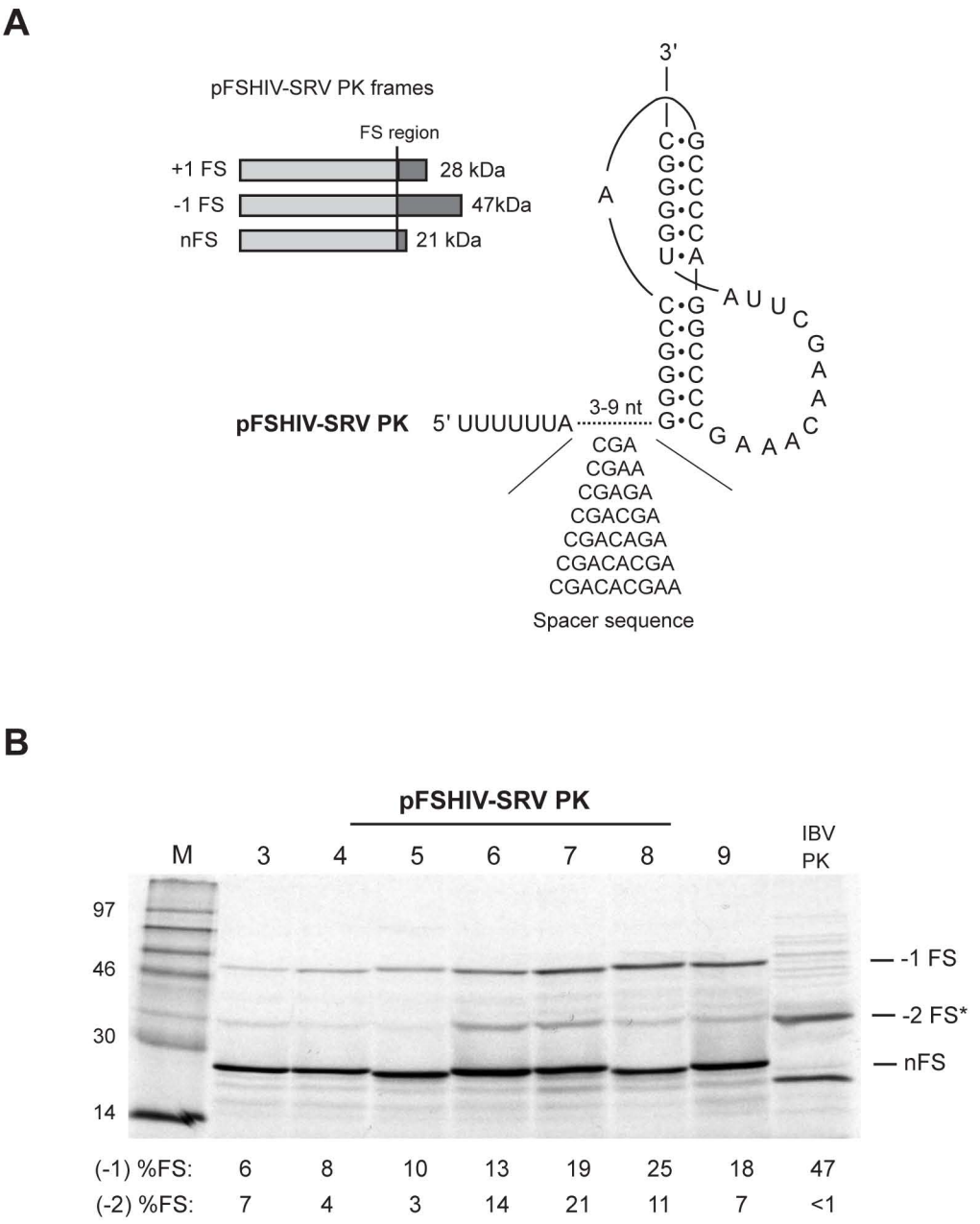
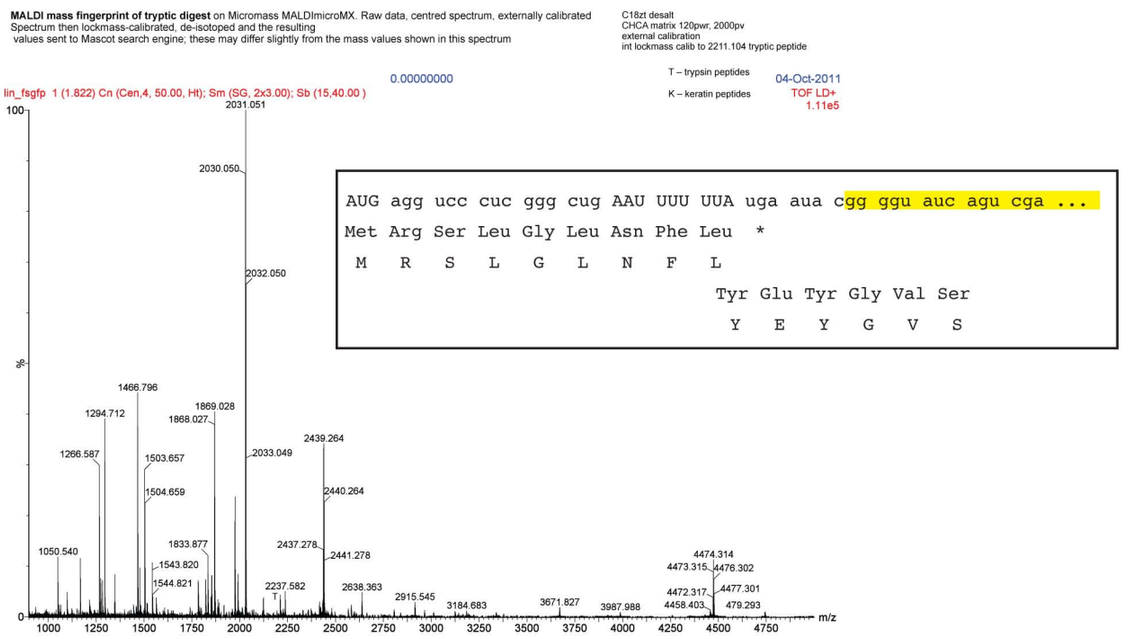

A
pFSHIV-SRV PK frames
FS region
+1 FS
28 kDa
-1 FS
47kDa
nFS
21 kDa
3'
A
C•G
G•C
G•C
G•C
G•C
U•A
A U U C
C•G
C•G
G•C
G•C
G•C
G•C
G
A
A
C
G A A A
pFSHIV-SRV PK
5' UUUUUUA
3-9 nt
CGA
CGAA
CGAGA
CGACGA
CGACAGA
CGACACGA
CGACACGAA
Spacer sequence
B
pFSHIV-SRV PK
M
3
4
5
6
7
8
9
IBV PK
97
46
30
14
— -1 FS
— -2 FS*
— nFS
(-1) %FS: 6 8 10 13 19 25 18 47
(-2) %FS: 7 4 3 14 21 11 7 <1

Figure S3

A



B

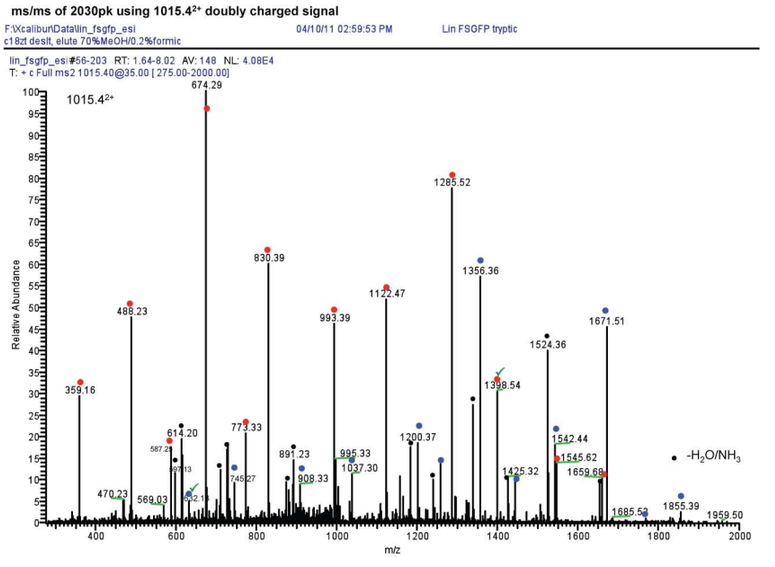
```
1   MRSLGLNFLY EYGVSVEPSR LGWYPGIHRP VATSVSKGEE LFTGVVPILV ELDGDVNGHK FSVSGEGEGD ATYGKLTLKF
81  ICTTGKLPVP WPTLVTTLTY GVQCFSRYPD HMKQHDFFKS AMPEGYVQER TIFFKDDGNY KTRAEVKFEG DTLVNRIELK
161 GIDFKEDGNI LGHKLEYNYN SHNVYIMADK QKNGIKVNFK IRHNIEDGSV QLADHYQQNT PIGDGPVLLP DNHYLSTQSA
241 LSKDPNEKRD HMVLLEFVTA AGITLGMDEL YK
```

red = match by mass only

The matched peptides cover **77%** (210/272AA's) of the protein.

m/z Submitted	MH+ Matched	Intensity	Delta ppm	Modifications	Start	End	Missed Cleavages	Sequence
804.3860	804.3675	14000.0	23.0	1Gln->pyro-Glu	114	119	0	(K)QHDFFK(S)
821.4120	821.3941	14000.0	21.8		114	119	0	(K)QHDFFK(S)
1050.5410	1050.5214	22030.0	18.6		148	156	0	(K)FEGDTLVNR(I)
1266.5891	1266.5783	66900.0	8.52		120	130	0	(K)SAMPEGYVQER(T)
1282.5758	1282.5732	16900.0	2.01	1Oxidation	120	130	0	(K)SAMPEGYVQER(T)
1347.6638	1347.6579	18140.0	4.35		131	141	1	(R)TIFFKDDGNYK(T)
1477.7775	1477.7645	24810.0	8.77		144	156	1	(R)AEVKFEGDTLVNR(I)
1503.6611	1503.6598	74930.0	0.869		61	75	0	(K)FSVSGEGEGDATYGK(L)
1542.8223	1542.7911	31720.0	20.2		161	174	1	(K)GIDFKEDGNILGHK(L)
1868.0270	1868.0177	133300.0	4.96		21	37	0	(R)LGWYPGIHRPVATSVSK(G)
1973.8934	1973.9062	81590.0	-6.48		175	190	0	(K)LEYNYNSHNVYIMADK(Q)
1989.8812	1989.9011	30540.0	-10.0	1Oxidation	175	190	0	(K)LEYNYNSHNVYIMADK(Q)
2030.0496	2030.0229	333000.0	13.1		3	20	0	(R)SLGLNFLYEYGVSVEPSR(L) ← mass matched
2437.2780	2437.2609	120500.0	7.00		38	60	0	(K)GEELFTGVVPILVELDGDVNGHK(F)
2566.3176	2566.2932	3174.0	9.52		250	272	0	(R)DHMVLLEFVTAAGITLGMDELYK(-)
2582.3013	2582.2881	6546.0	5.12	1Oxidation	250	272	0	(R)DHMVLLEFVTAAGITLGMDELYK(-)
4472.3110	4472.1753	54600.0	30.3		203	243	0	(R)HNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSK(D)
4741.5360	4741.3605	3245.0	37.0		201	243	1	(K)IRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSK(D)

C



D

	b			y
	---	1	S 18	---
201.1234	2	L	17	1942.9909
258.1448	3	G	16	1829.9068
371.2289	4	L	15	1772.8854
485.2718	5	N	14	1659.8013
632.3402	6	F	13	1545.7584
745.4243	7	L	12	1398.6900
908.4876	8	Y	11	1285.6059
1037.5302	9	E	10	1122.5426
1200.5936	10	Y	9	993.5000
1257.6150	11	G	8	830.4367
1356.6834	12	V	7	773.4152
1443.7155	13	S	6	674.3468
1542.7839	14	V	5	587.3148
1671.8265	15	E	4	488.2463
1768.8792	16	P	3	359.2037
1855.9113	17	S	2	262.1510
---	18	R	1	175.1190