Supplemental Information:

Structural analysis of the lncRNA SChLAP1 reveals protein binding interfaces and a conformationally heterogenous retroviral insertion

James P. Falese[†,1], Emily J. McFadden[†,1], Christopher A. D'Inzeo[2], and Amanda E. Hargrove*[1,2,3]

Affiliations

1. Department of Biochemistry, Duke University School of Medicine, Durham, North Carolina, 27710, USA
2. Department of Chemistry, Duke University, Durham, North Carolina, 27708, USA
3. Department of Chemistry, University of Toronto, Mississauga, Ontario, L5L1C6, Canada

†: These authors contributed equally to this work.
*To whom correspondence should be addressed: amanda.hargrove@utoronto.ca

Supplementary Text 1

Pseudoknot prediction in SChLAP1

We additionally processed our SHAPE data for pseudoknot formation using SHAPEKnots (Hajdin et al. 2013). One long-range pseudoknot (nucleotides 429-434 and 599-604, 165 nucleotides between stems) was predicted from our *in vitro* data (Fig. S6A). A pseudoknot of similar 5′ coordinates was predicted in our *ex cellulo* data (nucleotides 416-420 and 644-648, 224 nucleotides between predicted stems) (Fig. S6A). However, for both predicted pseudoknots, the SHAPE reactivities of these predicted base-paired nucleotides was significantly increased in our *in cellulo* data and was thus incompatible with pseudoknot formation in cells (Fig. S6B and S6C). We thus did not constrain our any of our structure models herein to contain pseudoknots.

Supplementary Text 2

Commentary on SChLAP1:HNRNPL Interactions

HNRNPL was previously identified as a SChLAP1-binding protein in glioma, and exon 2 was found to be sufficient for binding by biotin pulldowns (Ji et al 2019). While exon 2 lacks the consensus CA-repeat sequence preference for HNRNPL previously identified by RIP-seq in LNCaP cells (Fei et al 2017), we noted a separate CA-rich ΔSHAPE site within exon 2 that we hypothesized could facilitate binding (Fig. 4). Our Western Blot data suggests that the E2-E5 junction binds HNRNPL *in vitro*. Upon closer inspection of our ΔSHAPE data (Fig. 8C), we realized that the ΔSHAPE within the E2-E5 junction overlaps with an ACACA motif, i.e. the consensus sequence for HNRNPL. However, this motif is present within exon 5, rather than exon 2 of SChLAP1. Exon 5 alone was not sufficient for recovering SChLAP1 in the previous study (Ji et al 2019, referred to as exon 3 in their work). We hypothesize several possible explanations for this disparity. One possibility is that there are multiple HNRNPL binding sites within SChLAP1. A second possibility is that RNA structural context is important for characterizing HNRNPL interactions. Pulldowns of exon 2 or exon 5 alone, by definition, will lack the E2-E5 junction, and

these exons in isolation may form structures that alter HNRNPL binding. We expect that additional studies will more specifically characterize the binding sites for HNRNPL within SChLAP1.

A

±	score	% div.	% del.	% ins.	query sequence	position in query- begin	end	(left)	C	matching repeat	class/family	-position in repeat- (left) begin	end	begin (left)	linkage id/graphic
±	2699	4.3	0.0	0.0	SChLAP1	1	322	(1114)	+	LTR12C	LTR/ERV1	1258	1579	(0)	1
±	16	0.0	0.0	0.0	SChLAP1	1088	1104	(332)	+	(A)n	Simple_repeat	1	17	(0)	2
±	1973	11.2	3.2	3.5	SChLAP1	1123	1436	(0)	+	THE1B	LTR/ERVL-MaLR	1	313	(51)	3

B

```
SChLAP1                1 GCTTTTATGAGCTGTAACACTCACCGCGAAGGTCCGCAGCTTCACTCCTG 50
                           i                                 i
LTR12C#LTR/ER       1258 GCCTTTATGAGCTGTAACACTCACCGCGAAGGTCTGCAGCTTCACTCCTG 1307

SChLAP1               51 AAGCCAGCGAGACCACGAGCCTACTGGGAGGAACGAACAACTCCCGACGC 100
                                              i  i                     v
LTR12C#LTR/ER       1308 AAGCCAGCGAGACCACGAGCCCACCGGGAGGAACGAACAACTCCAGACGC 1357

SChLAP1              101 GCCGCCTTAAGAGCTGTAACACTCACCGCGAAGGTCTGCAGCTTCACTCC 150

LTR12C#LTR/ER       1358 GCCGCCTTAAGAGCTGTAACACTCACCGCGAAGGTCTGCAGCTTCACTCC 1407

SChLAP1              151 TGAGCCAGCGAGACCACGAACCCACCAGAAGGAAAAAACTCCGAACACAT 200
                                                       i
LTR12C#LTR/ER       1408 TGAGCCAGCGAGACCACGAACCCACCAGAAGGAAGAAACTCCGAACACAT 1457

SChLAP1              201 CTGAACATCAGAAGCAACAAACTCCGGACACGCCGCCTTTAAGAACTGTA 250
                          i            v             i             i
LTR12C#LTR/ER       1458 CCGAACATCAGAAGGAACAAACTCCGGACGCGCCGCCTTTAAGAGCTGTA 1507

SChLAP1              251 ACACTCACTGCGAGGGTCCGCGGCTTCATTCTTGAAGTGAGTGAGACCAA 300
                                 i                                 v
LTR12C#LTR/ER       1508 ACACTCACCGCGAGGGTCCGCGGCTTCATTCTTGAAGTCAGTGAGACCAA 1557

SChLAP1              301 GAACCCACCAGTTCTGGACACA 322
                                  i   i
LTR12C#LTR/ER       1558 GAACCCACCAATTCCGGACACA 1579
```

C

```
SChLAP1             1123 TGATATGGTTTGGCTGTGTCCCCACCCAAATATCATCTTGAATTGTAGCT 1172
                                                      v
THE1B#LTR/ERV          1 TGATATGGTTTGGCTGTGTCCCCACCCAAATCTCATCTTGAATTGTAGCT 50

SChLAP1             1173 CCCATAATTCCCACGTGTTGTGGGAGGGACCCGGTGGGAGATAATTGTAT 1222
                                      i                     i      v
THE1B#LTR/ERV         51 CCCATAATTCCCACGTGTCGTGGGAGGGACCCGGTGGGAGGTAATTGAAT 100

SChLAP1             1223 CATGGGGGTGGTTC--CCCCATACTATTCTCATAGTAGTGAATAAGTCTC 1270
                             i  v  --i   i i  i     i ii
THE1B#LTR/ERV         101 CATGGGGGCGGGTCTTTCCCGTGCTGTTCTCGTGATAGTGAATAAGTCTC 150

SChLAP1             1271 ACAAAATCTGATGGTTTTATGAGGGAAAACCCCTTTCACCTGGTTCTCAT 1320
                           i i               i i  ii i-----   v    --- v  ?
THE1B#LTR/ERV         151 ACGAGATCTGATGGTTTTATAAAGGGGAG-----TTCCCCTG---CACAN 192

SChLAP1             1321 TCTCTTCTCTGGTCTGTCGTCATGTAAGACATGCCTTT----CACCTTC- 1365
                           v ---     v i   i  i          i  ?    ---- v     -
THE1B#LTR/ERV         193 GC---TCTCTTGCCTGCCGCCATGTAAGACGTGNCTTTGCTCCTCCTTCG 239

SChLAP1             1366 ---TCCACCATGACTGTGAGGCCTCCCCAGCCACGTGGAACTGTGAGCCC 1412
                           ---   i      i                              i
THE1B#LTR/ERV         240 CCTTCCGCCATGATTGTGAGGCCTCCCCAGCCACGTGGAACTGTGAGTCC 289

SChLAP1             1413 ATTAAACCTCTTTCACTTATAAAT 1436
                                 vi
THE1B#LTR/ERV         290 ATTAAACCTCTTTCCTTTATAAAT 313
```

Figure S1. RepeatMasker analysis of SChLAP1 Iso. 1. A) Complete results table of RepeatMasker analysis of SChLAP1 (condensed version shown in Fig. 1C). B) LTR12C alignment of SChLAP1. C) THE1B alignment of SChLAP1.

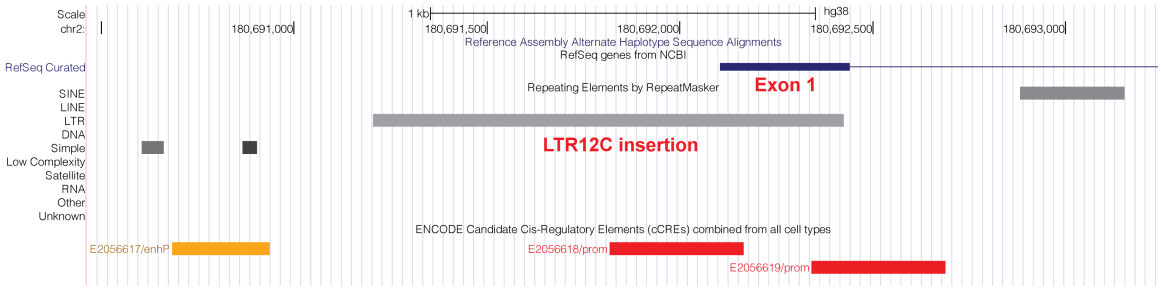


Figure S2. UCSC genome browser (Kent et al. 2002) display of SChLAP1 promoter and exon 1. Position of LTR12C insertion and exon 1 is indicated.

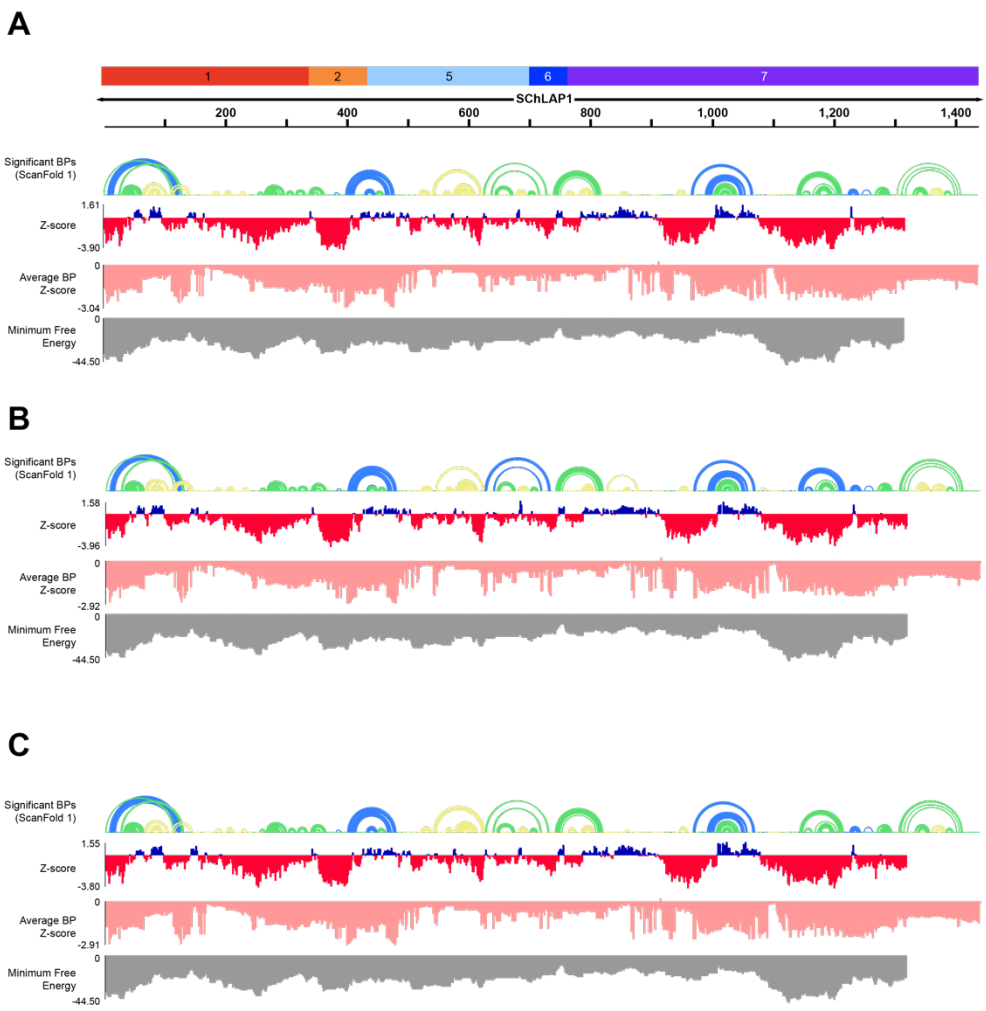


Figure S3. ScanFold 1 analysis of SChLAP1. The SChLAP1 Isoform 1 sequence was analyzed in 3 independent runs (A-C) of the ScanFold 1 algorithm (Andrews et al. 2018). Arcs are color coded as in Figure 2. For Z-scores, red indicates negative values, while blue indicates positive values, for the 120-nucleotide window centered around the given nucleotide.

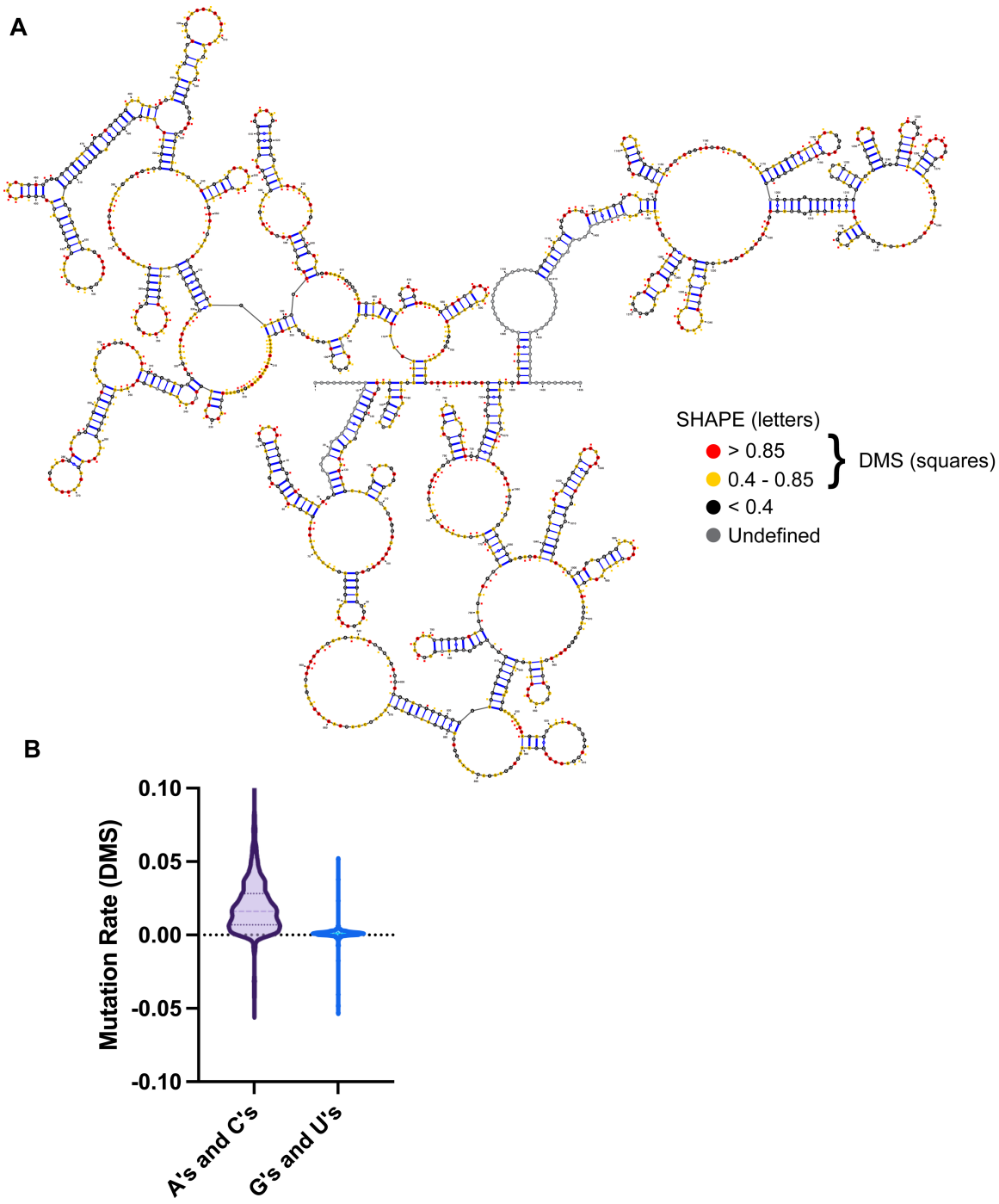


Figure S4. MFE structure model of *in vitro* SChLAP1. A) SHAPE-informed MFE structure of *in vitro* SChLAP1 generated through semi-native purification. DMS reactivities are overlaid over SHAPE-informed structure. MFE structure was generated in RNAstructure (Reuter and Mathews 2010) and visualized in VARNA (Darty et al. 2009). Arc diagram for *in vitro* data shown in Fig. S5.

B) Mutation rate of adenosines (A's) and cytidines (C's) versus guanosines (G's) and uridines (U's) from DMS probing.

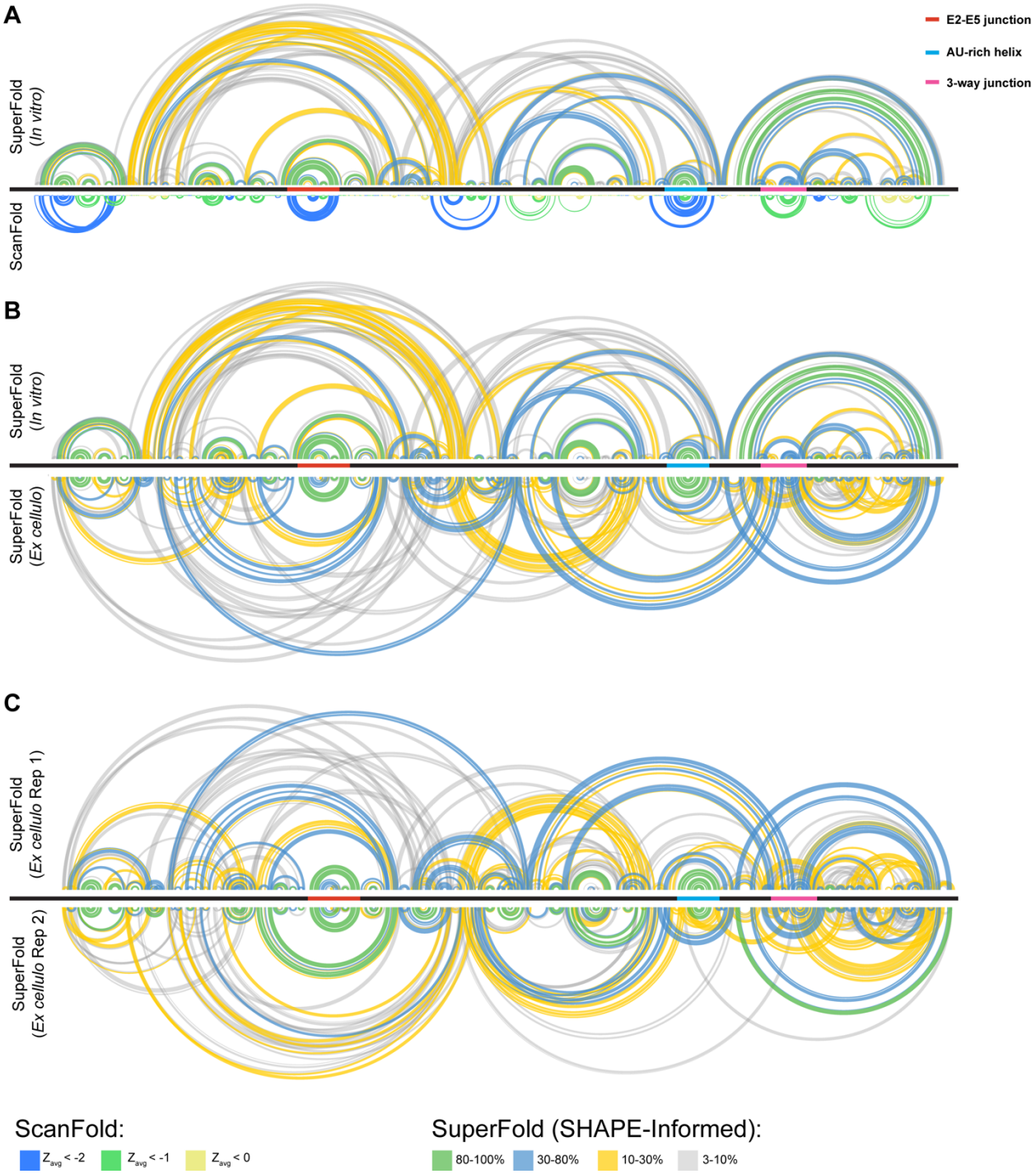


Figure S5. Comparison of arc diagrams between A) *in vitro* SHAPE-probed SChLAP1 and ScanFold-prediction, B) *in vitro* SHAPE probed and representative replicate of *ex cellulo* SHAPE

probing (same as Fig. 3), and C) both replicates of *ex cellulo* SHAPE probing. Positions of E2-E5 junction, AU-rich helix, and 3WJ are indicated with colored bars. Figure generated in Integrative Genomics Viewer (Robinson et al. 2011).

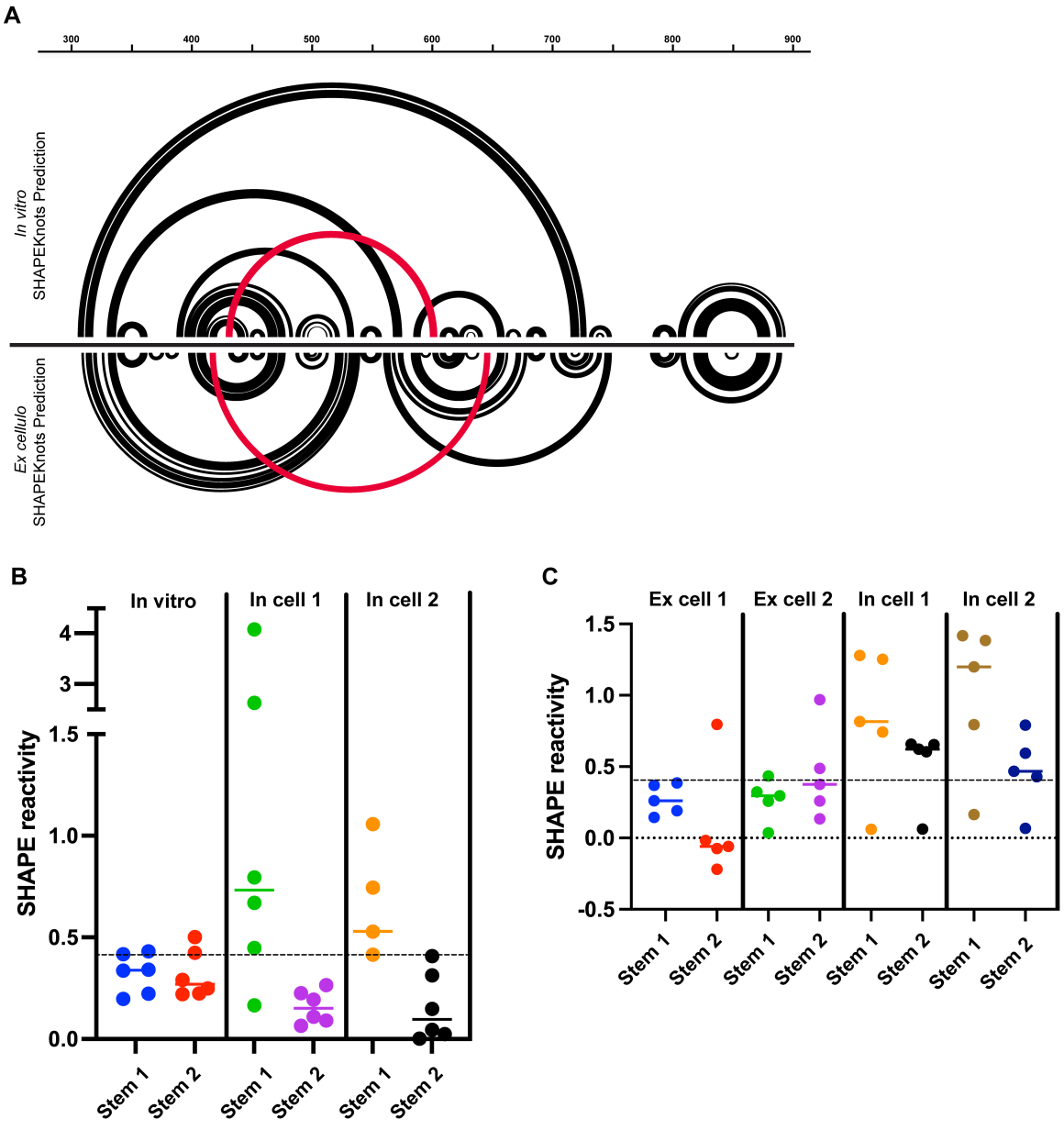


Figure S6. SHAPEknots analysis of SChLAP1. A) Representative arc diagrams of SChLAP1 prediction window generated through SHAPEknots (Hajdin et al. 2013). Red arcs denote predicted pseudoknots in *in vitro* and *ex cellulo* datasets. Figure generated in IGV (Robinson et

al. 2011). B) SHAPE reactivities for the *in vitro*-predicted pseudoknot and the same nucleotides in *in cellulo* replicates 1 and 2, divided by stem of the predicted pseudoknot. Colored lines denote median. A line at SHAPE = 0.4 is drawn for reference as this reactivity is where base-pairing is no longer favorable (Hajdin et al. 2013). For *in cellulo* replicate 2, stem 1, one nucleotide of undefined reactivity was not plotted, and a second nucleotide with outlier SHAPE reactivity (see Methods) was also not plotted, although it was maintained for SHAPEknots prediction and calculation of the median. C) SHAPE reactivities for the *ex cellulo* predicted pseudoknot for both *ex cellulo* and *in cellulo* replicates, divided by putative stem as in panel B.

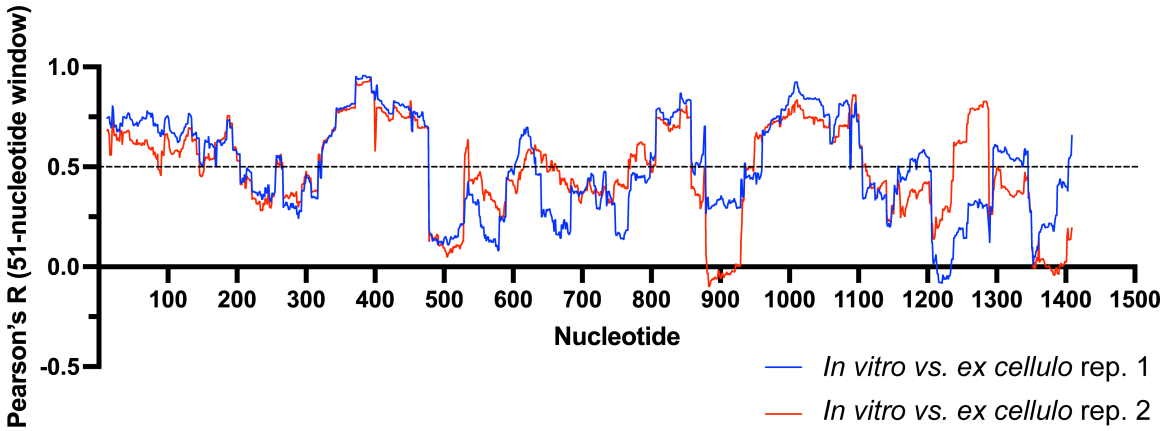


Figure S7. Correlation between *ex cellulo* and *in vitro* SChLAP1. Pearson correlation coefficients were calculated in 51-nucleotide sliding windows (25 nucleotides on each side of a central nucleotide). Correlation windows were truncated at the 5′ and 3′ ends until no fewer than 10 nucleotides were included in the calculation. A dashed line at R = 0.5 is drawn for reference.

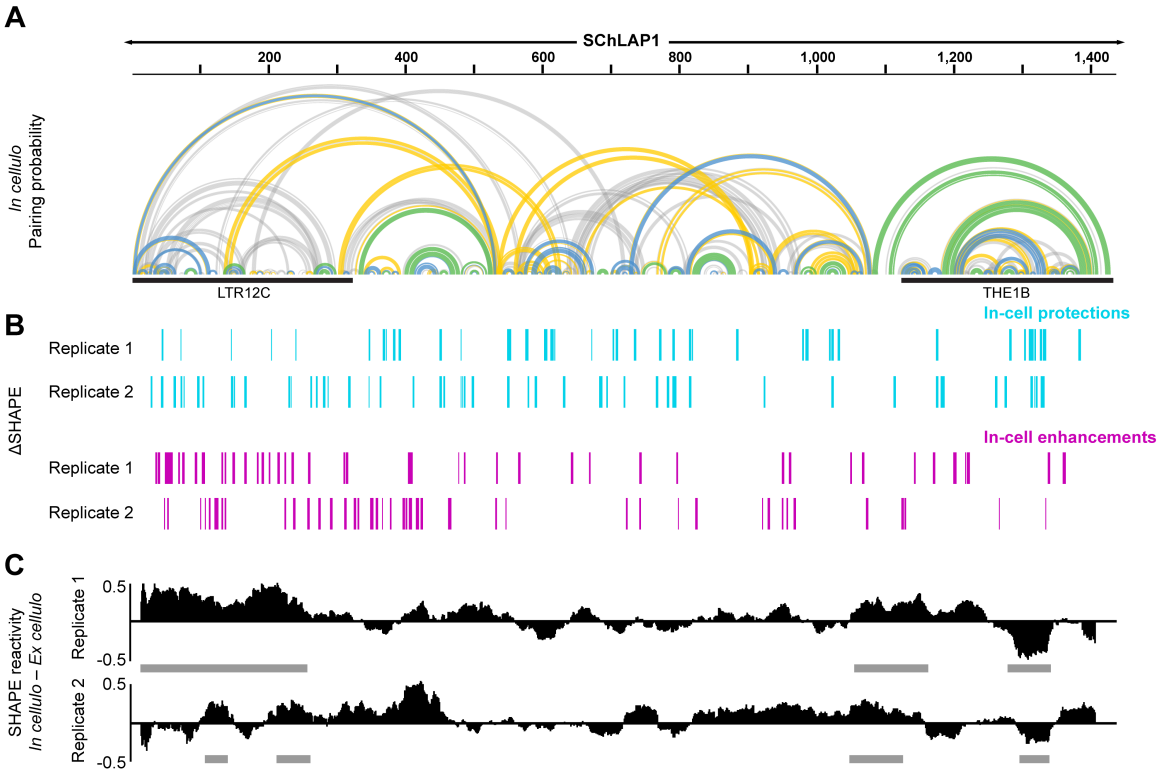


Figure S8. *In cellulo* SHAPE and comparison to *ex cellulo* models. A) Arc diagram of *in cellulo* probed SChLAP1, representative of two biological replicates. Color scheme is the same as Fig. 3A) Location of in-cell protections and in-cell enhancements from each replicate of ΔSHAPE analysis. Overlapping regions between replicates were merged and displayed in Figure 4 (see Methods). C) Difference plot of SHAPE reactivities between each replicate of *in cellulo* and *ex cellulo* reactivities. Reactivities were smoothened over 51-nucleotide windows (see Methods). Positive values indicate regions that are more reactive *in cellulo*. Regions of significant change (see Methods) overlapping between both replicates are indicated by grey bars. Figure generated in Integrative Genomics Viewer (Robinson et al. 2011).

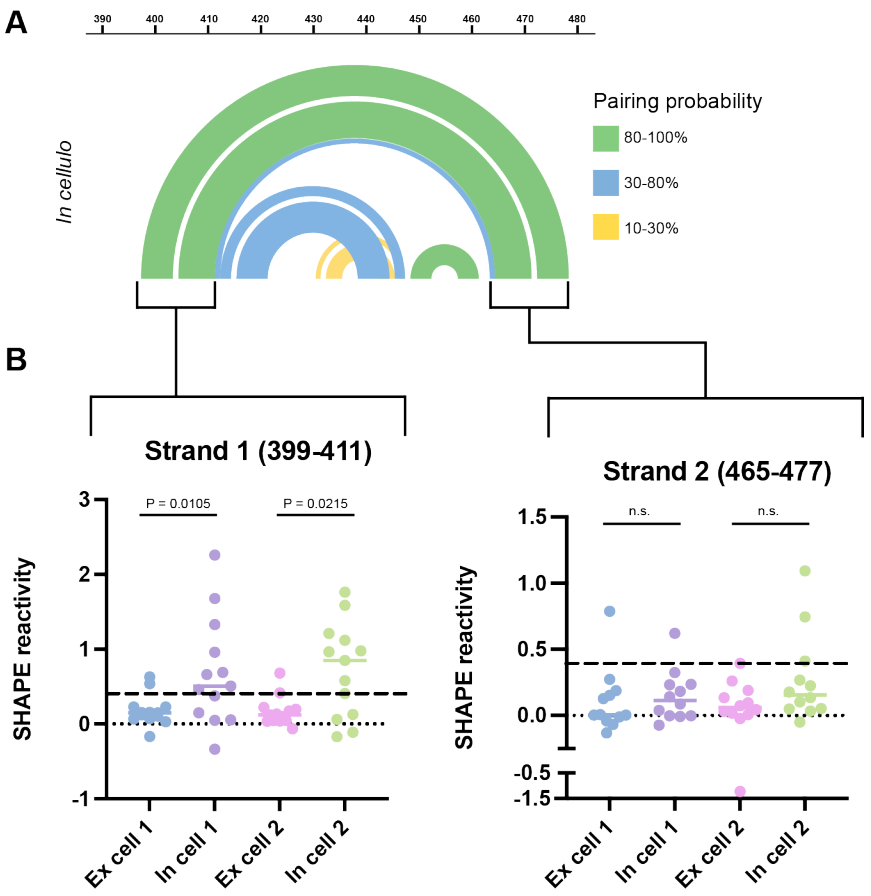


Figure S9. Analysis of SHAPE reactivities of the E2-E5 junction in cells. A) SHAPE-informed arc diagram of the E2-E5 junction (nucleotides 398-478) from a representative replicate of *in cellulo* probing. Pairing probabilities less than 10% are not depicted. B) Analysis of SHAPE reactivities of the largest stem within in this structure, separated by Strand 1 (399-411) and Strand 2 (465-477). Nucleotides 398, 412, 464, and 478 are ignored in this analysis as they are located at the end of helices and thus may have higher SHAPE reactivities, i.e. only internal positions were considered. Statistical comparisons were performed with a Wilcoxon t-test (comparing *ex cellulo* 1 with *in cellulo* 1 and *ex cellulo* 2 with *in cellulo* 2, as was performed with ΔSHAPE) with a cutoff of 0.05 for significance.

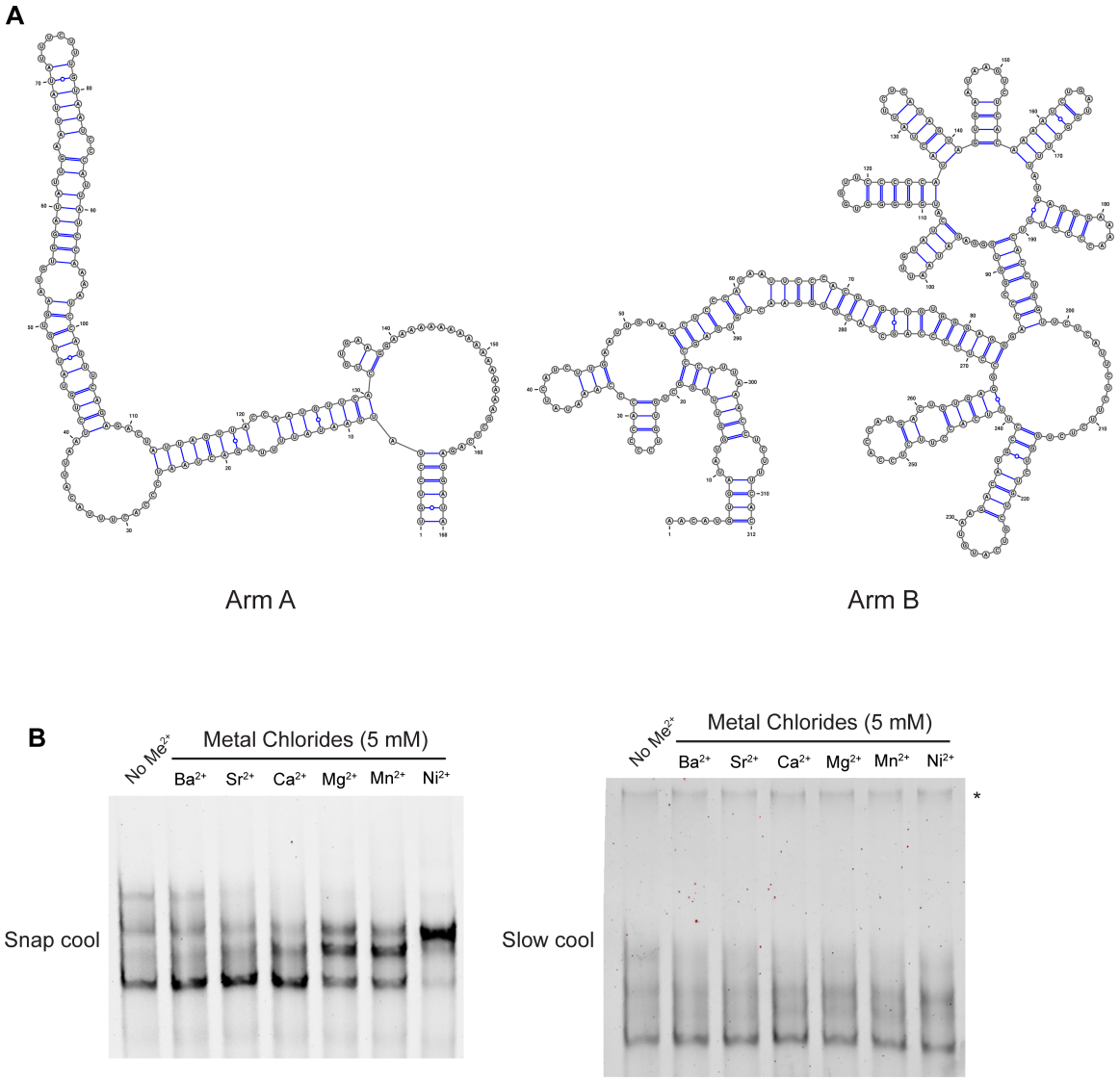
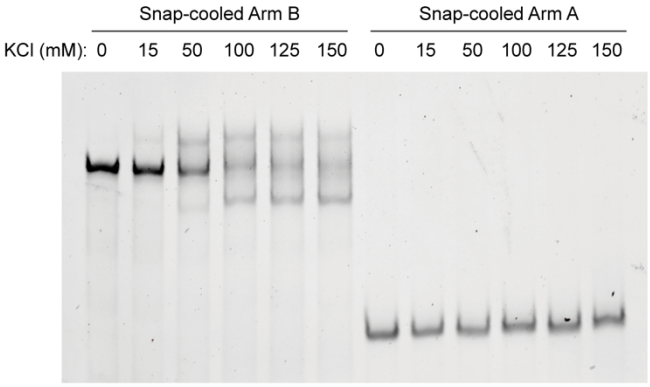


Figure S10. Characterization of Arm A and Arm B by native gel electrophoresis. A) MFE structures of Arm A (left) and Arm B (right) calculated in RNAstructure (Bellaousov et al. 2013) and visualized in VARNA (Darty et al. 2009). B) Native gel electrophoresis of kit-purified Arm B that was either snap cooled (left) or slow cooled (right) and thereafter incubated with the given divalent cations. Arm B was generated through the “kit purified” method. Each gel was performed in triplicate. Asterisks denote dimer, which was typically favored in slow cooling.

A



B

Maximum length = 30 nucleotides

Position	Length	QGRS
449	30	GGACACAAGGUCCACCUUCCAGGAAUAUGG
638	19	GGAGCGGGAUGGAGAAAGG
719	30	GGAGAUGGUGAACCCAAGAGGCUCUGCUGG
1195	16	GGAGGGACCCGGUGGG

Maximum length = 45 nucleotides

Position	Length	QGRS
447	32	GGGGACACAAGGUCCACCUUCCAGGAAUAUGG
638	19	GGAGCGGGAUGGAGAAAGG
719	30	GGAGAUGGUGAACCCAAGAGGCUCUGCUGG
786	40	GGUGAAUAUCAGCACUGGGACCAGUUAGCAGAGGAAAAGG
1194	35	GGGAGGGACCCGGUGGGAGAUAAUUGUAUCAUGGG

C

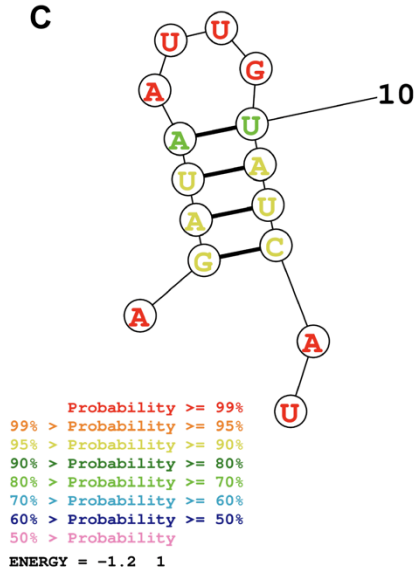
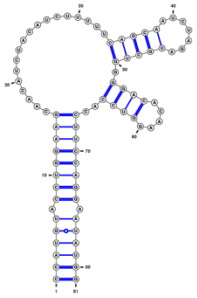


Figure S11. G-quadruplex analysis of Arm B. A) Native gel electrophoresis of kit-purified Arm A and Arm B in varying KCl concentrations. Gel is representative of three biological replicates. B) QGRS Mapper (Kikin et al. 2006) prediction of SChLAP1 with 30 nucleotides (left) and 45

nucleotides (right) set as the maximum allowed G-quadruplex length. Blue letters denote putative G-quadruplex forming regions, and green letters denote predicted hairpin shown in Panel C. C) Structure of predicted hairpin generated in RNAstructure (Bellaousov et al. 2013).

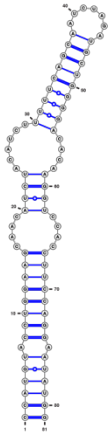
E2-E5 junction
Structure 1

-28.7



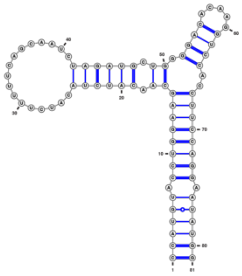
Structure 2

-28.6



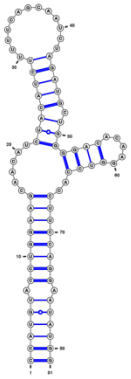
Structure 3

-28.1



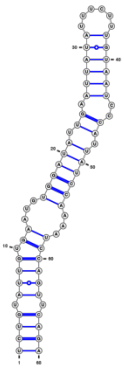
Structure 4

-27.3



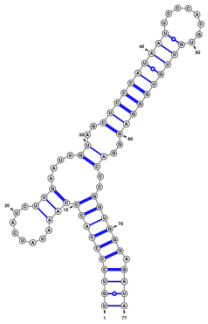
AU-rich helix
Structure 1

-17.9



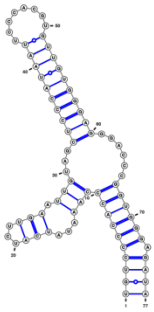
3WJ
Structure 1

-27.3



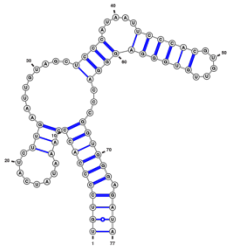
Structure 2

-26.4



Structure 3

-25.6

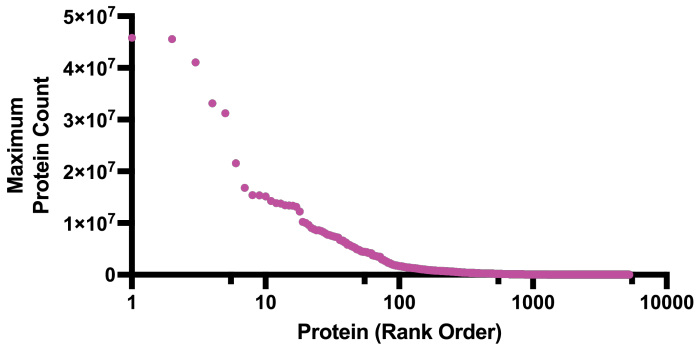


CU-strand
Structure 1 (MaxExpect)

2.0

Figure S12. Structure models of pulldown fragments. MFE structures of E2-E5 junction (4 predicted), AU-rich helix (1 predicted) and 3WJ (3 predicted) are shown. The calculated energies in kcal/mol are shown above each structure. Structures were caluclated using the RNAstructure webserver (Bellaousov et al. 2013) and visualized in VARNA (Darty et al. 2009). No MFE structure was predicted for the CU-strand, consistent with its hypothesized single-stranded character. However, a MaxExpect structure was predicted, which is shown above.

A



B

Term ID	Term Description	False Discovery Rate
GO:0003723	RNA binding	7.28E-72
GO:0003729	mRNA binding	1.40E-45
GO:1901363	Heterocyclic compound binding	1.66E-30
GO:0097159	Organic cyclic compound binding	3.55E-30
GO:0003730	mRNA 3-UTR binding	3.36E-24
GO:0036002	pre-mRNA binding	2.66E-15
GO:0005488	Binding	2.23E-11
GO:0003727	Single-stranded RNA binding	2.63E-11
GO:0035925	mRNA 3-UTR AU-rich region binding	1.83E-09
GO:0003677	DNA binding	5.16E-06
GO:0003697	Single-stranded DNA binding	5.29E-06
GO:0008187	Poly-pyrimidine tract binding	5.29E-06
GO:0070717	Poly-purine tract binding	6.78E-06
GO:0035613	RNA stem-loop binding	1.16E-05
GO:0042162	Telomeric DNA binding	1.81E-05
GO:0004386	Helicase activity	2.18E-05
GO:0003724	RNA helicase activity	4.26E-05
GO:0003725	Double-stranded RNA binding	4.26E-05
GO:0061980	Regulatory RNA binding	4.26E-05
GO:0008143	poly(A) binding	5.86E-05
GO:0035198	miRNA binding	0.00023
GO:0097157	pre-mRNA intronic binding	0.00024
GO:0030280	Structural constituent of skin epidermis	0.00035
GO:0017069	snRNA binding	0.00099
GO:0030620	U2 snRNA binding	0.0017
GO:0140098	Catalytic activity, acting on RNA	0.0035
GO:0140640	Catalytic activity, acting on a nucleic acid	0.0045
GO:0003712	Transcription coregulator activity	0.0089

C

		Normalized Score (0-1)				
PG.ProteinDescriptions	PG.ProteinNames	E2-E5	AU-helix	3WJ	CU-strand	Top Target RNA
Far upstream element-binding protein 2	FUBP2_HUMAN	0.00933166	1	0.03848494	0.13917343	AU-helix
Far upstream element-binding protein 1	FUBP1_HUMAN	0.00223474	1	0.02893625	0.28164987	AU-helix
Heterogeneous nuclear ribonucleoprotein K	HNRPK_HUMAN	0.03037555	0.07491982	1	0.12943358	3WJ
Vigilin	VIGLN_HUMAN	0.01306153	0.14154077	0.03595456	1	CU-strand
Poly(rC)-binding protein 2	PCBP2_HUMAN	0.02652506	0.16545538	1	0.08730309	3WJ
Splicing factor U2AF 65 kDa subunit	U2AF2_HUMAN	0.28369717	0.26509298	0.28228458	1	CU-strand
RNA-binding protein 3	RBM3_HUMAN	0.09357177	0.23489459	0.16038806	1	CU-strand
Splicing factor U2AF 35 kDa subunit	U2AF1_HUMAN	0.26617038	0.24342458	0.29275918	1	CU-strand
Heterogeneous nuclear ribonucleoprotein L	HNRPL_HUMAN	1	0.10289705	0.23418155	0.26317334	E2-E5
Cleavage stimulation factor subunit 3	CSTF3_HUMAN	0.0112874	0.22698068	0.04624496	1	CU-strand
TAR DNA-binding protein 43	TADBP_HUMAN	0.00864454	1	0.02031518	0.0288799	AU-helix
Cleavage stimulation factor subunit 1	CSTF1_HUMAN	0.01428133	0.24258205	0.04698942	1	CU-strand
Cold shock domain-containing protein E1	CSDE1_HUMAN	1	0.05555691	0.07937737	0.03875408	E2-E5
Cytotoxic granule associated RNA binding protein TIA1	TIA1_HUMAN	0.03743663	0.20761968	0.06963211	1	CU-strand
Cleavage stimulation factor subunit 2	CSTF2_HUMAN	0.01557426	0.21773099	0.03984374	1	CU-strand
Keratin, type I cytoskeletal 13	K1C13_HUMAN	0.21169121	1	0.27429559	0.18858605	AU-helix
Pinin	PININ_HUMAN	0.12188715	0.21662813	0.15576706	1	CU-strand

		Normalized Score (0-1)			
PG.ProteinDescriptions	PG.ProteinNames	E2-E5	AU-helix	3WJ	CU-strand
ATP-dependent RNA helicase DHX15	DHX15_HUMAN	0.85918521	0.62418318	0.73316612	1
Serine/arginine-rich splicing factor 3	SRSF3_HUMAN	1	0.01975779	0.67468499	0.65931936
ELAV-like protein 1	ELAV1_HUMAN	0.16243942	0.68427507	0.16622816	1

Figure S13. Proteomics analysis and antibody selection. A) Maximum protein count for all proteins identified in proteomics across the four pulldowns. B) Enriched molecular functions by gene ontology of the top 100 proteins (excluding spiked in control proteins, see Methods) with $p < 0.01$. C) Analysis of qualitative proteomics for selecting proteins to follow up on. Top: proteins in the top 100 identified proteins that passed selection criteria. Proteins that were most enriched for CU-strand were not selected for follow up study. Bottom: Example proteins that were not considered for antibody selection due to high abundance in multiple pulldowns.

Table S1. Human and putative primate SChLAP1 sequences found in BLAST search

Name	Organism	Common Name	Accession Number
SChLAP1 isoform 1	*Homo sapiens*	Human	NR_104319.1
SChLAP1 isoform 2	*Homo sapiens*	Human	NR_104320.1
SChLAP1 isoform 3	*Homo sapiens*	Human	NR_104321.1
SChLAP1 isoform 4	*Homo sapiens*	Human	NR_104322.1
SChLAP1 isoform 5	*Homo sapiens*	Human	NR_104323.1
SChLAP1 isoform 6	*Homo sapiens*	Human	NR_104324.1
SChLAP1 isoform 7	*Homo sapiens*	Human	NR_104325.1
PREDICTED: Cercocebus atys uncharacterized LOC105579891 (LOC105579891), ncRNA	*Cercocebus atys*	Sooty Mangabey	XR_001012906.1
PREDICTED: Mandrillus leucophaeus uncharacterized LOC105529207 (LOC105529207), ncRNA	*Mandrillus leucophaeus*	Drill	XR_001004474.1

Name	Organism	Common Name	Accession Number
PREDICTED: Colobus angolensis palliatus uncharacterized LOC105524242 (LOC105524242), ncRNA	*Colobus angolensis palliatus*	Angolan Colobus	XR_001003724.1
PREDICTED: Pan troglodytes uncharacterized LOC107973853 (LOC107973853), ncRNA	*Pan troglodytes*	Chimpanzee	XR_001716344.3
PREDICTED: Hylobates moloch uncharacterized LOC116809510 (LOC116809510), ncRNA	*Hylobates moloch*	Silvery Gibbon	XR_004369869.1
PREDICTED: Nomascus leucogenys uncharacterized LOC105738214 (LOC105738214), ncRNA	*Nomascus leucogenys*	Northern White-Cheeked Gibbon	XR_001113961.2
PREDICTED: Pan paniscus uncharacterized LOC130541288 (LOC130541288), ncRNA	*Pan paniscus*	Pygmy Chimpanzee	XR_008955326.1
PREDICTED: Gorilla gorilla gorilla uncharacterized LOC109025949 (LOC109025949), ncRNA	*Gorilla gorilla gorilla*	Western Lowland Gorilla	XR_008677752.1

Table S2. All DNA sequences.

Sequence Name	Sequence (5'→3'; Sense Strand)
SChLAP1 Iso. 1 template	GCTTTTATGAGCTGTAACACTCACCGCGAAGGTCCGCAGCTTCA CTCCTGAAGCCAGCGAGACCACGAGCCTACTGGGAGGAACGAA CAACTCCCGACGCGCCGCCTTAAGAGCTGTAACACTCACCGCG AAGGTCTGCAGCTTCACTCCTGAGCCAGCGAGACCACGAACCC ACCAGAAGGAAAAAACTCCGAACACATCTGAACATCAGAAGCAA CAAACTCCGGACACGCCGCCTTTAAGAACTGTAACACTCACTGC GAGGGTCCGCGGCTTCATTCTTGAAGTGAGTGAGACCAAGAAC CCACCAGTTCTGGACACAATTTCAAGTCCTCAGGTGCCATCAAT ATTCTGAAAATGGCAGTGATTTTTATTCAACCTGTATAAGGCACT TTCACCATGTACCTGGAAGCAACATCTACATCTTTTTCAGCAATC TAGATGCTGGGGACACAAGGTCCACCTTCCAGGAATATGGCCA TGACACCAGAAATCACAAACATGATGAGAATGGAATGACTGGGG AAGAAGTGCCAGATGCTTCACTTGTAAATGAAGACCCAGCCTCT

Sequence Name	Sequence (5'→3'; Sense Strand)
	GGGGATGCAGATACCACCTCCCTGAAGAAGCTGAATATCTGCA GATAAGTGGAGTTCACCAATGATGAGGAGCGGGATGGAGAAAG GAGGTAGGGAGAGTCATCCAAGGAACATGAGCAACATGTTAAA AGCCAAGTGGTTTAATTTCTGGAGATGGTGAACCCAAGAGGCTC TGCTGGGAGACAACAAAAATAATGAAGAATTGAACCAGAGTCCG GTGAATATCAGCACTGGGACCAGTTAGCAGAGGAAAAGGAAAG AATAAAAGCGAAAAGAATGAAGAGTCATATGATTACCAACTTTTC CTTTTTCATATAAATTGAGTGTATATGGGTCTGGAACAACCTGAA TTTCCATCAAGTCCTGGCTAACCTCATTATGTCCTATGAATATTT TTGACTAATCCCACTTTACATTAATCTGTATTGTGAATGTGGATA TTGAATTATATTTCTTTGTAATCCCATTATCCAAAATCCAGTTCAG AGACTATTAGTTACCAATGTTCACTGTGAAGGAAAAAAAAAAAAA AAAAGCTCAGAGGATAAACATGTGATATGGTTTGGCTGTGTCCC CACCCAAATATCATCTTGAATTGTAGCTCCCATAATTCCCACGTG TTGTGGGAGGGACCCGGTGGGAGATAATTGTATCATGGGGGTG GTTCCCCCATACTATTCTCATAGTAGTGAATAAGTCTCACAAAAT CTGATGGTTTTATGAGGGAAAACCCCTTTCACCTGGTTCTCATT CTCTTCTCTGGTCTGTCGTCATGTAAGACATGCCTTTCACCTTCT CCACCATGACTGTGAGGCCTCCCCAGCCACGTGGAACTGTGAG CCCATTAAACCTCTTTCACTTATAAAT
SChLAP1 Arm A (949-1116)	TGTCCTATGAATATTTTTGACTAATCCCACTTTACATTAATCTGTA TTGTGAATGTGGATATTGAATTATATTTCTTTGTAATCCCATTATC CAAAATCCAGTTCAGAGACTATTAGTTACCAATGTTCACTGTGAA GGAAAAAAAAAAAAAAAAAGCTCAGAGGATA
SChLAP1 Arm B (1117-1428)	AACATGTGATATGGTTTGGCTGTGTCCCCACCCAAATATCATCTT GAATTGTAGCTCCCATAATTCCCACGTGTTGTGGGAGGGACCC GGTGGGAGATAATTGTATCATGGGGGTGGTTCCCCCATACTATT CTCATAGTAGTGAATAAGTCTCACAAAATCTGATGGTTTTATGAG GGAAAACCCCTTTCACCTGGTTCTCATTCTCTTCTCTGGTCTGTC GTCATGTAAGACATGCCTTTCACCTTCTCCACCATGACTGTGAG GCCTCCCCAGCCACGTGGAACTGTGAGCCCATTAAACCTCTTTC AC
SChLAP1 Iso. 1 Amplicon 1	GCTTTTATGAGCTGTAACACTCACCGCGAAGGTCCGCAGCTTCA CTCCTGAAGCCAGCGAGACCACGAGCCTACTGGGAGGAACGAA CAACTCCCGACGCGCCGCCTTAAGAGCTGTAACACTCACCGCG AAGGTCTGCAGCTTCACTCCTGAGCCAGCGAGACCACGAACCC ACCAGAAGGAAAAAACTCCGAACACATCTGAACATCAGAAGCAA CAAACTCCGGACACGCCGCCTTTAAGAACTGTAACACTCACTGC GAGGGTCCGCGGCTTCATTCTTGAAGTGAGTGAGACCAAGAAC CCACCAGTTCTGGACACAATTTCAAGTCCTCAGGTGCCATCAAT ATTCTGAAAATGGCAGTGATTTTTATTCAACCTGTATAAGGCACT TTCACCATGTACCTGGAAGCAACATCTACATCTTTTTCAGCAATC

Sequence Name	Sequence (5'→3'; Sense Strand)
	TAGATGCTGGGGACACAAGGTCCACCTTCCAGGAATATGGCCA TGACACCAGAAATCACAAA
SChLAP1 Iso. 1 Amplicon 2	TGTACCTGGAAGCAACATCTACATCTTTTTCAGCAATCTAGATGC TGGGGACACAAGGTCCACCTTCCAGGAATATGGCCATGACACC AGAAATCACAAACATGATGAGAATGGAATGACTGGGGAAGAAGT GCCAGATGCTTCACTTGTAAATGAAGACCCAGCCTCTGGGGAT GCAGATACCACCTCCCTGAAGAAGCTGAATATCTGCAGATAAGT GGAGTTCACCAATGATGAGGAGCGGGATGGAGAAAGGAGGTAG GGAGAGTCATCCAAGGAACATGAGCAACATGTTAAAAGCCAAGT GGTTTAATTTCTGGAGATGGTGAACCCAAGAGGCTCTGCTGGG AGACAACAAAAATAATGAAGAATTGAACCAGAGTCCGGTGAATA TCAGCACTGGGACCAGTTAGCAGAGGAAAAGGAAAGAATAAAA GCGAAAAGAATGAAGAGTCATATGATTACCAACTTTTCCTTTTTC ATATAAATTGAGTGTATATGGG
SChLAP1 Iso. 1 Amplicon 3	CTGGGACCAGTTAGCAGAGGAAAAGGAAAGAATAAAAGCGAAA AGAATGAAGAGTCATATGATTACCAACTTTTCCTTTTTCATATAAA TTGAGTGTATATGGGTCTGGAACAACCTGAATTTCCATCAAGTC CTGGCTAACCTCATTATGTCCTATGAATATTTTTGACTAATCCCA CTTTACATTAATCTGTATTGTGAATGTGGATATTGAATTATATTTC TTTGTAATCCCATTATCCAAAATCCAGTTCAGAGACTATTAGTTA CCAATGTTCACTGTGAAGGAAAAAAAAAAAAAAAAAGCTCAGAG GATAAACATGTGATATGGTTTGGCTGTGTCCCCACCCAAATATC ATCTTGAATTGTAGCTCCCATAATTCCCACGTGTTGTGGGAGGG ACCCGGTGGGAGATAATTGTATCATGGGGGTGGTTCCCCCATA CTATTCTCATAGTAGTGAATAAGTCTCACAAAATCTGATGGTTTT ATGAGGGAAAAC
SChLAP1 Iso. 1 Amplicon 4	GACCCGGTGGGAGATAATTGTATCATGGGGGTGGTTCCCCCAT ACTATTCTCATAGTAGTGAATAAGTCTCACAAAATCTGATGGTTT TATGAGGGAAAACCCCTTTCACCTGGTTCTCATTCTCTTCTCTG GTCTGTCGTCATGTAAGACATGCCTTTCACCTTCTCCACCATGA CTGTGAGGCCTCCCCAGCCACGTGGAACTGTGAGCCCATTAAA CCTCTTTCACTTATAAAT

Table S3. All primer sequences for PCR and/or RT. [mX] indicates nucleotide with a 2'–OMe modification. N indicates randomization of A, C, G, or T.

Sequence Name	Forward Primer Sequence (5'→3')	Reverse Primer Sequence (5'→3')
SChLAP1 Iso. 1	CTAATACGACTCACTATAGCTTTTATGAGCTGTAACACTCACCGC	[mA][mU]TTATAAGTGAAAGAGGTTTAATGGGCTCACAGTTCC
Arm A	GAAATTAATACGACTCACTATAGTGTCCTATGAATATTTTTGACTAATCCCAC	[mU][mA]TCCTCTGAGCTTTTTTTTTTTTTTTTCCTTCAC
Arm B	GAAATTAATACGACTCACTATAGAACATGTGATATGGTTTGGCTGTGTC	[mG][mU]GAAAGAGGTTTAATGGGCTCACAGTTC
SChLAP1 Iso. 1 Amplicon 1 Blunt end	GCTTTTATGAGCTGTAACACTCACCGC	TTTGTGATTTCTGGTGTCATGGCCATATTCC
SChLAP1 Iso. 1 Amplicon 2 Blunt end	ATGTACCTGGAAGCAACATCTACATCTTTTTCAGC	CCCATATACACTCAATTTATATGAAAAAGGAAAAGTTGGTAATCATATGACTC
SChLAP1 Iso. 1 Amplicon 3 Blunt end	CTGGGACCAGTTAGCAGAGGAAAAGG	GTTTTCCCTCATAAAACCATCAGATTTTGTGAGACTTA
SChLAP1 Iso. 1 Amplicon 4 Blunt end	GACCCGGTGGGAGATAATTGTATCATG	ATTTATAAGTGAAAGAGGTTTAATGGGCTCACAGTTCC
SChLAP1 Iso. 1 Amplicon 1 PCR1	CCCTACACGACGCTCTTCCGATCTNNNNNGCTTTTATGAGCTGTAACACTCACCGC	GACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNTTTGTGATTTCTGGTGTCATGGCCATATTCC
SChLAP1 Iso. 1 Amplicon 2 PCR1	CCCTACACGACGCTCTTCCGATCTNNNNNATGTACCTGGAAGCAACATCTACATCTTTTTCAGC	GACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNCCCATATACACTCAATTTATATGAAAAAGGAAAAGTTGGTAATCATATGACTC

Sequence Name	Forward Primer Sequence (5'→3')	Reverse Primer Sequence (5'→3')
SChLAP1 Iso. 1 Amplicon 3 PCR1	CCCTACACGACGCTCTTCCGATCTNNNNNCTGGGACCAGTTAGCAGAGGAAAAGG	GACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNGTTTTCCCTCATAAAACCATCAGATTTTGTGAGACTTA
SChLAP1 Iso. 1 Amplicon 4 PCR1	CCCTACACGACGCTCTTCCGATCTNNNNNGACCCGGTGGGAGATAATTGTATCATG	GACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNATTTATAAGTGAAAGAGGTTTAATGGGCTCACAGTTCC

Table S4. RNA sequences made by *in vitro* transcription.

Sequence Name	Sequence (5'→3')
SChLAP1 Iso. 1	GCUUUUAUGAGCUGUAACACUCACCGCGAAGGUCCGCAG CUUCACUCCUGAAGCCAGCGAGACCACGAGCCUACUGGG AGGAACGAACAACUCCCGACGCGCCGCCUUAAGAGCUGU AACACUCACCGCGAAGGUCUGCAGCUUCACUCCUGAGCC AGCGAGACCACGAACCCACCAGAAGGAAAAAACUCCGAAC ACAUCUGAACAUCAGAAGCAACAAACUCCGGACACGCCGC CUUUAAGAACUGUAACACUCACUGCGAGGGUCCGCGGCU UCAUUCUUGAAGUGAGUGAGACCAAGAACCCACCAGUUC UGGACACAAUUUCAAGUCCUCAGGUGCCAUCAAUAUUCU GAAAAUGGCAGUGAUUUUUAUUCAACCUGUAUAAGGCAC UUUCACCAUGUACCUGGAAGCAACAUCUACAUCUUUUUC AGCAAUCUAGAUGCUGGGGACACAAGGUCCACCUUCCAG GAAUAUGGCCAUGACACCAGAAAUCACAAACAUGAUGAGA AUGGAAUGACUGGGGAAGAAGUGCCAGAUGCUUCACUUG UAAAUGAAGACCCAGCCUCUGGGGAUGCAGAUACCACCU CCCUGAAGAAGCUGAAUAUCUGCAGAUAAGUGGAGUUCA CCAAUGAUGAGGAGCGGGAUGGAGAAAGGAGGUAGGGA GAGUCAUCCAAGGAACAUGAGCAACAUGUUAAAAGCCAAG UGGUUUAAUUUCUGGAGAUGGUGAACCCAAGAGGCUCUG CUGGGAGACAACAAAAAUAAUGAAGAAUUGAACCAGAGUC CGGUGAAUAUCAGCACUGGGACCAGUUAGCAGAGGAAAA GGAAAGAAUAAAAGCGAAAAGAAUGAAGAGUCAUAUGAUU ACCAACUUUUCCUUUUUCAUAUAAAUUGAGUGUAUAUGG GUCUGGAACAACCUGAAUUUCCAUCAAGUCCUGGCUAAC CUCAUUAUGUCCUAUGAAUAUUUUUGACUAAUCCCACUU UACAUUAAUCUGUAUUGUGAAUGUGGAUAUUGAAUUAUA UUUCUUUGUAAUCCCAUUAUCCAAAAUCCAGUUCAGAGAC UAUUAGUUACCAAUGUUCACUGUGAAGGAAAAAAAAAAAA AAAAAGCUCAGAGGAUAAACAUGUGAUAUGGUUUGGCUG UGUCCCCACCCAAAUAUCAUCUUGAAUUGUAGCUCCCAU AAUUCCCACGUGUUGUGGGAGGGACCCGGUGGGAGAUA AUUGUAUCAUGGGGGUGGUUCCCCCAUACUAUUCUCAUA GUAGUGAAUAAGUCUCACAAAAUCUGAUGGUUUUAUGAG GGAAAACCCCUUUCACCUGGUUCUCAUUCUCUUCUCUGG UCUGUCGUCAUGUAAGACAUGCCUUUCACCUUCUCCACC AUGACUGUGAGGCCUCCCCAGCCACGUGGAACUGUGAGC CCAUUAAACCUCUUUCACUUAUAAAU
Arm A	UGUCCUAUGAAUAUUUUUGACUAAUCCCACUUUACAUUAA UCUGUAUUGUGAAUGUGGAUAUUGAAUUAUAUUUCUUUG UAAUCCCAUUAUCCAAAAUCCAGUUCAGAGACUAUUAGUU

Sequence Name	Sequence (5'→3')
	ACCAAUGUUCACUGUGAAGGAAAAAAAAAAAAAAAAGCU CAGAGGAUA
Arm B	AACAUGUGAUAUGGUUUGGCUGUGUCCCCACCCAAAUAU CAUCUUGAAUUGUAGCUCCCAUAAUUCCCACGUGUUGUG GGAGGGACCCGGUGGGAGAUAAUUGUAUCAUGGGGGUG GUUCCCCCAUACUAUUCUCAUAGUAGUGAAUAAGUCUCA CAAAAUCUGAUGGUUUUAUGAGGGAAAACCCCUUUCACC UGGUUCUCAUUCUCUUCUCUGGUCUGUCGUCAUGUAAGA CAUGCCUUUCACCUUCUCCACCAUGACUGUGAGGCCUCC CCAGCCACGUGGAACUGUGAGCCCAUUAAACCUCUUUCA C

Table S5. Pulldown probe sequences (IDT)

Name	Sequence	Coordinates	Synthesis type
E2-E5 junction	/5Biosg/rCrC rArUrG rUrArC rCrUrG rGrArA rGrCrA rArCrA rUrCrU rArCrA rUrCrU rUrUrU rUrCrA rGrCrA rArUrC rUrArG rArUrG rCrUrG rGrGrG rArCrA rCrArA rGrGrU rCrCrA rCrCrU rUrCrC rArGrG rArArU rArUrG rG	398-478	Ultramer RNA Oligo, 4 nmol
AU-rich helix	/5Biosg/rUrC rUrGrU rArUrU rGrUrG rArArU rGrUrG rGrArU rArUrU rGrArA rUrUrA rUrArU rUrUrC rUrUrU rGrUrA rArUrC rCrCrA rUrUrA rUrCrC rArArA rArUrC rCrArG rUrUrC rArGrA	989-1056	Ultramer RNA Oligo, 4 nmol
3WJ	/5Biosg/rUrG rUrCrC rCrCrA rCrCrC rArArA rUrArU rCrArU rCrUrU rGrArA rUrUrG rUrArG rCrUrC rCrCrA rUrArA rUrUrC rCrCrA rCrGrU rGrUrU rGrUrG rGrGrA rGrGrG rArCrC rCrGrG rUrGrG rGrArG rArUrA	1139-1215	Ultramer RNA Oligo, 4 nmol
CU-strand	/5Biosg/rGrG rUrUrC rUrCrA rUrUrC rUrCrU rUrCrU rCrUrG rGrUrC rU	1312-1335	100 nmol RNA Oligo

References

Andrews RJ, Roche J, Moss WN. 2018. ScanFold: an approach for genome-wide discovery of local RNA structural elements—applications to Zika virus and HIV. *PeerJ* **6:** e6136. doi: 10.7717/peerj.6136

Bellaousov S, Reuter JS, Seetin MG, Mathews DH. 2013. RNAstructure: Web servers for RNA secondary structure prediction and analysis. Nucleic Acids Res **41:** W471–W474. doi: 10.1093/nar/gkt290

Darty K, Denise A, Ponty Y. 2009. VARNA: Interactive drawing and editing of the RNA secondary structure. Bioinformatics **25:** 1974–1975. doi: 10.1093/bioinformatics/btp250

Fei T, Chen Y, Xiao T, Li W, Cato L, Zhang P, Cotter MB, Bowden M, Lis RT, Zhao SG, et al. 2017. Genome-wide CRISPR screen identifies HNRNPL as a prostate cancer dependency regulating RNA splicing. *Proc Natl Acad Sci U S A* **114:** E5207-E5215. doi: 10.1073/pnas.1617467114

Hajdin CE, Bellaousov S, Huggins W, Leonard CW, Mathews DH, Weeks KM. 2013. Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. *Proc Natl Acad Sci U S A* **110:** 5498–5503. doi: 10.1073/pnas.1219988110

Ji J, Xu R, Ding K, Bao G, Zhang X, Huang B, Wang X, Martinez A, Wang X, Li G, et al. 2019. Long noncoding RNA SChLAP1 forms a growth-promoting complex with HNRNPL in human glioblastoma through stabilization of ACTN4 and activation of NF-κB signaling. *Clin Cancer Res* **25:** 6868-6881. doi: 10.1158/1078-0432.CCR-19-0747

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. *Genome Res* **12:** 996-1006. doi: 10.1101/gr.229102

Kikin O, D'Antonio L, Bagga PS. 2006. QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res* **34:** W676-682. doi: 10.1093/nar/gkl253

Reuter JS, Mathews DH. 2010. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC bioinformatics* **11:** 129. doi: 10.1186/1471-2105-11-129

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29:** 24–26. doi: 10.1038/nbt.1754