STAR Protocols

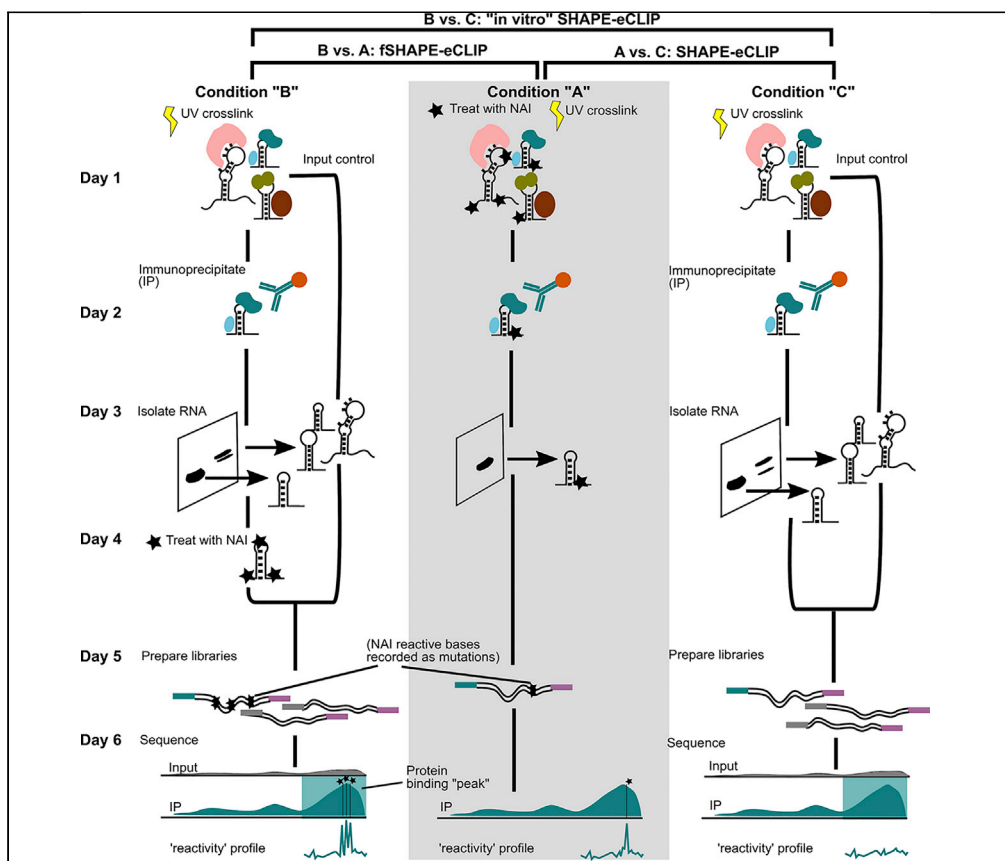## Protocol

fSHAPE, fSHAPE-eCLIP, and SHAPE-eCLIP probe transcript regions that interact with specific proteins

Meredith Corley,
Ryan A. Flynn,
Steven M. Blue,
Brian A. Yee,
Howard Y. Chang,
Gene W. Yeo

meracorley@gmail.com
(M.C.)
geneyeo@ucsd.edu
(G.W.Y.)

Highlights

Protocols for probing
protein-RNA
hydrogen bonds
(fSHAPE) and RNA
structure (SHAPE)

fSHAPE-eCLIP and
SHAPE-eCLIP probe
RNA regions bound by
immunoprecipitated
protein

Samples probed in
parallel under
different conditions is
the basis of each
protocol

Probed regions yield
fSHAPE or SHAPE
reactivities at
nucleotide resolution

Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) structure probing techniques characterize the secondary structure of RNA molecules, which influence their functions and interactions. A variation of SHAPE, footprinting SHAPE (fSHAPE), probes RNA in the presence and absence of protein to identify RNA bases that hydrogen-bond with protein. SHAPE or fSHAPE coupled with enhanced crosslinking and immunoprecipitation (SHAPE-eCLIP or fSHAPE-eCLIP) pulls down RNAs bound by any protein of interest and returns their structure or protein interaction information, respectively. Here, we describe detailed protocols for SHAPE-eCLIP and fSHAPE-eCLIP and an analysis protocol for fSHAPE.

Protocol

fSHAPE, fSHAPE-eCLIP, and SHAPE-eCLIP probe transcript regions that interact with specific proteins

Meredith Corley,[1,6,*] Ryan A. Flynn,[2,3] Steven M. Blue,[1] Brian A. Yee,[1] Howard Y. Chang,[4,5] and Gene W. Yeo[1,7,*]

[1]Department of Cellular and Molecular Medicine, Institute for Genomic Medicine, UCSD Stem Cell Program, University of California San Diego, La Jolla, CA 92093, USA

[2]Stem Cell Program, Boston Children's Hospital, Boston, MA 02115, USA

[3]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA

[4]Center for Personal Dynamic Regulomes, Stanford University School of Medicine, Stanford, CA 94305, USA

[5]Howard Hughes Medical Institute, Stanford University, Stanford, CA 94305, USA

[6]Technical contact

[7]Lead contact

*Correspondence: meracorley@gmail.com (M.C.), geneyeo@ucsd.edu (G.W.Y.)
https://doi.org/10.1016/j.xpro.2021.100762

SUMMARY

Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) structure probing techniques characterize the secondary structure of RNA molecules, which influence their functions and interactions. A variation of SHAPE, footprinting SHAPE (fSHAPE), probes RNA in the presence and absence of protein to identify RNA bases that hydrogen-bond with protein. SHAPE or fSHAPE coupled with enhanced cross-linking and immunoprecipitation (SHAPE-eCLIP or fSHAPE-eCLIP) pulls down RNAs bound by any protein of interest and returns their structure or protein interaction information, respectively. Here, we describe detailed protocols for SHAPE-eCLIP and fSHAPE-eCLIP and an analysis protocol for fSHAPE.

For complete details on the use and execution of these protocols, please refer to Corley et al. (2020).

BEFORE YOU BEGIN

fSHAPE, fSHAPE-eCLIP, and SHAPE-eCLIP are all methods that probe RNA in cells with small RNA-reactive reagents and prepare sequencing libraries from the probed RNA, whose analysis generates "reactivities" at nucleotide resolution. Pairwise analysis comparing the reactivities from samples probed under different conditions produces values that describe the propensity of each nucleobase to hydrogen bond with protein (fSHAPE) or its propensity to be base-paired (SHAPE).

fSHAPE:

To perform fSHAPE, follow the published icSHAPE protocol (Flynn et al., 2016) for both "in vitro" and "in vivo" icSHAPE, at least two replicates each, ignoring the steps for the DMSO-treated controls. Analyze fSHAPE sequencing data according to instructions under the quantification and statistical analysis section of this manuscript.

fSHAPE-eCLIP and SHAPE-eCLIP:

fSHAPE-eCLIP and SHAPE-eCLIP are very similar in execution, having only a few important differences. They will be presented here together with specific instructions where they differ and are referred to jointly as (f)SHAPE-eCLIP.

fSHAPE has been tested in K562, HepG2, HeLa, and HEK293T cells. (f)SHAPE-eCLIP has been tested in K562 cells.

The RNA probing reagent NAI-N$_3$ has been tested with fSHAPE. fSHAPE-eCLIP has been tested with the reagent NAI. SHAPE-eCLIP has been tested with NAI and dimethyl sulfate (DMS); NAI is preferred because DMS does not react with all RNA bases.

(f)SHAPE-eCLIP has been performed with antibodies for SLBP (described here) and RBFOX2. These methods are assumed to be compatible with any of the antibodies successfully employed for eCLIP, of which there are over 300 (Van Nostrand et al., 2020). It is highly recommended to verify your anti-body's immunoprecipitation efficiency before performing this protocol for a full set of samples.

It is recommended to store aliquots of primers at −80°C. Master mixes of enzymatic reactions, minus the enzymes, may be pre-mixed and aliquoted and stored at −80°C for later use.

Prepare primers

⏱ Timing: 1 h

1. Dilute primers to specified concentrations.
   a. RNA oligo "InvRiL19" for RNA linker ligation: Order 100 nmole, standard desalting. Resuspend at 200 μM; dilute to 40 μM for working stock. Store 5 μL aliquots at −80°C.
   b. DNA oligo "InvRand3Tr3" for cDNA ligation: Order 100 nmole, standard desalting. Resuspend at 200 μM; dilute to 80 μM for working stock. Store 5 μL aliquots at −80°C.
   c. DNA oligo "InvAR17" for reverse transcription: Order 25 nmole, standard desalting. Resuspend at 200 μM; dilute to 5 μM for working stock. Store 5 μL aliquots at −80°C.
   d. PCR_F_D501, PCR_F_D502, PCR_R_D701,PCR_R_D702: Order PAGE purified. Resuspend at 100 μM; dilute to 20 μM for working stock. Store 5 μL aliquots at −80°C.
      i. Any Illumina standard high-throughput RNA-sequencing primer, such as D503-508, D703-712, can substitute for the above.

Cell lines

⏱ Timing: days, depending on cell type

2. Grow K562 cells to 90% confluency. Aim for 40 million cells per sample condition. For 2 sample conditions with 2 replicates each at a minimum, this equates to 160 million cells in total.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-stem loop binding protein	MBLI	RRID:AB_10606342; RN045P Lot# 001
Anti-rabbit TrueBlot HRP	Rockland	RRID:AB_2610848; Cat# 18-8816-33
Chemicals, peptides, and recombinant proteins		
Dimethylsulfoxide (DMSO)	Sigma-Aldrich	Cat# 276855
2-Methylnicotinic acid imidazolide (NAI) in DMSO	EMD Millipore	Cat# 03-310
Buffer RLT	QIAGEN	cat # 79216
T4 polynucleotide kinase (PNK), 10 U/μL	New England Biolabs	Cat# M0201L
5X PNK 6.5 buffer	New England Biolabs	Cat# M0201L
1 M DTT	Thermo Scientific	Cat# P2325
100% Ethanol	Thermo Scientific	Cat# T038181000

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
100% Methanol	Thermo Scientific	Cat# T001021000
FastAP Thermosensitive Alkaline Phosphatase, 1 U/μL	Thermo Scientific	Ct# EF0651
Proteinase K, 0.8 U/μL	New England Biolabs	Cat# P8107S
SuperScript II Reverse Transcriptase, 200 U/μL	Invitrogen	Cat# 18064014
Turbo DNase, 2 U/μL	Life Technologies	Cat# AM2239
T4 RNA ligase 1, high concentration	New England Biolabs	Cat# M0437M
10 X ligase buffer (with DTT)	New England Biolabs	Cat# M0437M
100 mM ATP	New England Biolabs	Cat# M0437M
50% PEG 8000	New England Biolabs	Cat# M0437M
Murine RNase Inhibitor, 40 U/μL	New England Biolabs	Cat# M0314L
RNase I, 100 U/μL	Life Technologies	Cat# AM2295
Q5 PCR Master Mix	New England Biolabs	Cat# M0492L
Protease Inhibitor Cocktail III	EMD Millipore	Cat# 539134-1SET
ExoSAP-IT	Affymetrix	Cat# 78201
Tris-HCl buffer, 1 M pH 7.5	Invitrogen	Cat# 15567-027
Tween-20	Sigma-Aldrich	Cat# P1379
HEPES, 1 M pH 8.0	Life Technologies	Cat# 15630-080
Magnesium chloride, 1 M	Ambion	Cat# AM9530G
Sodium chloride, 5 M	Ambion	Cat# AM9759
Potassium chloride, 3M	Millipore Sigma	Cat# 60135
Nonidet P40 (NP-40)	Roche	Cat# 11332473001
EDTA, 500 mM	Millipore	Cat# 324504
HCl, 1.0 N	Sigma-Aldrich	Cat# H9892
NaOH, 1.0 N	Sigma-Aldrich	Cat# S2770
Sodium dodecyl sulfate (SDS) 10%	Sigma-Aldrich	Cat# 71736
Sodium deoxycholate	Sigma-Aldrich	Cat# 30970
20X MOPS Running Buffer	Thermo Fisher	Cat# NP0001
20X NuPAGE Transfer Buffer	Thermo Fisher	Cat# NP00061
10X TBST Buffer	Sigma-Aldrich	Cat# 91414
10X TBE Buffer	Millipore Sigma	Cat# SRE0062
Manganese (II) chloride solution 1M	Sigma-Aldrich	Cat# M1787
Glycerol	Thermo Scientific	Cat# 15514011
Power SYBR Green PCR Master Mix	Applied Biosystems	Cat# 4368577
Dynabeads M-280 Sheep Anti-Rabbit, or anti-mouse, 10 mg/mL	Life Technologies	Cat# 11204D
Dynabeads MyOne Silane, 40 mg/mL	Life Technologies	Cat# 37002D
Agencourt AMPure XP beads	Beckman Coulter	Cat# A63881
ECL Western Blotting Substrate	Thermo Scientific	Cat# 32106
Broad Range Protein Ladder	Thermo Scientific	Cat # 26634
Milk powder	Sigma-Aldrich	Cat# M7409
SYBR Safe DNA Gel Stain	Invitrogen	Cat# S33102
6X Orange G Buffer	New England Biolabs	Cat# B7022S
50 bp DNA Ladder	Thermo Scientific	Cat# SM0373
NuSieve GTG Agarose (low melting temperature)	Lonza	Cat# 50081
Critical commercial assays		
RNA Clean and Concentrator Kit	Zymo Research	Cat# R1019
MinElute Gel Extraction Kit	QIAGEN	Cat# 28604
Deposited data		
Sequencing data and processed fSHAPE and SHAPE reactivities from fSHAPE, fSHAPE-eCLIP, and SHAPE-eCLIP libraries	(Corley et al., 2020)	GEO: GSE149767
Experimental models: Cell lines		
K562	ATCC	CCL-243
Oligonucleotides		
"InvRiL19" (RNA oligo): /5Phos/rArGrArUrCrGrGrArArGrArGr CrArCrArCrGrUrC/3SpC3/, standard desalting	IDT	N/A

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
"InvRand3Tr3": /5Phos/NNNNNNNNNNAGATCGGAAGAGCGTCGTGT/3SpC3/, standard desalting	IDT	N/A
"InvAR17": CAGACGTGTGCTCTTCCGA, standard desalting	IDT	N/A
"PCR_F_D501": AATGATACGGCGACCACCGAGATCTACACTATAGCCTACACTCTTTCCCTACACGACGCTCTTCCGATCT, PAGE purified	IDT	N/A
"PCR_F_D502": AATGATACGGCGACCACCGAGATCTACACATAGAGGCACACTCTTTCCCTACACGACGCTCTTCCGATCT, PAGE purified	IDT	N/A
"PCR_R_D701": CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC, PAGE purified	IDT	N/A
"PCR_R_D702": CAAGCAGAAGACGGCATACGAGATTCTCCGGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC, PAGE purified	IDT	N/A
Software and algorithms		
Cutadapt 1.14	(Martin, 2011)	https://cutadapt.readthedocs.io/en/stable/
UMItools >= 0.5.0	(Smith et al., 2017)	https://github.com/CGATOxford/UMI-tools
STAR >= 2.4.0i	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
Bedtools >= 2.25.0	(Quinlan and Hall, 2010)	https://bedtools.readthedocs.io/en/latest/
Samtools >= 1.9	(Li et al., 2009)	https://samtools.github.io
Bamtools 2.3.0	(Barnett et al., 2011)	https://github.com/pezmaster31/bamtools
fSHAPE analysis pipeline	This manuscript	https://github.com/meracorley/fSHAPE
fSHAPE- and SHAPE-eCLIP analysis pipelines	This manuscript	https://github.com/meracorley/f-SHAPE-eCLIP/tree/pipeline
Python >= 3.5.4	N/A	N/A
Hmmlearn 0.2.1	Scikit-learn	https://hmmlearn.readthedocs.io/en/latest/
Other		
Laminar flow hood	Standard lab equipment	N/A
Cell culture incubator	Thermo Scientific	Cat# 51026331
Large centrifuge	Thermo Scientific	Cat# 1189M64
UV crosslinker	Thermo Scientific	Cat# 95034
Small centrifuge	Eppendorf	Cat# 5406000046
Bioruptor sonicator	Bioruptor	Cat# B01020001
1.5 mL Tube magnet	Thermo Scientific	Cat# CS15000
0.2 mL Tube magnet	EpiCypher	Cat# 10-0008
Tube rotator	Thermo Scientific	Cat# 88881001
Thermomixer	Eppendorf	Cat# 5382000023
PAGE gel tank	Invitrogen	Cat# EI0001
4%–12% Bis-Tris gel, 1.5 mm, 10 well	Invitrogen	Cat# NP0335BOX
Transfer tank	Thermo Scientific	Cat# VEP-2
Power supply for running gels/transfers	Invitrogen	Cat# PS0350
PVDF membrane	Invitrogen	Cat# PB9320
Nitrocellulose membrane	Invitrogen	Cat# PB7320
Agarose gel electrophoresis equipment	Standard lab equipment	N/A
Thermocycler	Standard lab equipment	N/A

MATERIALS AND EQUIPMENT

eCLIP Lysis Buffer (step 5), store at 4°C for up to 6 months

Reagent	Final concentration	Amount
1 M Tris-HCl pH 7.4	50mM	5 mL
5 M NaCl	100 mM	2 mL
NP-40	1%	1 mL
10% SDS	0.1%	1 mL
Sodium deoxycholate	0.5%	0.5 g
diH2O	n/a	91 mL
Total	n/a	100 mL

eCLIP High Salt Wash Buffer (step 10), store at 4°C for up to 6 months

Reagent	Final concentration	Amount
1 M Tris-HCl pH 7.4	50mM	5 mL
5 M NaCl	1 M	20 mL
500 mM EDTA	1 mM	200 μL
NP-40	1%	1 mL
SDS 10%	0.1%	1 mL
Sodium deoxycholate	0.5%	0.5 g
diH2O	n/a	72.8 mL
Total	n/a	100 mL

eCLIP Wash Buffer (step 10), store at 4°C for up to 6 months

Reagent	Final concentration	Amount
1 M Tris-HCl pH 7.4	20 mM	2 mL
5 M NaCl	5 mM	100 μL
1 M MgCl$_2$	10 mM	1 mL
1% Tween-20	0.2%	20 mL
diH2O	n/a	76.9 mL
Total	n/a	100 mL

RLTW Buffer (step 38), store at RT for up to 6 months

Reagent	Final concentration	Amount
RLT Buffer (1X)	1X	9.75 mL
1% Tween-20	0.025%	250 μL
Total	n/a	10 mL

PKS Buffer (step 28), store at RT for up to 6 months

Reagent	Final concentration	Amount
1 M Tris-HCl pH 7.4	100 mM	1 mL
5 M NaCl	50 mM	100 μL
500 mM EDTA	10 mM	200 μL
10% SDS	0.2%	200 μL
diH2O	n/a	8.5 mL
Total	n/a	10 mL

1X RNA Ligase Buffer (step 15), store at RT for up to 6 months

Reagent	Final concentration	Amount
1 M Tris-HCl pH 7.5	50 mM	500 μL
1 M MgCl$_2$	10 mM	100 μL
diH2O	n/a	9.4 mL
Total	n/a	10 mL

3X SHAPE Folding Buffer (step 31), store at RT for up to 6 months

Reagent	Final concentration	Amount
1 M HEPES pH 8.0	333 mM	333 μL
1 M MgCl$_2$	20 mM	20 μL
5 M NaCl	333 mM	66.6 μL
diH2O	n/a	580.4 μL
Total	n/a	1 mL

SHAPE FS Buffer (step 41), store at RT for up to 6 months

Reagent	Final concentration	Amount
1 M Tris-HCl pH 8.0	500 mM	500 μL
3 M KCl	750 mM	250 μL
diH$_2$O	n/a	250 μL
Total	n/a	1 mL

TT Elution Buffer (step 38), store at RT for up to 6 months

Reagent	Final concentration	Amount
1 M Tris-HCl pH 7.5	10 mM	10 μL
500 mM EDTA	0.1 mM	0.2 μL
1% Tween-20	0.01% Tween-20	10 μL
diH$_2$O	n/a	979.8 μL
Total	n/a	1 mL

PCR Elution Buffer (step 48), store at RT for up to 6 months

Reagent	Final concentration	Amount
1 M Tris-HCl pH 7.5	10 mM	10 μL
5 M NaCl	20 mM	4 μL
500 mM EDTA	0.1 mM	0.2 μL
diH$_2$O	n/a	985.8 μL
Total	n/a	1 mL

Enzymatic reactions:

"Early" FastAP treatment (on-bead) (step 13)

Reagent	Final concentration	Amount
diH$_2$O	n/a	38 μL
10X FastAP buffer	1X	5 μL
Murine RNase Inhibitor, 40 U/μL	1.6 U/μL	2 μL
Turbo DNase, 2 U/μL	0.08 U/μL	2 μL
FastAP enzyme, 1 U/μL	0.06 U/μL	3 μL
Total	n/a	50 μL

"Early" PNK treatment (on-bead) (step14)

Reagent	Final concentration	Amount
diH$_2$O	n/a	116 μL
5X PNK Buffer pH 6.5	1X	30 μL
T4 PNK enzyme, 10 U/μL	0.27 U/μL	4 μL
Total	n/a	150 μL

"Early" 3′ RNA linker ligation (on-bead) (step 16)

Reagent	Final concentration	Amount
diH$_2$O	n/a	8.4 μL
10X Ligase buffer (no DTT; self-made, see buffers above)	1X	3 μL
0.1 M ATP	1 mM	0.3 μL
100% DMSO	3%	0.9 μL
1% Tween-20	0.02%	0.6 μL
50% PEG 8000	15%	9 μL
Murine RNase Inhibitor, 40 U/μL	0.53 U/μL	0.4 μL
RNA ligase high conc., 30 U/μL	2.4 U/μL	2.4 μL
Combine above	n/a	**25 μL**

Continued

Reagent	Final concentration	Amount
Beads binding RNA	n/a	~2.5 μL
InvRiL19 oligo, 40 μM	3.33 μM	2.5 μL
Total	n/a	~30 μL

In vitro structure probing (step 31)

Reagent	Final concentration	Amount
RNA sample, heat denatured and flash-cooled	n/a	~10 μL
diH$_2$O	n/a	2.4 μL
3X SHAPE Folding Buffer	1X	6.6 μL
Murine RNase Inhibitor, 40 U/μL	2 U/μL	1.0 μL
Combine above	n/a	20 μL
2 M NAI in DMSO	100 mM	1 μL
Total	n/a	21 μL

⚠ CRITICAL: The reagent NAI should be handled with caution, avoiding contact with skin.

"Late" FastAP treatment (step 34)

Reagent	Final concentration	Amount
diH$_2$O	n/a	6 μL
10X FastAP Buffer	1X	2 μL
Murine RNase Inhibitor, 40 U/μL	2 U/μL	1 μL
FastAP enzyme, 1 U/μL	0.1 U/μL	2 μL
Combine above	n/a	11 μL
RNA sample	n/a	~10 μL
Total	n/a	21 μL

"Late" PNK treatment (step 35)

Reagent	Final concentration	Amount
diH$_2$O	n/a	45 μL
5X PNK Buffer pH 6.5	1X	20 μL
1 M DTT	5 mM	0.5 μL
Turbo DNase, 2 U/μL	0.02 U/μL	1 μL
T4 PNK enzyme, 10 U/μL	0.4 U/μL	4 μL
Combine above	n/a	75 μL
"Late" FastAP reaction	n/a	~21 μL
Total	n/a	~100 μL

"Late" 3′ RNA linker ligation for Input samples (step 37)

Reagent	Final concentration	Amount
diH$_2$O	n/a	2.8 μL
10X NEB Ligase Buffer (with DTT)	1X	2.0 μL
0.1 M ATP	1 mM	0.2 μL
100% DMSO	3%	0.6 μL
1% Tween-20	0.02%	0.4 μL
50% PEG 8000	15%	6.0 μL
Murine RNase Inhibitor, 40 U/μL	0.6 U/μL	0.3 μL
RNA ligase high conc., 30 U/μL	1.8 U/μL	1.2 μL
Combine above	n/a	13.5 μL
RNA sample	n/a	5 μL
100% DMSO	n/a	1.5 μL
InvRiL19 oligo, 40 μM	1 μM	0.5 μL
Total	n/a	20 μL

"Late" 3′ RNA linker ligation for IP samples (step 37)

Reagent	Final concentration	Amount
diH$_2$O	n/a	5.6 µL
10X NEB Ligase buffer (with DTT)	1X	4 µL
0.1 M ATP	1 mM	0.4 µL
100% DMSO	3%	1.2 µL
1% Tween-20	0.02%	0.8 µL
50% PEG 8000	15%	12.0 µL
Murine RNase Inhibitor, 40 U/µL	0.6 U/µL	0.6 µL
RNA ligase high conc., 30 U/µL	1.8 U/µL	2.4 µL
Combine above	n/a	**27 µL**
RNA sample	n/a	~10 µL
100% DMSO	n/a	2.4 µL
InvRiL19 oligo, 40 µM	1 µM	1.0 µL
Total	n/a	**~40 µL**

SHAPE reverse transcription (step 41)

Reagent	Final concentration	Amount
10X SHAPE FS Buffer	1X	2.0 µL
Murine RNase Inhibitor, 40 U/µL	0.4 U/µL	0.2 µL
0.1 M DTT	5 mM	1.0 µL
500 mN MnCl$_2$	6 mM	0.24 µL
Superscript II reverse transcriptase, 200 U/µL	10 U/µL	1.0 µL
diH$_2$O	n/a	5.36 µL
Combine above	n/a	**10 µL**
RNA sample	n/a	~9 µL
5 µM InvAR17 primer	0.25 µM	1 µL
10 mM dNTPs	0.5 mM	1 µL
Total	n/a	**~20 µL**

5′ cDNA linker ligation (step 44)

Reagent	Final concentration	Amount
diH$_2$O	n/a	1.4 µL
10X NEB Ligase buffer (with DTT)	1X	4 µL
0.1 M DTT	2 mM	0.2 µL
0.1 M ATP	1 mM	0.1 µL
1% Tween-20	0.02%	0.2 µL
50% PEG 8000	18%	3.6 µL
RNA ligase high conc., 30 U/µL	3.0 U/µL	1.0 µL
Combine above	n/a	**7.8 µL**
RNA sample in dried beads	n/a	0 µL
100% DMSO	8%	0.8 µL
InvRand3Tr3 oligo, 80 µM	4.8 µM	0.6 µL
TT Elution Buffer	n/a	1.1 µL
Total	n/a	**10.3 µL**

cDNA qPCR quantification (step 46)

Reagent	Final concentration	Amount
diH$_2$O	n/a	3.6 µL
2X PowerSybr master mix	1X	5.0 µL
20 µM D50x primer mixed in equal parts with 20 µM D70x primer	0.4 µM	0.4 µL
1:10 dilution cDNA sample	1:100	1.0 µL
Total	n/a	**10 µL**

cDNA PCR amplification (step 47)

Reagent	Final concentration	Amount
2X Q5 PCR master mix	1X	20 μL
20 μM right primer (D50x)	1 μM	2 μL
20 μM left primer (D70x)	1 μM	2 μL
cDNA sample	n/a	16 μL
Total	n/a	40 μL

Equipment:

Cell culture equipment: laminar flow hood, cell culture incubator

UV crosslinker (254 nm, UV-C)

Bioruptor sonicator, for RNA fragmentation

Centrifuge for 15 mL conicals; Centrifuge for 1.5 mL tubes; minifuge

Magnet for 1.5 mL tubes, 0.2 mL tubes

Tube rotator

Thermomixer

Western blot and RNA transfer equipment: SDS-PAGE gels 4–12% Bis-Tris, 10-well, 1.5 mm (1.0 mm if 1.5 mm not available), buffer tanks for PAGE and for western blots/RNA transfers, power supply for gel electrophoresis, PVDF and nitrocellulose membranes, sponges, roller

Thermocycler

Real time qPCR equipment

Agarose gel electrophoresis equipment

STEP-BY-STEP METHOD DETAILS
Treat cells with NAI or DMSO and UV crosslink

⏲ Timing: 3 h

Treat cells with the structure probing reagent NAI or mock treat with DMSO, followed by UV cross-linking to covalently link RNA binding proteins to bound RNA molecules. Complete these steps for all protocol types: fSHAPE-eCLIP, SHAPE-eCLIP, or "in vitro" SHAPE-eCLIP.

1. Cell sample requirements
   a. Grow cells to ∼30–40 million cells per sample. With two treatment conditions each needing at least two replicates, this adds up to at least four samples, 120–160 million cells total.
   b. fSHAPE-eCLIP and SHAPE-eCLIP both consist of two RNA treatment conditions performed in parallel. The different conditions are referred to here as "A", "B", or "C" (see Graphical Abstract):
      i. fSHAPE-eCLIP consists of one set of cells treated with NAI ("A"), i.e., RNA is treated *in vivo*. The other set is mock-treated with DMSO, but later treated with NAI once its RNA is *in vitro* ("B").

ii. SHAPE-eCLIP consists of one set of cells treated with NAI ("A") and the other mock-treated with DMSO ("C"). An "in vitro" variant of SHAPE-eCLIP consists of two sets of samples that are both mock-treated with DMSO at the cell stage, but one set later treats RNA with NAI *in vitro* ("B").

2. Setup
   a. Pre-chill on ice 100 mL 1X PBS.
   b. Set an incubator to 37°C.
   c. Pre-chill on ice metal plate that fits inside crosslinker. Set crosslinker to 400 mJ/cm2 and 2 min. Crosslinker should emit UV-C light at ~254 nm wavelengths.
   d. Fetch some dry ice.

3. Treat cells with the structure probing reagent (NAI) or mock-treat with DMSO
   a. Split cells into 30–40 million cells per sample, two samples for each treatment condition.Instructions are for suspension cells; see note below on adherent cells.
   b. Reduce each cell sample volume to 2 mL in plain media (no FBS) in 10 cm plates.
   c. Per sample pair: For "A" conditions (in fSHAPE-eCLIP or SHAPE-eCLIP), add 170 uL 2 M NAI structure probing reagent to each sample. For "B" or "C" conditions (in fSHAPE-eCLIP or the "in vitro" variant of SHAPE-eCLIP), add 170 uL DMSO to each sample. Swirl to mix.
   d. Incubate plates at 37°C for 10 min.

4. UV crosslink cells and wash
   a. Place pre-chilled metal plate into crosslinker and place plates on top. The chilled plate prevents cells from heating up during crosslinking.
   b. Remove lids from plates. Run the pre-set crosslinking program.
   c. Move cells from plates to 15 mL conicals. Spin down at 500 *g* for 3 min.
   d. Remove supernatant (NAI-treated supernatant should go into its own waste container).
   e. Resuspend cells in 5 mL cold PBS. Spin down and remove supernatant as before. Resuspend again with cold PBS.
   f. Spin down cells one more time, remove PBS supernatant, and flash freeze cells in dry ice. Store at −80°C.

⚠ CRITICAL: Removing lids from plates before crosslinking is critical to cells being properly crosslinked.

Note: Adherent cells should have media removed and replaced with 2 mL plain media (no FBS) and subsequently treated with 170 uL NAI or DMSO, swirling to mix. Cells should then be crosslinked as instructed, then scraped off plates with a cell scraper and transferred to 15 mL conicals, following the same washing and centrifugation steps outlined for suspension cells.

� Pause point: Cell pellets can be stored at −80°C for months.

Lyse cells and immunoprecipitate RNA

⏱ Timing: 3–16 h

Lyse protein-bound RNA from cells. Pull down RNA bound to the protein of interest with a monoclonal antibody coupled to magnetic beads. Complete these steps for all protocol types.

5. Setup
   a. Pre-chill eCLIP Lysis Buffer.
   b. Thaw Protease Inhibitor.
   c. Bioruptor sonicator water tank should be at 4°C (in a cold room).
   d. Pre-heat Thermomixer to 37°C.
   e. Pre-set 1.5 mL tube centrifuge to 4°C.

6. Couple antibody to Dynabeads M-280
   a. Prepare beads.
      i. Use 125 μL beads per sample, 500 μL for two conditions each with two replicates.
      ii. Apply bead solution to magnet; remove supernatant.
      iii. Wash beads twice in 500 μL cold eCLIP Lysis Buffer.
      iv. Resuspend beads in 100 μL cold eCLIP Lysis Buffer for every sample, typically 400 μL.
   b. Bind antibody.
      i. Add 10 μg antibody per sample to prepared beads.
      ii. Rotate at 15°C–25°C for 45 min.

7. Lyse cells, while beads are rotating
   a. For two replicates—four samples total—combine 4 mL eCLIP Lysis Buffer with 22 μL Protease Inhibitor Cocktail III and 44 μL Murine RNase Inhibitor.
   b. Retrieve cell pellets from −80°C, keeping on dry ice. Immediately add 1 mL master mix from 7a to each pellet, pipetting to resuspend. Move each sample to a 1.5 mL tube.
   c. Place on ice (not dry ice) and allow to lyse for 5 min.

8. Fragment RNA
   a. Sonicate lysed samples in Bioruptor on "low", 4°C, 5 cycles of 30 s on and 30 s off. This assists cell lysis and shears chromatin, aiding release of RNA-protein complexes.

   Note: Alternatively, sonication may be done on ice with a probe sonicator, sonicating each sample twice at 5 decibels in 10 second bursts. Avoid touching the side of tube with sonicator.

   b. Dilute 1 μL RNase I in PBS at 1:50 on ice.
   c. Add 10 μL Turbo DNase to each lysed sonicated sample.
   d. Add 10 μL diluted RNase I (20 U total) to each sample. Immediately incubate in a Thermomixer at 37°C, 1200 rpm, 5 min. Place on ice.
   e. Centrifuge samples at 15,000 g, 4°C, for 3 min.
   f. Meanwhile, proceed with preparing antibody-bound magnetic beads (below).

9. Immunoprecipitate RNA bound by protein of interest
   a. Wash antibody beads from 6b two times with 500 μL in cold eCLIP Lysis Buffer.
   b. Remove as much supernatant as possible from beads and resuspend once more in 100 μL cold eCLIP Lysis Buffer per sample, typically 400 μL.
   c. Distribute beads very well by pipetting up and down (taking tube off of magnet) and divide bead solution evenly between separate 1.5 mL tubes, one tube per sample.
   d. Only once lysates from 8e are finished centrifuging, apply each bead sample to the magnet and remove supernatant. Add cleared lysate for each sample from 8e to its own antibody bead sample.
   e. Rotate at 4°C for 2 h, or up to 15 h in cold room.

   Note: If doing fSHAPE-eCLIP or "in vitro" variant of SHAPE-eCLIP, opt for 2 hours of rotation and skip to "Isolate RNA" section.

Ligate 3′ RNA adapter

⏱ Timing: 3 h

Complete this section only for SHAPE-eCLIP samples at this time in the protocol. Treat RNA and ligate an adapter, later used for reverse transcription.

10. Setup
    a. Verify High Salt Wash Buffer, Wash Buffer, and 1x FastAP Buffers are cold.
    b. Pre-heat Thermomixer to 37°C.

11. Set aside non-immunoprecipitated sample controls, referred to as "Input" samples
    a. Of the sample conditions ("A","B", or "C"; see 1b for naming convention), pick one condition to take sample Input controls from. Typically "B" or "C" conditions. One can carry forward Input controls for all conditions if one chooses, but this creates more controls than is necessary.
    b. Gently mix samples from 9e well. Take **two** 20 μL aliquots from each replicate of the chosen condition, four **aliquots** total, adding each to separate tubes labeled "Transfer-Input" and "Western-Input" and their replicate number. **Set Inputs aside on ice until step 20.**
    c. The remainder of the samples are referred to as immunoprecipitated (IP) samples.
12. Wash antibody-bound beads
    a. Perform the following washes on ice if possible.
    b. Apply all IP samples to magnet. Wait for beads to pellet on the magnet, removing the supernatant when it is clear.
    c. Wash bead pellets with 900 μL **cold** High Salt Wash Buffer by pipetting the buffer onto the bead pellets. Wait until supernatant is clear (∼1 min) and remove supernatant by pipetting. Repeat.
    d. Wash beads 2 times as above with 500 μL cold Wash Buffer. During the second wash, move tubes back and forth on magnet (Methods Video S1). One time is sufficient, such that the bead pellet is forced to re-form. Wait until beads have reformed pellet until removing supernatant, ∼1 min.
    e. Resuspend beads with 500 μL 1X FastAP Buffer, pipetting to completely resuspend beads in buffer.
    f. Allow supernatant to clear and remove supernatant. Add 500 μL 1X FastAP Buffer and place tubes on ice.
13. FastAP treat RNA on beads
    a. Prepare "Early" FastAP master mix on ice (see materials and equipment). Mix well.
    b. Place IP bead samples on magnet, removing supernatant when clear. Add 50 μL FastAP master mix to each IP bead sample.
    c. Incubate samples in Thermomixer at 1200 rpm at 37°C for 10 min.
14. PNK treat RNA on beads
    a. While FastAP reaction is incubating, prepare "Early" PNK master mix.
    b. Mix well and add 150 μL master mix to each sample.
    c. Incubate samples in Thermomixer at 1200 rpm at 37°C for 20 min.
15. Wash beads
    a. Apply all IP samples to magnet, removing supernatant when clear.
    b. Wash bead pellets with 500 μL cold Wash Buffer. Move tubes back and forth on magnet, allow supernatant to clear and remove. Repeat wash, moving tubes back and forth on magnet, followed by removal of clear supernatant.
    c. Wash beads 2 times as above with 500 μL cold High Salt Wash Buffer.
    d. Wash beads 2 times as above with 500 μL cold Wash Buffer.
    e. Wash beads with 300 μL 1X Ligase Buffer (no DTT), moving back and forth on magnet and removing cleared supernatant.
    f. Wash beads 2 times with 300 μL 1X Ligase Buffer (no DTT). Keep beads in buffer on ice until next step.
16. Ligate 3′ RNA adapter
    a. Prepare "Early" 3′ RNA Ligation master mix. Mix well by pipetting or flicking.
    b. Remove supernatant from IP bead samples. Briefly spin samples in minifuge, apply to magnet, and use a small pipette tip to remove all residual liquid in tubes.
    c. Add 25 μL of master mix to each IP bead sample.
    d. Add 2.5 μL 40 μM InvRiL19 to each sample.
    e. Incubate at 15°C–25°C for 75 min, flicking tubes every 10 min to mix.
    f. Wash beads 2 times with 500 μL cold Wash Buffer.

Isolate RNA

⏱ Timing: 4–24 h

Complete this section for all IP and Input samples and for all protocol types: fSHAPE-eCLIP, SHAPE-eCLIP, or "in vitro" SHAPE-eCLIP. Verify efficacy of the antibody with a Western blot and isolate IP and Input RNA with a gel transfer.

17. Setup
    a. Pre-heat thermomixer to 70°C.
    b. Retrieve gel tank(s) and power supply.
    c. Dilute 20X MOPS Running Buffer to 2 L of 1X buffer.
    d. Prepare 2 L of **cold** 1X NuPAGE Transfer Buffer with 10% methanol. Keep at 4°C.
18. For only fSHAPE-eCLIP or "in vitro" SHAPE-eCLIP samples from **9e**: set aside non-immunoprecipitated samples as controls, referred to as "Input" samples
    a. Of the sample conditions ("A","B", or "C"; see 1b for naming convention) from step 9e, pick one condition to take sample Input controls from. Typically "B" or "C" conditions. One can carry forward Input controls for all conditions if one chooses, but this creates more controls than is necessary.
    b. Gently mix samples from 9e well. Take **two** 20 μL aliquots from each replicate of the chosen condition, **four** aliquots total, adding each to separate tubes labeled "Transfer-Input" or "Western-Input" and their replicate number. **Set Inputs aside on ice until step 20.**
    c. The remainder of the samples is referred to as immunoprecipitated (IP) samples.
19. Wash IP bead samples, from step 16f or 18c
    a. Apply IP samples to magnet, remove supernatant, and wash two times with 500 μL High Salt Wash Buffer. Remove cleared supernatant.
    b. Resuspend beads in 500 μL cold High Salt Wash Buffer. Add 500 μL cold Wash Buffer. Allow bead pellets to reform on magnet and remove cleared supernatant.
    c. Wash three times with cold Wash Buffer.
20. Prepare IP and Input samples for gel separation
    a. Prepare two 4%–12% Bis-Tris SDS-PAGE gels, 12-well, 1.5 mm. One will be used for a Western Blot ("Western") and the other for a gel transfer of RNA ("Transfer").
    b. Remove all supernatant from IP bead samples. Take beads off magnet and resuspend beads well in 100 μL Wash Buffer.
    c. For each resuspended IP bead sample, move 20 μL to a new tube. Label these aliquots with their treatment type, replicate number, and "Western." These samples will be used in the Western Blot.
    d. For the remaining 80 μL of IP bead samples, apply to magnet, remove all supernatant, and resuspend beads in 20 μL Wash Buffer. These samples will be used for the RNA transfer.
    e. Prepare a loading buffer master mix of 7.5 μL 4X NuPAGE LDS Buffer with 3 μL 1 M DTT for every sample (IPs and Inputs for the Western Blot and RNA Transfer; typically 12 samples total).
    f. Add 10 μL DTT-Buffer mix to each sample and denature all samples in Thermomixer for 10 min at 70°C, 1200 rpm. Samples are now decoupled from beads.
    g. Cool samples on ice for 1 min followed by a brief spin in a minifuge.
    h. Place all tubes (including Inputs) on magnet on ice. Desired material is contained in the supernatant.
21. Gel separate IP and Input samples
    a. Set up the two Bis-Tris gels. Label wells according to Tables 1 and 2. Fill tanks with 1X MOPS Running Buffer.
    b. Load IP and Input RNA Transfer samples into gel according to map in Table 1. Run at 150 V for 75 min, or until dye front is at bottom.
    c. Load half of IP and Input Western Blot samples (15 μL) into gel according to map in Table 2. Save the remainder of Western Blot samples at −20°C for troubleshooting. Run at 150 V for 75 min, or until dye front is at bottom.

Table 1. Gel set up for Transfer samples

Column	1	2	3	4	5	6	7	8	9	10	11	12
Sample	M	Input rep1	M	IP, treatment A/B rep1	M	IP, treatment B/C rep1	M	Input rep2	M	IP, treatment A/B rep2	M	IP, treatment B/C rep2
Volume to Load	5 μL	30 μL	5 μL	30 μL	5 μL	30 μL	5 μL	30 μL	5 μL	30 μL	5 μL	30 μL
% of Sample Represented		2%		80%		80%		2%		80%		80%

M = protein marker (3 uL) in loading buffer (2 uL).

22. Set up membranes while gels are running
    a. For the Transfer gel: cut out four Whatman papers, one nitrocellulose membrane, and two sponges, all sized to fit the gel.
    b. For the Western gel: cut out four Whatman papers, one PVDF membrane, and two sponges, all sized to fit the gel.
    c. Fill two separate shallow trays half-way with cold 1X Transfer buffer with 10% methanol. Place cassettes for Western Blot/RNA Transfer open in trays.
    d. Place sponges, Whatman papers, and nitrocellulose/PVDF membranes in trays to soak. Use rollers to roll bubbles out of sponges if necessary.
23. Western Blot and RNA Transfer
    a. Once gels are finished running, carefully remove one side of their casing. Using a straight edge, remove extraneous portions of the gels: the bottom below the dye front, the wells. Make sure to note the samples' orientation in the gels.
    b. Set up the RNA Transfer inside open cassette in the shallow tray. Carefully stack layers as follows, using a small roller carefully roll out bubbles between additions, such that each layer is flush to the one below.
        i. Sponge
        ii. 2 Whatman papers
        iii. Gel from 21b
        iv. Nitrocellulose membrane
        v. 2 Whatman papers
        vi. Sponge
    c. Set up the Western Blot inside open cassette in the shallow tray. Carefully stack layers as follows, using a small roller carefully roll out bubbles between additions, such that each layer is flush to the one below.
        i. Sponge
        ii. 2 Whatman papers
        iii. Gel from 21c
        iv. PVDF membrane
        v. 2 Whatman papers
        vi. Sponge
    d. Close each cassette. Cassettes should snugly fit the RNA Transfer/Western Blot, holding everything in place. If loose, add another sponge on top. **Note the side of the cassette that the nitrocellulose/PVDF membrane is closest to.**

Table 2. Gel set up for Western Blot

Column	1	2	3	4	5	6	7	9	10
Sample	M	Input rep1	IP, treatment A/B rep1	IP, treatment B/C rep1	M	Input rep2	IP, treatment A/B rep2	IP, treatment B/C rep2	M
Volume to Load	5 μL	15 μL	15 μL	15 μL	5 μL	15 μL	15 μL	15 μL	5 μL
% of Sample Represented		1%	10%	10%		1%	10%	10%	

M = protein marker in loading buffer.

e. Place cassettes in a wet transfer tank. **The side of the cassette that is closest to the nitrocellulose/PVDF membrane should face the positive electrode in the tank**, such that material flows from the gel into the membrane.

f. Fill transfer tank with cold 1X Transfer Buffer with 10% methanol.

g. Run at 30V 10–20 h in cold room (preferred for Transfer), or at 200 mA for 2 h.

Develop the western blot

⏱ **Timing: 3 h**

Develop the Western Blot to confirm that immunoprecipitation was successful.

24. Setup
    a. Preheat Thermomixer to 37°C.
    b. In the meantime, continue to run the RNA Transfer (recommended). Or, remove the RNA Transfer nitrocellulose membrane, rinse briefly with sterile 1X PBS, wrap in plastic wrap, and store at −20°C.
25. Develop western blot
    a. Stop Western Blot transfer and remove PVDF membrane from stack.
    b. Place a small cut on the membrane for noting its orientation.
    c. Incubate membrane in Blocking Buffer (5% milk in 1X TBST Buffer) at 15°C–25°C for 30 min with gentle mixing.
    d. Dilute primary antibody in Blocking Buffer to a final antibody concentration of 0.2–0.5 ug/mL. Incubate membrane with antibody at 15°C–25°C for 1 h with gentle mixing.
    e. Move membrane to fresh container in 1X TBST Buffer. Let incubate for 5 min with gentle mixing. Repeat 2 times.
    f. Dilute secondary antibody, rabbit or mouse TruBlot HRP (depending on whether primary antibody is derived from rabbit or mouse etc), to a final dilution of 1:2000 in Blocking Buffer. Incubate at 15°C–25°C for 1–3 h with gentle mixing.
    g. Move membrane to fresh container in 1X TBST Buffer. Let incubate for 5 min with gentle mixing. Repeat 2 times.
    h. Mix equal parts ECL Reagent 1 and 2, 0.5 mL each. Apply to membrane and incubate in place for 1 min.
    i. Remove membrane from solution, touching edges of membrane to paper towels to absorb excess solution. Encase membrane in protective clear plastic.
    j. Expose membrane to film in dark room for 30 s and 5 min time spans. Adjust exposure times as necessary to render a clear visualization of bands.
    k. Verify that protein band appears for IP samples and that it indicates the expected size for the given protein being immunoprecipitated. See Figure 1A and troubleshooting 1

Note: Bands in the Western Blot may be much darker for the IP samples that were not treated with NAI *in vivo*, indicating greater yields for these samples. This is normal and likely due to *in vivo* NAI treatment unavoidably resulting in some cellular stress/death.

⧗ **Pause point:** RNA transfer may continue to run at 4°C for 10–20 hours total. Alternatively, the RNA Transfer nitrocellulose membrane may be rinsed briefly with sterile 1X PBS, wrapped in plastic wrap, and stored at −20°C for ~1 day.

Isolate RNA

⏱ **Timing: 4 h**

Extract RNA from nitrocellulose membrane for all IP and Input samples.
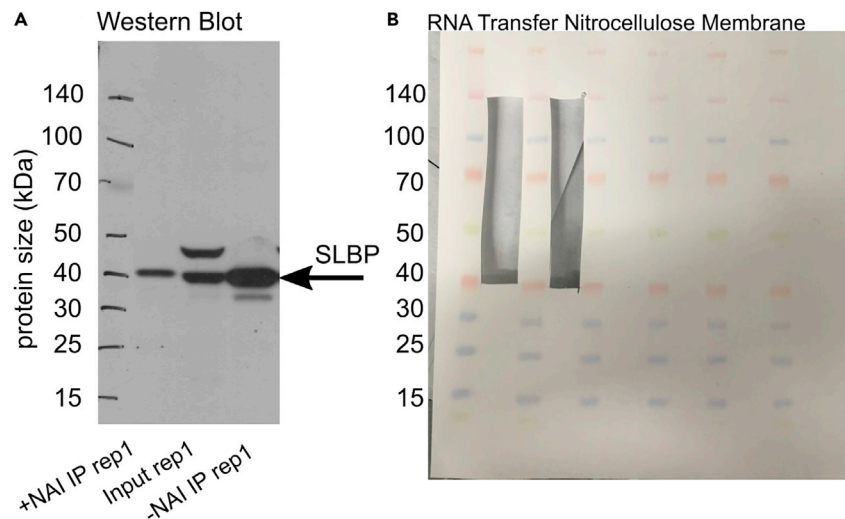
Figure 1. RNA isolation steps
(A) Western blot of immunoprecipitated (IP) and Input control (f)SHAPE-eCLIP samples for stem loop binding protein (SLBP).
(B) (f)SHAPE-eCLIP RNA samples transferred to nitrocellulose membrane, with excised samples.

26. Setup
    a. Preheat Thermomixer to 37°C.
    b. Fetch sterile razor blades, one per sample (six blades).
    c. Clean a glass plate (or other hard surface to cut against) with ethanol and/or RNase-removing solution.
    d. Fetch 1.5 mL tubes, one per sample.
27. Cut RNA Transfer nitrocellulose membrane
    a. Fetch nitrocellulose membrane from RNA transfer. Note expected positioning of protein band within each lane based on the western blot. See Figure 1A.
    b. Cut the IP and Input sample lanes using a fresh razor blade for each lane.
        i. Cuts should start at the expected size of the immunoprecipitated protein (40 kDa for SLBP) and extend to 75 kDa above (Figure 1B). (75 kDa correspond to the mass of an RNA molecule that is ~230 bases long).
        ii. Further cut each excised lane into 1–2 mm pieces. Place pieces into 1.5 mL tubes and label with sample name (Methods Video S2).
28. Extract RNA from membrane
    a. Prepare 150 µL Proteinase K SDS mix per sample, consisting of 130 µL PKS Buffer + 20 µL Proteinase K.
    b. Add Proteinase K mixture to membrane pieces, ensuring all pieces are submerged.
    c. Incubate in Thermomixer 37°C for 20 min, 1200 rpm.
    d. Further incubate in Thermomixer at 50°C for 20 min, 1200 rpm.
    e. Transfer solution in each tube to new 1.5 mL tube.
    f. Add 55 µL water to leftover membrane pieces. Transfer water to supernatant above.
29. Purify RNA with RNA Clean and Concentrator Kit column cleanup
    a. To each sample add 400 µL (2x starting volume) RNA binding buffer to supernatant from membrane (step 28 f), pipetting to mix.
    b. Add 700 µL (3.5x starting volume) of 100% ethanol and mix well.
    c. Transfer 650 µL of each sample to a spin column (the max volume column can accommodate at once).
    d. Centrifuge columns for 30 s at 5000 g at 15°C–25°C. Add flow-through back into columns and centrifuge again. Discard flow-through.

e. Load remaining sample to the same column and centrifuge, re-load, centrifuge as before. Discard flow-through.

f. Wash columns with the following buffers, centrifuging for 30 s at 5000 g, discarding flow-through each time:
   i. 400 µL RNA Prep Buffer
   ii. 500 µL RNA Wash Buffer
   iii. 500 µL RNA Wash Buffer
   iv. 200 µL RNA Wash Buffer. Centrifuge for 1 min at 9000 g.

g. Centrifuge for 2 min at max speed to dry columns.

h. Transfer columns to new 1.5 mL tubes.

i. Add 10 µL water to columns, sit for 1 min, and centrifuge for 30 s at 9000 g.

j. Add flow-through back onto columns and let sit and centrifuge at 9000 g as before. Flow-through contains purified RNA.

k. Samples may be stored at −80°C.

⏸ **Pause point:** Samples may be kept at −80°C for a few days.

Note: If following SHAPE-eCLIP protocol skip to "Ligate RNA (late)". fSHAPE-eCLIP and the "in vitro" variant of SHAPE-eCLIP proceed to "In vitro RNA treament."

In vitro RNA treatment with NAI

⏱ **Timing: 1 h**

Complete this section only for treatment condition "B" IP samples, either for fSHAPE-eCLIP or the "in vitro" variant of SHAPE-eCLIP (see step 1). Refold IP RNA samples and treat with structure probing reagent NAI.

30. Setup
    a. Samples not undergoing this step should be set aside on ice.
    b. Preheat thermocycler or heat block to 95°C.

31. Probe RNA with NAI
    a. Heat ∼10 µL "B" IP samples to 95°C for 2 min, then place on ice for 1 min.
    b. Prepare "in vitro structure probing" reaction with heated samples, without NAI (see materials and equipment).
    c. Heat to 37°C for 5 min to re-fold RNA.
    d. Add 1 µL 2 M NAI to each sample and continue heating at 37°C for 10 min.

32. Purify RNA with RNA Clean and Concentrator Kit column cleanup
    a. To each sample add 42 µL (2x starting volume) RNA binding buffer, pipetting to mix.
    b. Add 73.5 µL (3.5x starting volume) of 100% ethanol and mix well.
    c. Transfer the entirety of each sample to a spin column.
    d. Centrifuge columns for 30 s at 5000 g at 15°C–25°C . Add flow-through back into columns and centrifuge again. Discard flow-through.
    e. Load remaining sample to column and centrifuge, re-load, centrifuge as before. Discard flow-through.
    f. Wash columns with the following buffers, centrifuging for 30 s at 5000 g, discarding flow-through each time:
       i. 400 µL RNA Prep Buffer
       ii. 500 µL RNA Wash Buffer
       iii. 500 µL RNA Wash Buffer
       iv. 200 µL RNA Wash Buffer. Centrifuge for 1 min at 9000 g.
    g. Centrifuge for 2 min at max speed to dry columns.
    h. Transfer columns to new 1.5 mL tubes.

   i. Add 10 μL water to columns, sit for 1 min, and centrifuge for 30 s at 9000 *g*.

   j. Add flow-through back onto columns and let sit and centrifuge as before. Flow-through contains purified RNA.

   k. Samples may be stored at −80°C.

▱▱ **Pause point:** Samples may be kept at −80°C for a few days.

Ligate 3′ RNA adapter (late)

⊙ Timing: 3 h

For the SHAPE-eCLIP protocol complete this section for Input samples only. For fSHAPE-eCLIP or "in vitro" SHAPE-eCLIP protocols complete this section for all IP and Input samples. Ligate an adapter, later used for reverse transcription.

33. Setup
    a. Preheat Thermomixer to 37°C.
    b. If following SHAPE-eCLIP protocol, set IP samples aside on ice until "Reverse Transcribe RNA" section. These sample will have already undergone RNA adapter ligation.
34. FastAP treat RNA
    a. Prepare "Late" FastAP master mix on ice (see materials and equipment). Mix well.
    b. Add 11 μL FastAP master mix to each sample.
    c. Incubate samples in Thermomixer at 1200 rpm at 37°C for 10 min.
35. PNK treat RNA
    a. While FastAP reaction is incubating, prepare "Late" PNK master mix.
    b. Mix well and add 75 μL master mix to each sample.
    c. Incubate samples in Thermomixer at 1200 rpm at 37°C for 20 min.
36. Purify RNA with RNA Clean and Concentrator Kit column cleanup
    a. To each sample add 200 μL (2x starting volume) RNA binding buffer, pipetting to mix.
    b. Add 300 μL (3x starting volume) of 100% ethanol and mix well.
    c. Transfer the entirety of each sample to a spin column.
    d. Centrifuge columns for 30 s at 5000 *g* at 15°C–25°C. Add flow-through back into columns and centrifuge again. Discard flow-through.
    e. Load remaining sample to column and centrifuge, re-load, centrifuge as before. Discard flow-through.
    f. Wash columns with the following buffers, centrifuging for 30 s at 5000 *g*, discarding flow-through each time:
       i. 400 μL RNA Prep Buffer
       ii. 500 μL RNA Wash Buffer
       iii. 500 μL RNA Wash Buffer
       iv. 200 μL RNA Wash Buffer. Centrifuge for 1 min at 9000 *g*.
    g. Centrifuge for 2 min at max speed to dry columns.
    h. Transfer columns to new 1.5 mL tubes.
    i. Add 10 μL water to columns, sit for 1 min, and centrifuge for 30 s at 9000 *g*.
    j. Add flow-through back onto columns and let sit and centrifuge as before. Flow-through contains purified RNA.

▱▱ **Pause point:** Treated RNA may be stored at −80°C.

37. Ligate 3′ RNA adapter
    a. Preheat Thermomixer or heat block to 65°C.
    b. Prepare "Late" 3′ RNA Ligation master mix for Inputs and "Late" 3′ RNA Ligation mix for IP samples if treating IP samples at this stage (i.e., for fSHAPE-eCLIP or "in vitro" SHAPE-eCLIP).

   c. For Input samples: remove half of samples (5 μL) and store at −80°C as backups.

   d. For IP samples: use entire sample.

   e. Incubate samples with InvRiL19 adapter:

     i. Add 2 μL of 40 μM InvRiL19 adapter and 1.5 μL DMSO to Input samples.

     ii. Add 4 μL of 40 μM InvRiL19 adapter and 2.4 μL DMSO to IP samples (for fSHAPE-eCLIP or "in vitro" SHAPE-eCLIP).

     iii. Incubate at 65°C for 2 min.

     iv. Snap cool on ice for 1 min.

   f. Add 13.5 μL of master mix to Input samples and 27 μL to IP samples. Mix well.

   g. Incubate at 15°C–25°C for 60 min, flicking tubes every 15 min to mix.

38. Cleanup samples with Silane magnetic beads

   a. Prepare beads:

     i. Take 10 μL MyONE Silane Beads per Input sample and 20 μL per sample.

     ii. Add 5x volumes RLT Buffer to beads. (500 μL RLT Buffer to 100 μL of beads). Mix well.

     iii. Apply to magnet and remove cleared supernatant.

     iv. Add 6.3x volumes RLTW Buffer to beads. (630 μL RTLW Buffer to originally 100 μL of beads). Mix well.

   b. Bind RNA samples to beads:

     i. Add 61 μL beads in RLTW Buffer to each Input sample, 122 μL to each IP sample. Mix well.

     ii. Add 65 μL 100% ethanol to each Input sample, 130 μL to each IP sample.

     iii. Pipette mix each sample 10 times and leave pipette tip in tube.

     iv. Incubate with beads for 10 min, pipetting to mix every 3–5 min.

   c. Wash beads:

     i. Prepare 20 mL 80% ethanol.

     ii. Apply samples to magnet and discard cleared supernatant.

     iii. Add 300 μL 80% ethanol to samples if in 0.2 mL tubes, or 1 mL if in 1.5 mL tubes. Mix well.

     iv. Wait 30 s, then apply samples to magnet and discard cleared supernatant.

     v. Repeat wash with 80% ethanol. Discard supernatant.

     vi. Repeat wash with 100 μL or 750 μL 80% ethanol. Discard supernatant.

     vii. Briefly spin tubes, apply to magnet, and remove residual supernatant with a fine tip.

     viii. Leave tubes open to let beads dry, ~5 min. Beads are dry when they no longer shine or look wet. (Beads are over-dried if they change color from brown to burnt orange.)

   d. Elute RNA from beads:

     i. Resuspend beads in 9.5 μL TT Elution Buffer. Sit for 5 min.

     ii. Apply samples to magnet and capture supernatant, which contains RNA samples. Move supernatant to new 0.2 mL tubes.

⏸️ **Pause point:** Samples may be stored at −80°C for up to two weeks.

Reverse transcribe RNA and ligate 5′ cDNA adapter

⏱ **Timing: 4 h + 10–20 h incubation**

Complete this section for all samples and all protocol types. Reverse transcribe RNA with manganese to perform "mutational profiling," followed by 5′ adapter ligation to the cDNA.

39. Setup

   a. Preheat thermocycler or heat block to 65°C.

40. Anneal primer

   a. To each ~9 μL sample add 1 μL 5 μM InvAR17 primer and 1 μL 10 mM dNTPs.

   b. Heat to 65°C for 2 min. Snap cool on ice for one minute. (Do not let cool down on heat block).

41. Reverse transcribe

   a. Set thermocycler or heat block to 45°C.

b. Prepare SHAPE reverse transcription master mix (see materials and equipment).

c. Add 10 μL master mix to each sample. Mix well.

d. Incubate at 45°C for 3 h.

⚠ CRITICAL: The reverse transcription here requires all 3 hours, since manganese-induced errors slow down the RT.

Note: Errors are induced in the cDNA by manganese when the RT encounters NAI adducts in the RNA. This "mutational profiling" records the location of NAI reactions with RNA bases and is a crucial step of RNA secondary structure probing.

42. Isolate cDNA

    a. Treat with ExoSAP-IT

      i. Set heat block to 37°C.

      ii. Add 3.5 μL ExoSAP-IT to each sample. Vortex and briefly spin down.

      iii. Incubate at 37°C for 15 min.

      iv. Add 1 μL 0.5 M EDTA to each sample and mix.

    b. Remove RNA

      i. Set heat block to 70°C.

      ii. Add 3 μL 1 M NaOH to each sample and mix.

      iii. Incubate at 70°C for 10 min.

      iv. Add 3 μL 1 M HCl and mix, to neutralize NaOH.

43. Cleanup cDNA with Silane magnetic beads

    a. Prepare beads:

      i. Take 5 μL MyONE Silane Beads per sample.

      ii. Add 5x volumes RLT Buffer to beads. (150 μL RLT Buffer to 30 μL of beads). Mix well.

      iii. Apply to magnet and remove cleared supernatant.

      iv. Add 18.6x volumes RLT**W** Buffer to beads. (558 μL RTLW Buffer to originally 30 μL of beads). Mix well.

    b. Bind cDNA samples to beads:

      i. Add 90 μL beads in RLTW Buffer to each sample. Mix well.

      ii. Add 108 μL 100% ethanol to each sample.

      iii. Pipette mix each sample 10 times and leave pipette tip in tube.

      iv. Incubate with beads for 10 min, pipetting to mix every 3–5 min.

    c. Wash beads:

      i. Prepare 20 mL 80% ethanol.

      ii. Apply samples to magnet and discard cleared supernatant.

      iii. Add 300 μL 80% ethanol to samples if in 0.2 mL tubes, or 1 mL if in 1.5 mL tubes. Mix well.

      iv. Wait 30 s, then apply samples to magnet and discard cleared supernatant.

      v. Repeat wash with 80% ethanol. Discard supernatant.

      vi. Repeat wash with 100 μL or 750 μL 80% ethanol. Discard supernatant.

      vii. Briefly spin tubes, apply to magnet, and remove residual supernatant with a fine tip.

      viii. Leave tubes open to let beads dry, ∼5 min. Beads are dry when they no longer shine or look wet. (Beads are over-dried if they change color from brown to burnt orange.)

44. Ligate 5′ cDNA linker on beads

    a. Set heat block to 70°C.

    b. Make cDNA adapter mix, per sample:

      i. 1.1 μL TT Elution Buffer

      ii. 0.6 μL InvRand3Tr3 adapter

      iii. 0.8 μL 100% DMSO

      iv. Heat at 70°C for 2 min.

      v. Snap cool on ice for 1 min.

    c. Add 3 μL cDNA adapter mix to each dried Silane bead sample.

d. Prepare 5′ cDNA linker ligation master mix (See materials and equipment).

e. Add 7.8 master mix to each bead sample by first stirring beads with pipette tip, then slowly pipetting the viscous master mix while stirring until homogenous.

f. Incubate at 15°C–25°C 10–20 h on a rotator.

45. Cleanup ligated cDNA with Silane magnetic beads

    a. Add 5 μL TT Elution Buffer to each sample to total to 15 μL.

    b. Prepare beads:

       i. Prepare 2.5 μL MyONE Silane Beads per sample.

      ii. Add 5x volumes RLT Buffer to beads. (75 μL RLT Buffer to 15 μL of beads). Mix well.

     iii. Apply to magnet and remove cleared supernatant.

     iv. Add 18.8x volumes RLT**W** Buffer to beads. (282 μL RTLW Buffer to originally 15 μL of beads). Mix well.

    c. Bind cDNA samples to beads:

       i. Add 45 μL beads in RLTW Buffer to each sample. Mix well.

      ii. Add 45 μL 100% ethanol to each sample.

     iii. Pipette mix each sample 10 times and leave pipette tip in tube.

     iv. Incubate with beads for 10 min, pipetting to mix every 3–5 min.

    d. Wash beads:

       i. Prepare 20 mL 80% ethanol.

      ii. Apply samples to magnet and discard cleared supernatant.

     iii. Add 300 μL 80% ethanol to samples if in 0.2 mL tubes, or 1 mL if in 1.5 mL tubes. Mix well.

     iv. Wait 30 s, then apply samples to magnet and discard cleared supernatant.

      v. Repeat wash with 80% ethanol. Discard supernatant.

     vi. Repeat wash with 100 μL or 750 μL 80% ethanol. Discard supernatant.

     vii. Briefly spin tubes, apply to magnet, and remove residual supernatant with a fine tip.

    viii. Leave tubes open to let beads dry, ∼5 min. Beads are dry when they no longer shine or look wet. (Beads are over-dried if they change color from brown to burnt orange.)

    e. Elute cDNA from beads:

       i. Resuspend beads in 25 μL TT Elution Buffer. Sit for 5 min.

      ii. Apply samples to magnet and capture supernatant, which contains cDNA samples. Move supernatant to new 1.5 mL tubes.

⧗⧗ **Pause point:** Store samples at −80°C.

Quantify and amplify cDNA library

⊙ **Timing: 8 h**

Complete this section for all samples and all protocol types. Quantify cDNA with quantitative PCR (qPCR), also known as "real time PCR", followed by amplification and gel purification of libraries.

46. qPCR

    a. Prepare cDNA qPCR master mix (see materials and equipment).

    b. For each sample, add 9 μL master mix to a well in a 384 well plate, or other plate compatible with qPCR instrument.

    c. Add 1 μL of a 1:10 dilution of each cDNA sample to the wells in the 384 well plate.

    d. Seal the plate with clear film.

    e. Execute qPCR according to the preset protocol in Table 3.

    f. Record Ct calculated for each sample. Typical Input sample Ct: 12. IP sample Ct: 19. See expected outcomes.

47. PCR amplify cDNA library

    a. Prepare cDNA PCR master mix, selecting a unique combination of forward and reverse (D50x/D70x) primers for each sample.

Table 3. Quantify libraries with qPCR

qPCR protocol

Step	Temperature	Time	Cycles
Initial Denaturation	95 C	10 min	1
Denaturation	95 C	15 s	30
Annealing & extension	60 C	1 min	
Take image	NA	NA	

    b. Add PCR master mix to 16 μL of each cDNA sample for 40 μL total.

    c. The PCR cycle number for each sample should be 3 cycles less than the qPCR-calculated Ct for its 1:10 dilution. Example:

      i. Typical Input: 12−3 = 9 total PCR cycles

      ii. Typical IP: 19−3 = 16 total PCR cycles

    d. The PCR consists of two stages, the first of which always consists of 6 cycles. The second stage should be repeated such that the total number of cycles attains the preferred total cycles calculated above. Example:

      i. Typical Input: 6 cycles + 3 cycles = 9 total PCR cycles

      ii. Typical IP: 6 cycles + 10 cycles = 16 total PCR cycles

    e. Amplify libraries according to PCR protocol in Table 4.

48. Cleanup library with AmpureXP beads

    a. Prepare 10 mL 80% ethanol.

    b. Leave beads out at 15°C–25°C 15 min before use.

    c. Vortex beads well.

    d. Add 72 μL bead suspension to each PCR reaction. Mix well.

    e. Incubate at 15°C–25°C for 10 min, mixing 2–3 times during this time.

    f. Apply samples to magnet and remove cleared supernatant.

    g. Wash beads on magnet 3 times with 100 μL 80 ethanol (if in 0.2 mL tubes) or 500 μL (if in 1.5 mL tubes).

    h. Air dry beads for 5 min on magnet. Avoid over-drying (pellet will appear cracked).

    i. Resuspend beads in 20 μL PCR Elution Buffer, incubating for 5 min at 15°C–25°C .

    j. Apply samples to magnet and move supernatant containing cleaned cDNA library to new tubes.

    ⏸⏸ Pause point: Store samples at −80°C.

49. Gel purify cDNA library

    a. Prepare 3% low melting temperature agarose gel in 1X TBE Buffer.

      i. Prepare 50 mL total for 7 × 10 cm tray, 80 mL for 15 × 7 cm tray, 120 mL for 15 × 10 cm tray

      ii. Heat TBE Buffer and gradually add agarose powder while stirring.

Table 4. Amplify libraries with PCR

PCR protocol

Step	Temperature	Time	Cycles
Initial Denaturation	98 C	30 s	1
Denaturation	98 C	15 s	6
Annealing	68 C	30 s	
Extension	72 C	40 s	
Denaturation	98 C	15 s	Ct# - 3 - 6
Annealing & extension	72 C	60 s	
Final Extension	72 C	60 s	1
Hold	4 C	Forever	1

     iii. Completely dissolve agarose by intermittently heating solution in microwave in 30 s intervals followed by stirring/swirling.

     iv. Add 1:10000 SYBR Safe gel stain. Mix well and pour agarose solution into tray.

b. Separate library samples on gel

     i. Add 6 μL 6X OrangeG Buffer to each cleaned sample from step 48j (∼24 μL total). Mix well.

     ii. Prepare 50 base-pair ladder in OrangeG Buffer, enough for two wells.

     iii. Move agarose gel into gel tank in 1X TBE Buffer. Load samples into wells with ladder on each end.

     iv. Run gel at ∼95 V for 50 min. More time will result in finer separation, but larger gel cuts. See example in expected outcomes.

c. Cut gel

     i. Visualizing with blue light, cut each lane from 175–350 base-pairs in size with a fresh razor blade for each sample. Place cut gel pieces into 15 mL conical tube.

Note: primer-dimers run at 142 base-pairs. Avoid cutting too close to this size.

     ii. Weigh the gel slice in each conical, taring with an empty 15 mL conical tube.

d. Extract libraries from gel with Qiagen MinElute gel extraction kit

     i. Add 6x μL of Buffer QG for every 100 mg of gel. Ex: 600 μL Buffer QG for 100 mg gel.

     ii. Incubate gel in buffer at 15°C–25°C . Can shake or swirl to help dissolve; no vortexing.

     iii. Add 1x μL of isopropanol and mix very well. Ex: 100 μL isopropanol for 100 mg original weight of gel slice.

     iv. Load a maximum of 750 μL dissolved gel solution onto column and spin in centrifuge at max speed for 1 min. Discard flow-through. Repeat until all solution has been spun through.

     v. For gel weights >400 mg, spin column with 500 μL Buffer QG after every 4 spins.

     vi. After all sample has been spun through column, wash column once with 500 μL Buffer QG at max speed for 1 min.

     vii. Add 750 μL Buffer PE to column and spin at max speed for 1 min. Discard flow-through and repeat wash, spinning for 2 min.

     viii. Move column to a new 1.5 mL tube. If column has any residual Buffer PE on the outside, repeat 2 min spin to dry completely.

     ix. With a fine pipette tip, move tip around the rim lining the filter inside the column taking up any residual liquid around rim. (No residual Buffer PE should be present in final libraries).

     x. Air dry columns for 2 min.

     xi. Add 12.5 μL Buffer EB to the center of column. Incubate for 2 min. Spin at max speed for 1 min. Add eluate back into column and repeat spin for improved yield.

e. Quantitate libraries

     i. Run a sample of each library on a D1000 DNA TapeStation, Bioanalyzer, or equivalent.

     ii. Quantify library peaks between ∼150–600 base-pairs. See example in expected outcomes and troubleshooting 2 and 3. Sufficient library concentration is >= 5 nM per library.

f. Sequence

     i. Pool libraries according to a final molarity specified by your sequencing provider. Each Input sample library may be pooled at half the concentration that IP sample libraries are pooled at. For example, if pooling samples to a final concentration of 10 nM, four IP libraries could each comprise 2 nM and two Input libraries could each comprise 1 nM.

     ii. Sequence single-end libraries aiming for a depth of at least 50 million reads per library.

g. Process sequencing reads with the (f)SHAPE-eCLIP pipeline. See quantification and statistical analysis.

EXPECTED OUTCOMES

Success of the (f)SHAPE-eCLIP libraries will be first apparent after qPCR. The library preparation has likely been successful if the Ct number of its 1:10 dilution is in the expected range (Table 5). Libraries are size-selected on an agarose gel (below).

Table 5. Expected PCR cycle numbers for (un)successful libraries

Sample library type	Good qPCR Ct	Bad qPCR Ct
Input	9-15	>15
IP	<21	>=21

Successful cDNA libraries after PCR amplification and bead purification are visualized on a gel as a smear in the 150–350 bp range (Figure 2). Dotted outlines in the figure above indicate the approximate region boundaries for excision of libraries (175–350 bp). Star indicates adapter dimer and cross indicates unincorporated PCR primers remaining after bead purification. With conservative gel excision, both of these should be separated from the libraries with high confidence.

Successful final libraries as visualized and quantified on a TapeStation instrument are expected to resemble Figure 3. The "region molarity" value is the concentration used to determine the pooling strategy for individual samples prior to sequencing and will ideally be >5 nM for any given library.
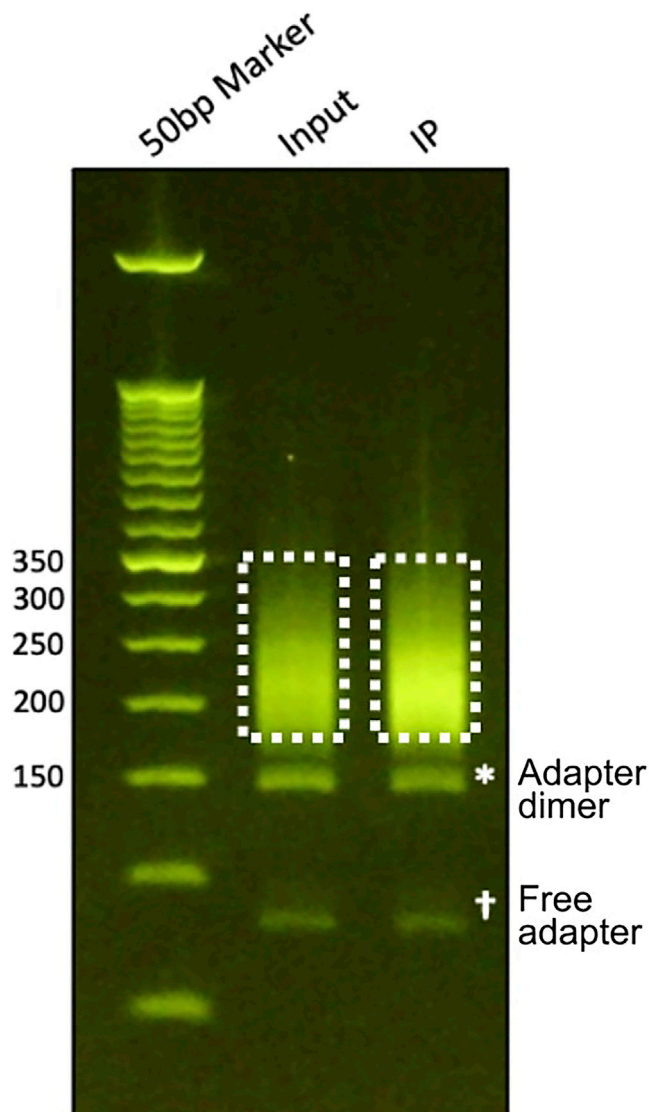
Figure 2. (f)SHAPE-eCLIP libraries purified on agarose gel
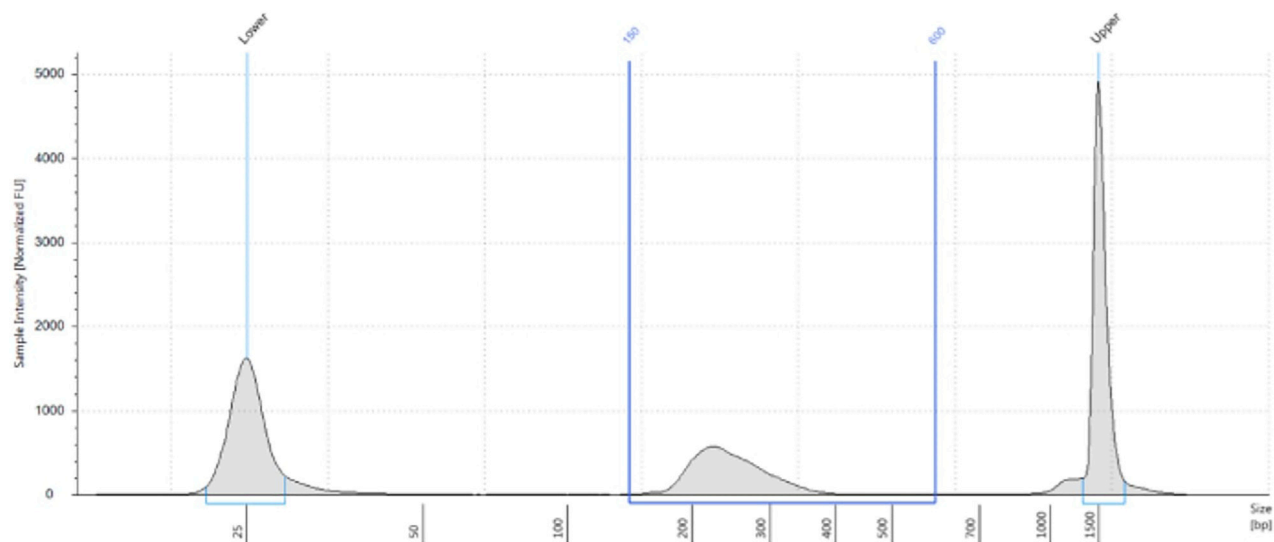
Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
150	600	251	4.15	26.4	75.06		■

Figure 3. Visualized final (f)SHAPE-eCLIP library

fSHAPE Expected Outcomes:

The (f)SHAPE-eCLIP wet lab protocol, not fSHAPE, is described in this manuscript. For an fSHAPE protocol, follow the published icSHAPE protocol (Flynn et al., 2016) for both "in vitro" and "in vivo" icSHAPE, ignoring steps for the DMSO-treated control. While most of the fSHAPE library preparation does not have visualization steps, after PCR and gel purification it will be clear if the protocol was successful. Successful fSHAPE libraries are indicated by ample dsDNA products >150 bp in size. See troubleshooting 4.

QUANTIFICATION AND STATISTICAL ANALYSIS

fSHAPE-eCLIP and SHAPE-eCLIP analysis:

The pipeline for (f)SHAPE-eCLIP is written in common workflow language (CWL) and run with cwltool and pre-built Docker containers holding the software necessary for each step of the pipeline. The pipeline requires at least 8 cores and 60 Gb of memory. If opting for an AWS EC2 instance we recommend a Centos7 AMI (ami-02676464a065c9c05, located on us-west-1) with either c3.8xlarge or r4.2xlarge machines and 500 Gb of attached storage.

Dependencies:

Docker

cwltool==3.0.20200706173533

pip==20.1.1

Logging into an AWS EC2 instance will likely require a pem key. SSH instructions should be given on the last page of setup for the instance. Log into your instance with the ip address given when you set up the instance:

$ ssh -i "my.pem" centos@<ip address>

Setup up your remote compute instance as follows:

$ sudo yum install -y install screen wget; sudo yum -y group install "Development Tools"

$ wget https://repo.anaconda.com/miniconda/Miniconda3-latest-Linux-x86_64.sh ; sh Miniconda3-latest-Linux-x86_64.sh

$ source ~/.bashrc

$ pip install pip==20.1.1

$ sudo yum install -y yum-utils ; sudo yum-config-manager --add-repo https://download.docker.com/linux/centos/docker-ce.repo ; sudo yum install -y docker-ce docker-ce-cli containerd.io

$ sudo systemctl start docker

$ sudo groupadd docker ; sudo usermod -aG docker ${USER}

Log out and then back into the machine. Use command "groups" and verify that "docker" is one of the groups. Test that Docker is working properly with:

$ docker run hello-world

$ pip install schema-salad==7.0.20210124093443

$ pip install cwltool==3.0.20200706173533

$ git clone https://github.com/YeoLab/eclip

$ git clone https://github.com/meracorley/f-SHAPE-eCLIP

$ cd f-SHAPE-eCLIP

$ git checkout pipeline

Softlink the CWL definition files from eclip into f-SHAPE-eCLIP:

$ cd /home/centos/f-SHAPE-eCLIP/cwl

$ ln -s /home/centos/eclip/cwl/*.cwl ./

Add the paths to the f-shape-eclip executables to PATH variable. The following is the command if https://github.com/yeolab/f-shape-eclip is cloned to the home directory of a centos machine. If on your system they were instead cloned to /home/user/mypath/, for example, then /home/centos/eclip/bin would become /home/user/mypath/eclip/bin, et cetera.

$ echo export PATH=/home/centos/eclip/bin:/home/centos/f-SHAPE-eCLIP:/home/centos/f-SHAPE-eCLIP/cwl:/home/centos/f-SHAPE-eCLIP/wf:＼$PATH >> ~/.bashrc

$ source ~/.bashrc

Build reference genome and repBase reference. For human (GRCh38/Hg38) as follows:

$ mkdir refs; cd refs

$ wget https://www.encodeproject.org/files/GRCh38_no_alt_analysis_set_GCA_000001405.15/@@download/GRCh38_no_alt_analysis_set_GCA_000001405.15.fasta.gz

$ wget https://external-collaborator-data.s3-us-west-1.amazonaws.com/reference-data/gencode.v29e.annotation.gtf.db

$ gunzip GRCh38_no_alt_analysis_set_GCA_000001405.15.fasta.gz

$ mkdir STAR

$ docker pull brianyee/star:2.7.6a

docker run -v /home/centos/refs:/refs brianyee/star:2.7.6a STAR \

--runThreadN 8 \

--runMode genomeGenerate \

--genomeDir refs/STAR \

--genomeFastaFiles refs/GRCh38_no_alt_analysis_set_GCA_000001405.15.fasta

$ wget https://external-collaborator-data.s3-us-west-1.amazonaws.com/reference-data/homo_sapiens_repbase_fixed_v2.fasta

$ mkdir STAR_repeat

$ docker run -v /home/centos/refs:/refs brianyee/star:2.7.6a STAR \

--genomeSAindexNbases 9 \

--runThreadN 8 \

--runMode genomeGenerate \

--genomeDir refs/STAR_repeat \

--genomeFastaFiles refs/homo_sapiens_repbase_fixed_v2.fasta

Download blacklisted regions for your species, GRCh38 (Hg38) as follows:

$ wget https://www.encodeproject.org/files/ENCFF269URO/@@download/ENCFF269URO.bed.gz

$ gunzip ENCFF269URO.bed.gz

$ cd ../

Fetch your data. Use scp or AWS CLI to copy your fastq files onto your machine.

Or use the fSHAPE-eCLIP example dataset stored in GEO: GSE149767, samples SRR11682378, SRR11682379, SRR11682382, SRR11682383, SRR11682384, SRR11682385:

$ cd ~/

$ wget http://ftp-trace.ncbi.nlm.nih.gov/sra/sdk/current/sratoolkit.current-centos_linux64.tar.gz ; tar xzf sratoolkit.current-centos_linux64.tar.gz

$ echo export PATH=/home/centos/sratoolkit.2.11.0-centos_linux64/bin:\$PATH >> ~/.bashrc

$ source ~/.bashrc

$ mkdir data; cd data/

$ for f in {SRR11682378,SRR11682379,SRR11682382,SRR11682383,SRR11682384,SRR11682385}; do fastq-dump $f ; gzip $f; done

Format a yaml metadata file to tell the pipeline which files represent which sample types. An example yaml file can be found in https://github.com/meracorley/f-SHAPE-eCLIP/tree/pipeline/examples/wf_fSHAPE.yaml

The yaml file requires two "treated" fastq files, two "untreated" fastq files, and two Input fastq files. For SHAPE-eCLIP, the "treated" samples will be treatment conditions "A" or "B" and "untreated" samples will be condition "C." See step 1b of this protocol for letter names behind treatment conditions. For fSHAPE-eCLIP, "treated" samples will be condition "B" and "untreated" will be condition "A" (Table 6). For the example files from GEO above, SRR11682382.fastq and SRR11682383.fastq are the "treated" files, SRR11682378.fastq and SRR11682379.fastq are the "untreated", and SRR11682384.fastq and SRR11682385.fastq are the Input sample files.

Order very much matters in the yaml file. The yaml file also specifies paths to the STAR index, chromosome sizes of the STAR index, blacklisted regions, the STAR_repeat index, and a GTF file for the given reference genome. Note that "GRCh38_v29e" is a Clipper parameter and is the recognized species name for hg38 in the "species" field of the yaml file ("_v29e" refers to "Gencode V29 Encode" as we recommend matching Clipper references to the assembly.

Make executable:

$ chmod +x wf_fSHAPE.yaml

The pipeline is ready to be run. Start a Screen session to run the pipeline:

$ screen -S pipeline

Execute the pipeline:

Table 6. How to label "treated" and "untreated" samples in yaml file

Protocol	Conditions (see step 1b)	"Treated" condition in yaml	"Untreated" condition in yaml	Example "treated" files	Example "untreated" files
fSHAPE-eCLIP	A,B	B	A	SRR11682382.fastq, SRR11682383.fastq	SRR11682378.fastq, SRR11682379.fastq
SHAPE-eCLIP	A,C	A	C	NA	NA
"in vitro" SHAPE-eCLIP	B,C	B	C	NA	NA

$ ./wf_fSHAPE.yaml

The pipeline puts output files within the ./wf_fSHAPE/results/ directory. Binding sites for the protein that was immunoprecipitated are output to BED files, one for each replicate. fSHAPE/SHAPE reactivities are output for genes that have at least on annotated protein binding site in .map file format, one file per gene. Reactivities across genes are also combined in bedgraph and bigWig format (for + and – strands separately) for the purpose of viewing reactivities on a genome browser.

The map file format is a tab-delimited file with four columns and one row per base. Columns are described in Table 7. Traditionally, reactivity values of "-999.0" are understood to indicate "no data." High fSHAPE reactivities (typically > 2.0) indicate bases that hydrogen bond with protein. High SHAPE reactivities (typically > 0.85) indicate bases that are most likely unpaired. Reactivity measurements are most certain for bases having higher coverage (>200 reads) and good agreement between replicates. For SHAPE reactivities, map files contain everything the user needs to predict structures informed by structure probing data, such as with RNAstructure (DiChiacchio and Mathews, 2016) or RNAfold (Lorenz et al., 2011).

fSHAPE analysis:

fSHAPE requires a very minimal set of bioinformatics tools and python libraries. Open a linux terminal instance and install the following software:

Python >= 3.5.4

Python libraries: sci-kit learn >= 0.20.3, joblib >= 0.16.0, hmmlearn 0.2.1

pybedtools >= 0.8.0

Cutadapt 1.14

UMItools >= 0.5.0

STAR >= 2.4.0i

Bedtools >= 2.25.0

Samtools >= 1.9

Bamtools 2.3.0

If this software cannot be successfully installed for some reason, it may work to build a linux-based Docker image that includes all the required software, or to launch an AWS EC2 instance or equivalent cloud computing instance and install the software there.

Table 7. Example of .map output format

Relative position in region	Reactivity value, averaged across replicates	Standard error	Base
30	-999.0	0.02499	T
31	0.7903	0.00035	T
32	2.65775	0.09731	G
33	1.54583	0.2568	C
34	4.80619	0.96297	T
35	-0.4285	0.13546	T

Download scripts and example data for fSHAPE analysis with:

git clone https://github.com/meracorley/fSHAPE

Locate example fastq files for each sample condition. There should be "in vitro/-protein" condition samples and "in vivo/+protein" condition samples, at least two replicates each for a total of at least four fastq files. Example fastq files to try out the pipeline are included in the fSHAPE directory: K562_vitro_N1.ftl.fastq.gz, K562_vitro_N2.ftl.fastq.gz, K562_vivo_N1.ftl.fastq.gz,K562_vivo_N2.ftl.fastq.gz

Create a directory for storing analyzed data and cd into it. Move fastq files into this directory.

Download reference fasta file, hg38 in the case of the example files:

$ wget https://hgdownload.cse.ucsc.edu/goldenpath/hg38/bigZips/hg38.fa.gz

$ gunzip hg38.fa.gz

> *Note:* All the following commands are time and resource intensive, and for shared compute environments these commands should be submitted as jobs with qsub or equivalent. On private setups it is recommended to run the commands flanked by "nohup <command> &" or in a Screen session.

Build a STAR reference for the reference (requires > 32 Gb of memory for human genome):

$ STAR –runThreadN 8 –runMode genomeGenerate –genomeDir STAR_hg38 --genomeFastaFiles hg38.fa

Raw reads for each sample must first be trimmed, mapped, de-duplicated, and coverage quantifiedover each chromosome in the reference genome. This is achieved by running the processReads Pipeline.sh script as follows (use "/bin/bash processReadsPipeline.sh" -h for help message):

$ /bin/bash processReadsPipeline.sh -i sample1.fastq,sample2.fast2,…,sampleN.fastq -r STAR_reference -a adapterSequence -u UMIpattern -p threads(optional) -l logfile.txt

See UMItools documentation for how to specify the UMI specific to your reads.

With the example fastq files the processReadsPipeline command would be as follows:

$ /bin/bash processReadsPipeline.sh -i K562_vitro_N1.ftl.fastq,K562_vitro_N2.ftl.fastq,K562_vivo_N2.ftl.fastq,K562_vivo_N1.ftl.fastq -r STAR_hg38/ -a AGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCTCGTATGCCGTCTTCTGCTTG -u NNNCCCCNNNNN -p 8 -l LOGFILE

This calculates coverages across each chromosome for each input fastq file, which are output to ./sampleN/coverage/. For example, for K562_vitro_N1.ftl its coverage files are output to ./K562_vitro_N1.ftl/coverage/

For very large input fastq files, running a separate instance of processReadsPipeline.sh per sample will complete in a more reasonable amount of time, i.e., 24 h.

Generate reactivities for regions or full transcripts with bedReactivities.py:

```
$ python bedReactivities.py -i input_bed_regions.bed -g reference_genome.fa -v
invivo_sample1/coverage/,invivo_sample2/coverage/,...,invivo_sampleNcoverage/ -t
invitro_sample1/coverage/,...invitro_sampleN/coverage/
```

The unless the "-s" option is used, the input bed regions are grouped together by name (transcript identifier in column 4 of the bed region) such that reactivities are grouped, calculated, and normalized by transcript.

With the provided example files, generating reactivities for the mRNA *FTL* would be done as follows:

```
$ python bedReactivities.py -i ../fSHAPE/ftl.bed -g hg38.fa -v
K562_vivo_N1.ftl/coverage/,K562_vivo_N2.ftl/coverage/ -t
K562_vitro_N1.ftl/coverage/,K562_vitro_N2.ftl/coverage/
```

This outputs .map reactivity files for each transcript with regions in the input bed file. If a transcript has no coverage, however, no files are output for it. The map file format is a tab-delimited file with four columns and one row per base. Columns are described in Table 7.

Traditionally, reactivity values of "-999.0" are understood to indicate "no data." High fSHAPE reactivities (typically > 2.0) indicate bases that hydrogen bond with protein. fSHAPE reactivities are most certain for bases with higher coverage (>200 reads) and good agreement between replicates.

LIMITATIONS

(f)SHAPE-eCLIP will yield unreliable results for samples that begin with an inadequate number of cells (<∼40 million). Some cell types, such as motor neurons, are intrinsically sparse in RNA content, and thus will be difficult to generate (f)SHAPE-eCLIP libraries for. Antibody quality is also key, where the antibody for the desired protein being studied should be verified beforehand. Low quality antibodies may exhibit non-specific binding, such that the RNA molecules they immunoprecipitate are not in fact bound by the protein of interest.

The fSHAPE library preparation protocol uses cDNA circularization to generate the initial DNA template for PCR, however CircLigase has known sequence biases (poorly ligates terminal cytosines). This impacts both the overall yield of the ligation reaction and may also bias RT-stops against cytosines. fSHAPE produces the most reliable data for transcripts with high expression, although the sensitivity of fSHAPE can be increased with higher sequencing depths. Our fSHAPE libraries were sequenced to depths of 200 million per sample condition, and transcripts with expression above ∼200 transcripts per million (TPM) typically produced at least partial fSHAPE profiles.

TROUBLESHOOTING

Problem 1

No or unexpected band on (f)SHAPE-eCLIP western blot, step 25k.

Potential solution

If no bands appear on the western blot, this indicates that the western blot was done incorrectly, or the antibody is ineffective, or the protein being targeted by the antibody is simply not expressed in the given cell type. The Input samples are generally expected to resemble the IP samples in the Western blot, but will occasionally show additional band(s) that are due to non-specific recognition of abundant proteins in the lysate, such as for the SLBP Input (Figure 1A). This is not concerning. Multiple bands in the IP sample could indicate multiple isoforms of the desired protein, but could also indicate non-specific immunoprecipitation of other proteins. We highly recommend IP validating antibodies prior to performing (f)SHAPE-eCLIP and comparing the Western to the vendor's data on the given antibody. If an antibody fails validation, it is best to find an alternative antibody that can both

withstand the increased stringency and varied environments of (f)SHAPE-eCLIP and deliver a clean pull-down. Alternatively, epitope-tagging of proteins with V5 or Halo tags could be a viable option to circumvent antibody viability issues (England et al., 2015). Note that this may introduce unforeseen binding effects due to potential conformational changes to the protein of interest.

Problem 2

Excessive adapter dimer in (f)SHAPE-eCLIP library, step 49e.

Potential solution

Incomplete removal of adapter dimer will appear in the final library trace as a sharp peak ~164 bp (Figure 4). Adapter dimers will consume sequencing resources at the expense of the desired libraries and should be removed. We generally tolerate adapter dimers that make up < 10% of the integrated area for any given sample. To remove primer dimers repeat the gel extraction, adjusting the lower boundary of the excision window higher to more confidently remove the adapter dimers. Gel extraction causes an estimated ~50% sample loss, so samples with low library yield (5–10 nM) may require an additional PCR with a few cycles prior to the secondary gel extraction, following times and temperatures of the latter Denaturation and Annealing/Extension steps in the PCR instructions in Table 4.

Problem 3

Final (f)SHAPE-eCLIP libraries are low-yield, quantified at < 5 nM, step 49e.
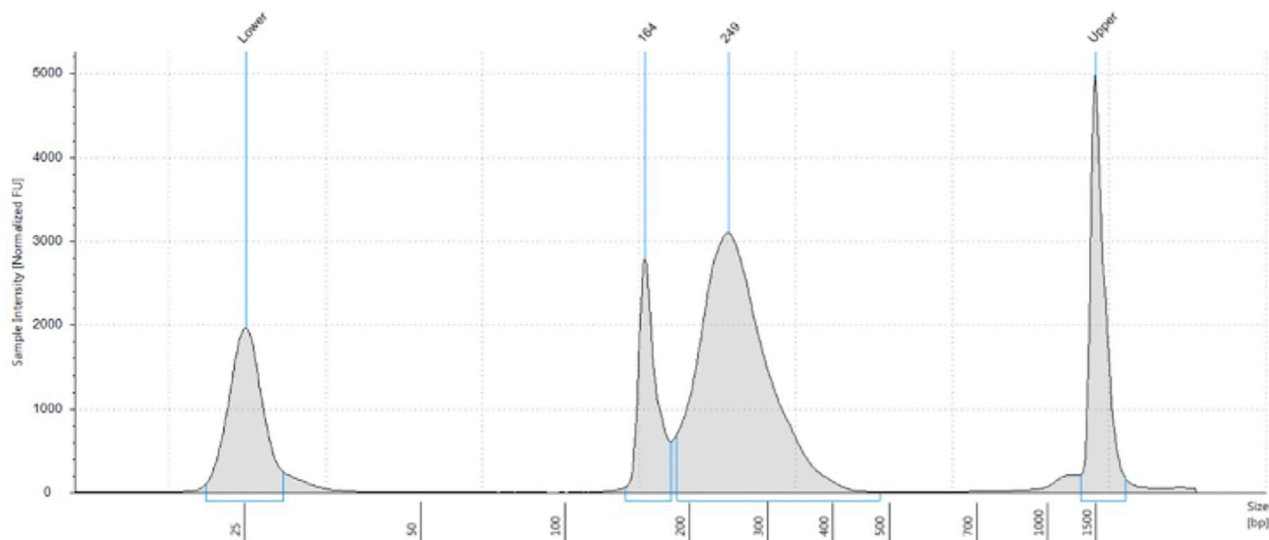
Sample Table

Well	Conc. [ng/µl]
E1	23.3

Peak Table

Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	6.12	-	377	-		Lower Marker
164	4.45	-	41.7	19.13		
249	18.8	-	116	80.87		
1500	6.50	6.50	6.67	-		Upper Marker

Figure 4. Example (f)SHAPE-eCLIP library contaminated with primer dimers

Potential solution

Low library yield could indicate that the protein of interest either does not bind RNA or binds very few specific targets. If the original PCR cycle number was relatively low, 2 or 3 additional PCR cycles can be done to the existing amplified material to increase yield. However, if the original PCR cycle number was relatively high (12+ for input, 18+ for IP), additional amplification will likely only create PCR duplicates and it is best to pool as many femtomoles of library as your pooling strategy allows and to concentrate the pool by bead purification, if necessary to meet sequencing requirements. It may be best to restart the experiment, scaling up the starting material, to increase the probability of sufficiently high yield for sequencing.

Problem 4

Low uniquely mapped yield (<50%) after sequencing and data analysis.

Potential solution

Having low uniquely mapped read yields after mapping and analyzing reads is an indicator that the library had a large number of PCR duplicates. This means that the library underwent overamplification during PCR. Unfortunately, the best way to solve this problem is to re-make the library, scaling up the starting material and taking care not to use too many cycles during library PCR amplification. Sequencing the sample much more deeply, up to 100 million reads per library, will also produce more reads on unique molecules, but at the cost of millions of reads wasted on more PCR duplicates.

Problem 5

fSHAPE, Excessive short reads in library

Excessive short products in fSHAPE library, refer to Flynn et al., 2016.

Potential solution

fSHAPE library reads less than 150 bp in length will prove difficult to map and process in the analysis steps. Too large a fraction of short reads in the final libraries may result from a poorly executed gel-cut in the post-PCR purification PAGE gel. Short cDNAs may also come from too-high levels of SHAPE modifications, causing early truncation and reducing the cDNA length that is possible. Less time with the SHAPE reagent could improve this. Short cDNAs may also come from generating short RNA fragments during the fragmentation step after mRNA enrichment. Reducing the time for RNA fragmentation and lengthening the average fragment could help improve this.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Gene Yeo (geneyeo@ucsd.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All fSHAPE, fSHAPE-eCLIP, and SHAPE-eCLIP raw sequencing reads and final processed data generated during this study are available for download in the Gene Expression Omnibus (GEO), accession number GEO: GSE149767. All custom software used to process this data is available at github.com/meracorley/fshape, github.com/meracorley/f-shape-eclip/tree/pipeline/, and github.com/YeoLab/f-shape-eclip, and upon request.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xpro.2021.100762.

AUTHOR CONTRIBUTIONS

Conceptualization, M.C., R.A.F., H.Y.C., and G.W.Y.; experimentation, M.C. and S.M.B.; analysis, M.C. and B.A.Y.; writing, M.C., R.A.F., and G.W.Y.; supervision, G.W.Y.

DECLARATION OF INTERESTS

G.W.Y. is co-founder, member of the Board of Directors, equity holder, and paid consultant for Locana and Eclipse Bioinnovations. G.W.Y. is a Distinguished Visiting Professor at the National University of Singapore. The terms of this arrangement have been reviewed and approved by the University of California, San Diego, in accordance with its conflict of interest policies. M.C. and G.W.Y. are inventors on a patent disclosure to UCSD on technologies described in this manuscript. H.Y.C. is an inventor on a patent owned by Stanford University on in vivo SHAPE. H.Y.C. is a co-founder of Accent Therapeutics and Boundless Bio and an advisor for 10× Genomics, Arsenal Biosciences, and Spring Discovery. H.Y.C. is an Investigator of the Howard Hughes Medical Institute.

REFERENCES

Barnett, D.W., Garrison, E.K., Quinlan, A.R., Strömberg, M.P., and Marth, G.T. (2011). BamTools. Bioinformatics 27, 1691–1692.

Corley, M., Flynn, R.A., Lee, B., Blue, S.M., Chang, H.Y., and Yeo, G.W. (2020). Footprinting SHAPE-eCLIP Reveals Transcriptome-wide Hydrogen Bonds at RNA-Protein Interfaces. Mol. Cell 80, 903–914.

DiChiacchio, L., and Mathews, D.H. (2016). Predicting RNA–RNA Interactions Using RNAstructure. In RNA Structure Determination (Humana Press), pp. 51–62.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.

England, C.G., Luo, H., and Cai, W. (2015). HaloTag technology: a versatile platform for biomedical applications. Bioconj. Chem. 26, 975–986.

Flynn, R.A., Zhang, Q.C., Spitale, R.C., Lee, B., Mumbach, M.R., and Chang, H.Y. (2016). Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE. Nat. Protoc. 11, 273–290.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

Lorenz, R., Bernhart, S.H., Zu Siederdissen, C.H., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol. Biol. 6, 1–14.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. 17, 10–12.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.

Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res. 27, 491–499.

Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R., Blue, S.M., Chen, J.Y., Cody, N.A., Dominguez, D., et al. (2020). A large-scale binding and functional map of human RNA-binding proteins. Nature 583, 711–719.