nature chemical biology

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41589-018-0121-5

In the format provided by the authors and unedited.

Continuous directed evolution of proteins with improved soluble expression

Tina Wang[1,2], Ahmed H. Badran[1,2], Tony P. Huang[1,2] and David R. Liu[1,2,3]*

[1]Merkin Institute of Transformative Technologies in Healthcare, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [2]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA. [3]Howard Hughes Medical Institute, Harvard University, Cambridge, MA, USA.
*e-mail: drliu@fas.harvard.edu

Supplementary Table 1. Properties of purified anti-GCN4 scFv variants.

scFv	T_m (°C)	Yield (mg)	EC_{50} (nM) [a]
Ωg	59.3	0.02 ± 0.013 [b]	4.4
29.1.2	57.1	0.11 ± 0.033 [b]	26.3
29.1.5	57.0	0.10 ± 0.019 [b]	4.4
m3	63.5	0.09 ± 0.012 [b]	4.7
Ωg L224V	n.d.	0.05 [b]	n.d.
Ωg G103S	n.d.	0.03 [b]	n.d.
Ωg F87S	n.d.	0.06 [b]	n.d.
m3-noCys [c]	52.4	0.06 [d]	10.7
58.1.1 [c]	57.5	n.d.	87.7
58.1.2 [c]	57.5	n.d.	67.8
58.2.8 [c]	51.8	0.06 [d]	n.d.
29.1.5-noCys [c]	40.8	0.06 [d]	n.d.
58.3.1 [c]	51.3	0.18 [d]	88.6
58.4.2 [c]	44.8	0.12 [d]	15.7
58.4.4 [c]	47.4	0.22 [d]	81.2

[a] Measured by ELISA using MBP–TEV–GCN4 7P14P as the antigen.
[b] Yield isolated from 220 mL overnight culture in LB at 16 °C, given as the mean ± standard deviation of three biological replicates (purifications performed on three different days using starter cultures from three different colonies)
[c] Disulfide-free scFvs.
[d] Yield isolated from 250 mL overnight culture in 2xYT at 16 °C

Supplementary Table 2. Properties of purified anti-htt scFv variants.

scFv	T_m (°C)	Yield (mg) [a]	EC_{50} (nM) [b]
C4	59.0	0.1	34
C4 V38F	51.5	0.3	n.d.
C4 A98V	55.0	0.4	n.d.
34.1.2	n.d.	n.d.	67
34.2.3	55.0	0.6	26
34.2.6	56.0	0.5	155

[a] Yield isolated from 250 mL culture in 2xYT at 37 °C.
[b] Measured by ELISA using biotinylated htt exon 1 peptide as the antigen.

Supplementary Table 3. Properties of purified MBP variants.

protein	T_m (˚C)	Yield (mg) [a]
MBP	61.5	1.2
MBP G32D/I33P	58.3	0.9
50.1.5	59.0	1.2
50.2.1	57.5	0.9
50.2.2	60.1	1.1

[a] Yield isolated from 250 mL culture in 2xYT at 20 °C.

Supplementary Table 4. Yields of purified BE3 variants.

protein	Yield (mg) [a]	
	Replicate 1	Replicate 2
BE3 (wt)	0.25	0.21
36.1-BE3	0.45	0.33
43.1-BE3	0.71	0.54
43.2-rev-BE3	0.56	0.39

[a] Yield isolated from 200 mL culture in 2xYT at 16 °C. Replicates denote purifications performed on two separate days using starter cultures from two different colonies.

Supplementary Table 5. Plasmids used in this work.

			ORF1		ORF2		ORF3	
Name	Class (res)	Origin	Prom	[RBS][1] Genes	Prom	[RBS] Genes	Prom	[RBS] Genes
pTW006a	AP ($carb^R$)	SC101	P_{T7}	[SD8] luxAB	P_{BAD}	[SD8] T7c	P_c	araC
pTW006a2	AP ($carb^R$)	SC101	P_{T7}	[SD8] luxAB	P_{BAD}	[SD8] T7c R632S	P_c	araC
pTW006a3	AP ($carb^R$)	SC101	P_{T7}	[SD8] luxAB	P_{BAD}	[SD8] T7c L637A	P_c	araC
pTW006a4	AP ($carb^R$)	SC101	P_{T7}	[SD8] luxAB	P_{BAD}	[SD8] T7c Y639F	P_c	araC
pTW006a5	AP ($carb^R$)	SC101	P_{T7}	[SD8] luxAB	P_{BAD}	[SD8] T7c Q649S	P_c	araC
pTW006a6	AP ($carb^R$)	SC101	P_{T7}	[SD8] luxAB	P_{BAD}	[SD8] T7c F644A	P_c	araC
pTW006aP1a	AP ($carb^R$)	SC101	P_{T7}	[SD8] gIII, luxAB	P_{pro1}	[SD8] T7c		
pJC175e[2]	AP ($carb^R$)	SC101	P_{psp}	[SD8] gIII, luxAB				
pTW026a3c	AP ($carb^R$)	SC101	P_{psp}	[SD8] gIII-N				
pTW048a3	AP ($carb^R$)	SC101	$P_{lacZ\text{-}opt}$ (OR1+2)	[sd5] gIII-N	P_{pro1}[3]	434cI–$SH2_{ABL1}$		
pTW055a3	AP ($carb^R$)	SC101	$P_{lacZ\text{-}opt}$ (OR1+2)	[sd8] gIII-N	P_{pro1}	434cI–GCN4 7P14P		
pTW074c	AP ($carb^R$)	SC101	$P_{lacZ\text{-}opt}$ (OR1)	[sd5] gIII-N	P_{pro1}	434cI–htt 1-17		
pTW084b	AP ($carb^R$)	SC101	$P_{lacZ\text{-}opt}$ (OR1+2)	[sd5] gIII-N	P_{pro1}	434cI–m3		
pTW032b2c	AP ($spec^R$)	ColE1	P_{psp}	[sd5] gIII-C				
pTW051a	AP ($spec^R$)	ColE1	P_{T7}	[sd5] gIII-C	P_{pro1}	T7c		
pTW051b2	AP ($spec^R$)	ColE1	P_{T7a}	[sd5] gIII-C	P_{pro1}	T7c R632S		
pTW051b4	AP ($spec^R$)	ColE1	P_{T7d}	[sd5] gIII-C	P_{pro1}	T7c R632S		
pTW051d	AP ($spec^R$)	ColE1	P_{T7a}	[sd5] gIII-C	P_{pro1}	T7c R632S+Q649S		
pTW051d2	AP ($spec^R$)	ColE1	P_{T7d}	[sd5] gIII-C	P_{pro1}	T7c R632S+Q649S		
pTW004a	CP ($spec^R$)	ColE1	P_{tet}	[SD8] MBP–T7n				
pTW004b	CP ($spec^R$)	ColE1	P_{tet}	[SD8] MBP G32D+I33P–T7n				
pTW004d	CP ($spec^R$)	ColE1	P_{tet}	[SD8] MBP V8G–T7n				
pTW004e	CP ($spec^R$)	ColE1	P_{tet}	[SD8] MBP G19C–T7n				
pTW004f	CP ($spec^R$)	ColE1	P_{tet}	[SD8] MBP A276G–T7n				
pTW004g	CP ($spec^R$)	ColE1	P_{tet}	[SD8] MBP Y283D–T7n				

pTW004h	CP ($spec^R$)	ColE1	P_{tet}	[SD8] MBP T345I–T7n		
pTW035d	EP (kan^R)	pBR322	P_{T7lac}	6xHis–MBP G32D+I33P		
pTW035e	EP (kan^R)	pBR322	P_{T7lac}	6xHis–MBP		
pTW035g	EP (kan^R)	pBR322	P_{T7lac}	6xHis–MBP V8G		
pTW035h	EP (kan^R)	pBR322	P_{T7lac}	6xHis–MBP G19C		
pTW035i	EP (kan^R)	pBR322	P_{T7lac}	6xHis–MBP A276G		
pTW035j	EP (kan^R)	pBR322	P_{T7lac}	6xHis–MBP Y283D		
pTW035b	EP (kan^R)	pBR322	P_{T7lac}	6xHis–MBP P33T+T275I		
pTW035c	EP (kan^R)	pBR322	P_{T7lac}	6xHis–MBP P33T+V76I+ A167V+V373I		
pTW035aa	EP (kan^R)	pBR322	P_{T7lac}	6xHis–MBP D33T+F258Y		
pTW035ab	EP (kan^R)	pBR322	P_{T7lac}	6xHis–MBP D33L		
pTW035ac	EP (kan^R)	pBR322	P_{T7lac}	6xHis–MBP G24V+D33S		
pTW069a	EP (kan^R)	pBR322	P_{T7lac}	6xHis–Ωg		
pTW069h	EP (kan^R)	pBR322	P_{T7lac}	6xHis–m3		
pTW075l	EP (kan^R)	pBR322	P_{T7lac}	6xHis–29.1.2		
pTW075k	EP (kan^R)	pBR322	P_{T7lac}	6xHis–29.1.5		
pTW087a	EP (kan^R)	pBR322	P_{T7lac}	6xHis–MBP–TEV–GCN4 7P14P		
pTW081a	EP (kan^R)	pBR322	P_{T7lac}	6xHis–C4		
pTW081d	EP (kan^R)	pBR322	P_{T7lac}	6xHis–C4 V38F		
pTW081e	EP (kan^R)	pBR322	P_{T7lac}	6xHis–C4 A98V		
pTW081i	EP (kan^R)	pBR322	P_{T7lac}	6xHis–34.1.2		
pTW081j	EP (kan^R)	pBR322	P_{T7lac}	6xHis–34.2.3		
pTW081k	EP (kan^R)	pBR322	P_{T7lac}	6xHis–34.2.6		
pTW101a	EP ($spec^R$)	ColE1	$rhaP_{BAD}$	6xHis–rAPOBEC1	$rhaP_{RS}$	rhaS, rhaR
pTW101h	EP ($spec^R$)	ColE1	$rhaP_{BAD}$	6xHis–36.1	$rhaP_{RS}$	rhaS, rhaR
pTW101t	EP ($spec^R$)	ColE1	$rhaP_{BAD}$	6xHis–43.1	$rhaP_{RS}$	rhaS, rhaR
pTW101m	EP ($spec^R$)	ColE1	$rhaP_{BAD}$	6xHis–43.2	$rhaP_{RS}$	rhaS, rhaR

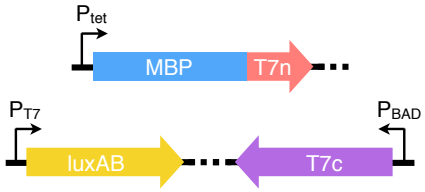
pTW101m2	EP (spec^R)	ColE1	rhaP$_{BAD}$	6xHis–43.2-rev	rhaP$_{RS}$	rhaS, rhaR		
pHR41[4]	EP (kan^R)	pBR322	P$_{T7lac}$	6xHis–BE3				
pTW142a	EP (kan^R)	pBR322	P$_{T7lac}$	6xHis–36.1-BE3				
pTW142b	EP (kan^R)	pBR322	P$_{T7lac}$	6xHis–43.1-BE3				
pTW142c	EP (kan^R)	pBR322	P$_{T7lac}$	6xHis–43.2-rev-BE3				
pNMG98[5]	EP (carb^R)	SC101	P$_{lac}$	gRNA	P$_{BAD}$	BE2	P$_c$	araC
pTW143a	EP (carb^R)	SC101	P$_{lac}$	gRNA	P$_{BAD}$	36.1-BE2	P$_c$	araC
pTW143b	EP (carb^R)	SC101	P$_{lac}$	gRNA	P$_{BAD}$	43.1-BE2	P$_c$	araC
pTW143c	EP (carb^R)	SC101	P$_{lac}$	gRNA	P$_{BAD}$	43.2-rev-BE2	P$_c$	araC
pTW143d	EP (carb^R)	SC101	P$_{lac}$	gRNA	P$_{BAD}$	43.2-BE2	P$_c$	araC
antibiotic selection uracil[5]	spec^R	RSF1030	P$_{kan}$	kan^R	P$_{cat}$	cam^R H193R		
pACK129[6]	carb^R	pBR322	P$_{CMV}$	BE3				
pTW107a	carb^R	pBR322	P$_{CMV}$	36.1 BE3				
pTW107b	carb^R	pBR322	P$_{CMV}$	43.1 BE3				
pTW107c	carb^R	pBR322	P$_{CMV}$	43.2-rev BE3				
MP6[7]	MP (chlor^R)	cloDF13	P$_{BAD}$	dnaQ926, dam, seqA, emrR, ugi, cda1	P$_c$	araC		
DP6[7]	MP (chlor^R)	cloDF13	P$_{BAD}$	dnaQ926, dam, seqA, emrR, ugi, cda1	P$_c$	araC	P$_{psp-tet}$	[sd8] gIII
SP01a	SP (none)	M13 f1	P$_{gIII}$	MBP–T7n				
SP01z	SP (none)	M13 f1	P$_{gIII}$	T7n				
SP02a	SP (none)	M13 f1	P$_{gIII}$	MBP G32D+ I33P–T7n				
SP02b	SP (none)	M13 f1	P$_{gIII}$	MBP Y283D–T7n				
SP09b3	SP (none)	M13 f1	P$_{gIII}$	HA4–T7n–rpoZ				
SP10b2	SP (none)	M13 f1	P$_{gIII}$	HA4 Y87A–T7n–rpoZ				
SP16c3	SP (none)	M13 f1	P$_{gIII}$	Ωg–eT7n–rpoZ				
SP24a3	SP (none)	M13 f1	P$_{gIII}$	C4–eT7n–rpoZ				
SP27b	SP (none)	M13 f1	P$_{gIII}$	GCN4 7P14P–MBP G32D+ I33P–eT7n–rpoZ				
SP30	SP (none)	M13 f1	P$_{gIII}$	GCN4 7P14P – rAPOBEC1–eT7n–rpoZ				

SP39b	SP (none)	M13 f1	P_{gIII}	m3-noCys–eT7n–rpoZ
SP39d	SP (none)	M13 f1	P_{gIII}	29.1.5-noCys–eT7n–rpoZ
SP13	SP (none)	M13 f1	P_{gIII}	kan
SP98[8]	SP (none)	M13 f1	P_{gIII}	rpoZ–HA4

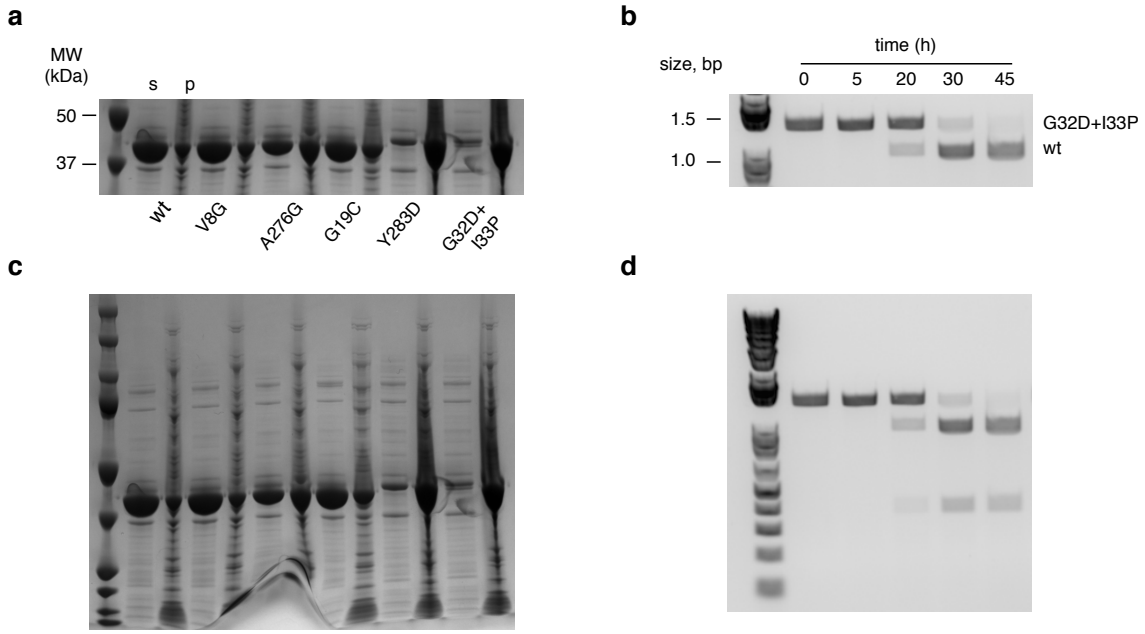
Supplementary Table 6. General reagents and equipment.

Chemicals and media		
Name	Source	Notes
Carbenicillin	Gold Biotechnology	
Spectinomycin	Gold Biotechnology	
Chloramphenicol	Gold Biotechnology	
Kanamycin	Sigma-Aldrich	
Tetracycline	Gold Biotechnology	
Streptomycin	Gold Biotechnology	
Rifampin	Alfa Aesar	
Isopropyl-β-D-thiogalactoside (IPTG)	Gold Biotechnology	
Anhydrotetracycline (aTc)	Fluka	
L-arabinose	Gold Biotechnology	
Rhamnose	Gold Biotechnology	
PEG 3350	Sigma-Aldrich	
Bluo-gal	Gold Biotechnology	
Dithiothreitol (DTT)	Sigma-Aldrich	
MATLEKLMKAFESLKSFK(Biotin)-NH_2	New England Peptide	Custom peptide synthesis
2xYT media	United States Biologicals	
LB media	United States Biologicals	
SOC media	New England BioLabs	
DRM	Reference 3	
Agar	United States Biologicals	
DMEM	Thermo Fisher Scientific	Supplemented with high glucose and GlutaMAX
Opti-MEM	Thermo Fisher Scientific	
Fetal Bovine Serum	Thermo Fisher Scientific	
Reagents and Supplies		
Name	Source	Notes
MinElute PCR Purification Kit	Qiagen	
HyClone water	GE Healthcare Life Sciences	
B-per lysis reagent	Thermo Fisher Scientific	
Protease inhibitor cocktail	Roche	
Laemmli sample loading buffer (4X)	Bio-Rad	
Precision Plus Protein Dual Color Standard	Bio-Rad	
Bolt MES SDS running buffer	Thermo Fisher Scientific	
InstantBlue Ultrafast protein stain	Expedeon	
Bolt 4-12% Bis-Tris Plus pre-cast gel	Thermo Fisher Scientific	
TALON Cobalt resin	Clontech	
Amicon Ultra-15 centrifugal filter unit	Millipore	
Quick Start Bradford reagent	Bio-Rad	
Odyssey blocking buffer in PBS	LI-COR	
SuperBlock Blocking Buffer in TBS	Thermo Fisher Scientific	
Protein Thermal Shift Dye Kit	Life Technologies	
Nunc MaxiSorp 96-well plates	Thermo Fisher Scientific	
1-Step Ultra TMB-ELISA Substrate	Thermo Fisher Scientific	
PVDF Ultrafree centrifugal filter	Millipore	
96-well deep well plate	Axygen	
96-well black-walled clear-bottom plate	Costar	
MaxiSorp 96-well plate	Nunc	
Lipofectamine 2000	Thermo Fisher Scientific	

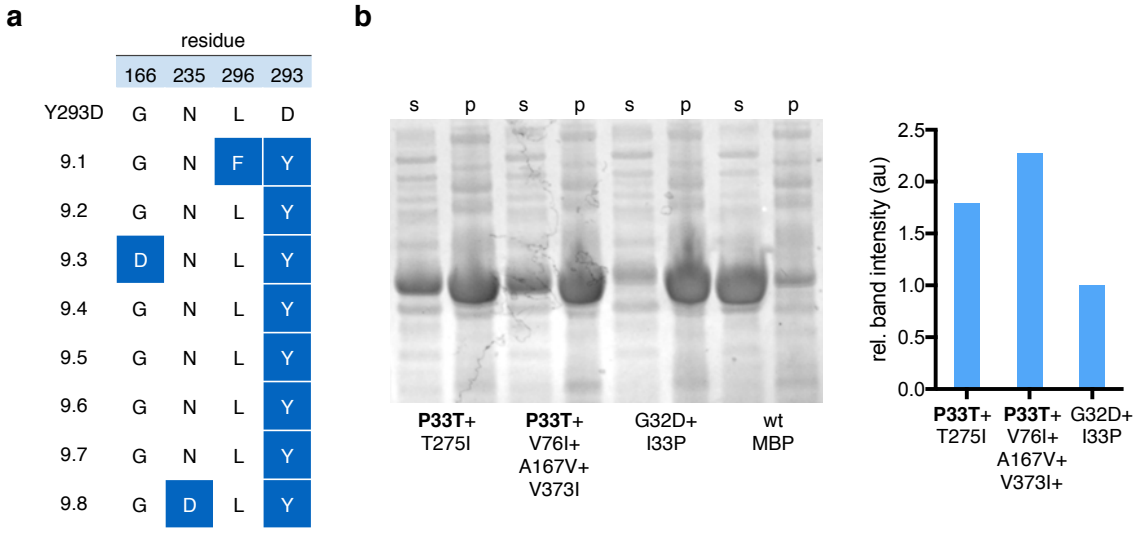
Antibodies and ELISA proteins		
Name	Source	Notes
Anti-6xHis	Abcam	Mouse monoclonal ab18184
Anti-GroEL	Sigma-Aldrich	Rabbit polyclonal G6532
Bovine Serum Albumin (BSA)	Sigma-Aldrich	
Anti-MBP, HRP-conjugated	Abcam	Mouse monoclonal ab49923
Streptavidin, HRP-conjugated	BioLegend	
Anti-mouse 680RD	LI-COR	Goat polyclonal 926-68070
Anti-rabbit 800CW	LI-COR	Donkey polyclonal 926-32213
Enzymes		
Name	Source	Notes
Phusion U Hot Start DNA polymerase	Thermo Fisher Scientific	
USER enzyme	New England BioLabs	
DpnI	New England BioLabs	
BsaWI	New England BioLabs	
Illustra Templiphi 100 Amplification Kit	GE Healthcare Life Sciences	
Cells		
Name	Source	Notes
Mach1	Thermo Fisher Scientific	
Turbo	New England BioLabs	
BL21 DE3	New England BioLabs	
S2060	Reference 4	
HEK293T	ATCC	
Equipment		
Name	Source	Notes
MilliQ water purification system	Millipore	
G:Box Chemi XRQ	Syngene	
iBlot 2 Gel Transfer Device	Life Technologies	
Odyssey Imaging System	LI-COR	
CFX96 Real-Time PCR Detection System	Bio-Rad	
Infinite M1000 Pro microplate reader	Tecan	



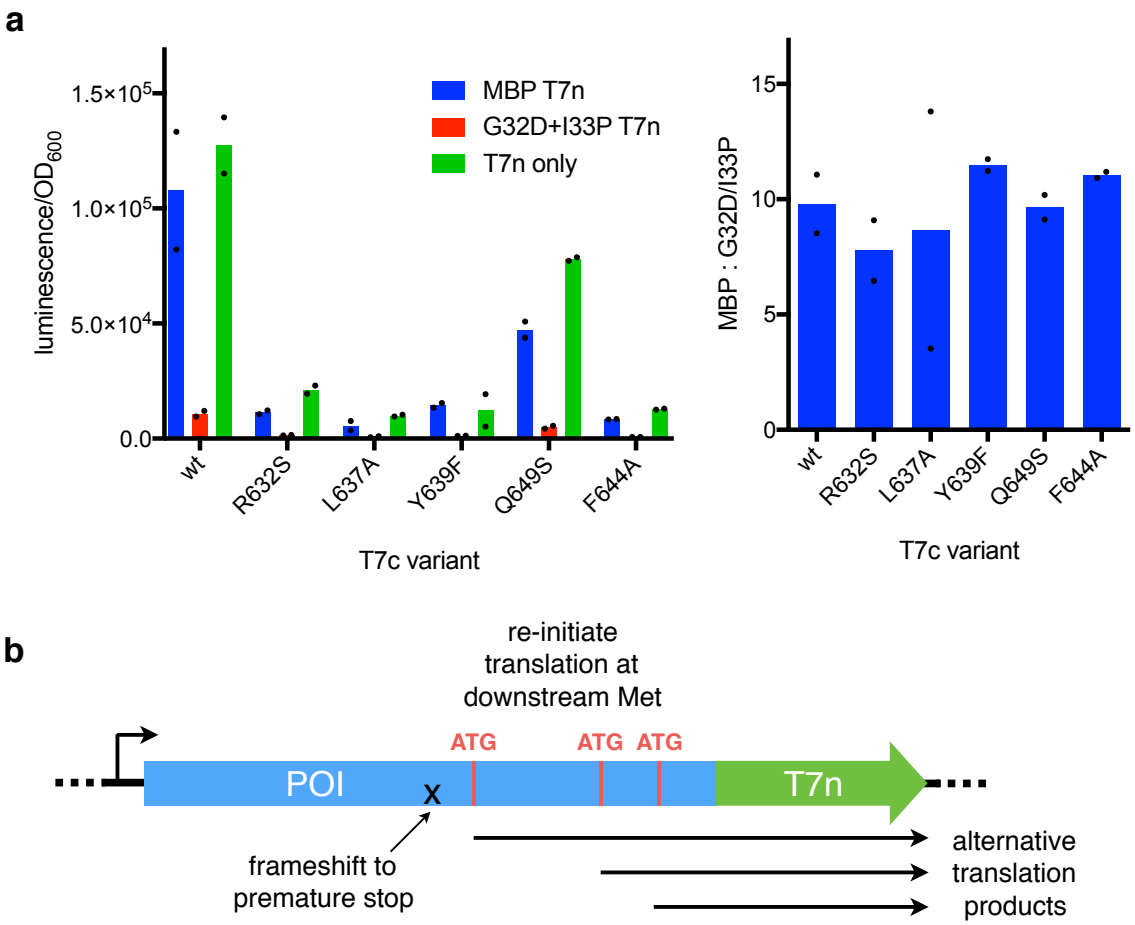
Supplementary Figure 1. Overview of the luciferase-based transcriptional activity reporter.



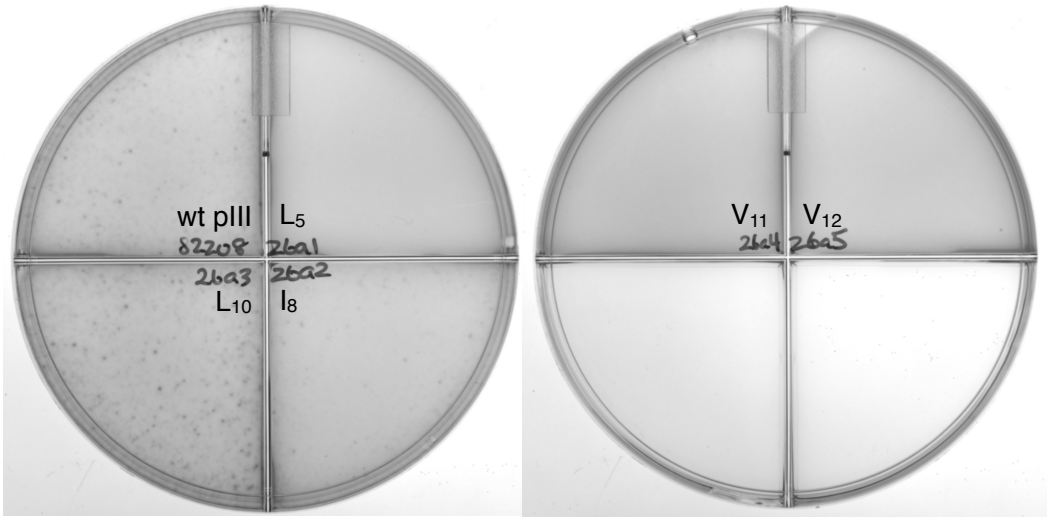
Supplementary Figure 2. Validation of the soluble protein expression PACE selection. (**a**) SDS-PAGE analysis of soluble and insoluble protein extracts of MBP variants expressed in *E. coli* at 37 °C. (**b**) SP encoding the poorly expressed MBP(G32D+I33P) mutant are outcompeted during PACE by SP encoding wild-type MBP, as shown by restriction digestion of PCR products with BsaWI, which cleaves wild-type MBP but not MBP(G32D+I33P). The starting ratio of MBP(G32D+I33P) SP to wild-type MBP SP was 1000:1. (**c**) Full-size image of the gel in panel (**a**). (**d**) Full-size image of the gel in panel (**b**). All experiments in this figure were repeated at least once with similar results.

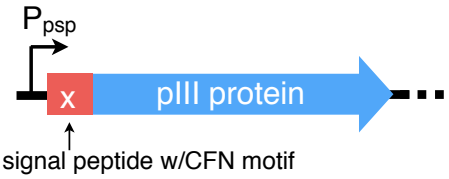


Supplementary Figure 3. SE-PACE of MBP Y283D (**a**) and G32D+I33P (**b**). SP isolated from PACE with MBP Y293D at 72 h show complete reversion to Tyr at position 293. The P33T acquired during MBP(G32D+I33P) SE-PACE confers modest improvements in soluble expression in BL21 DE3 cells at 37 °C. This experiment was repeated once with similar results.



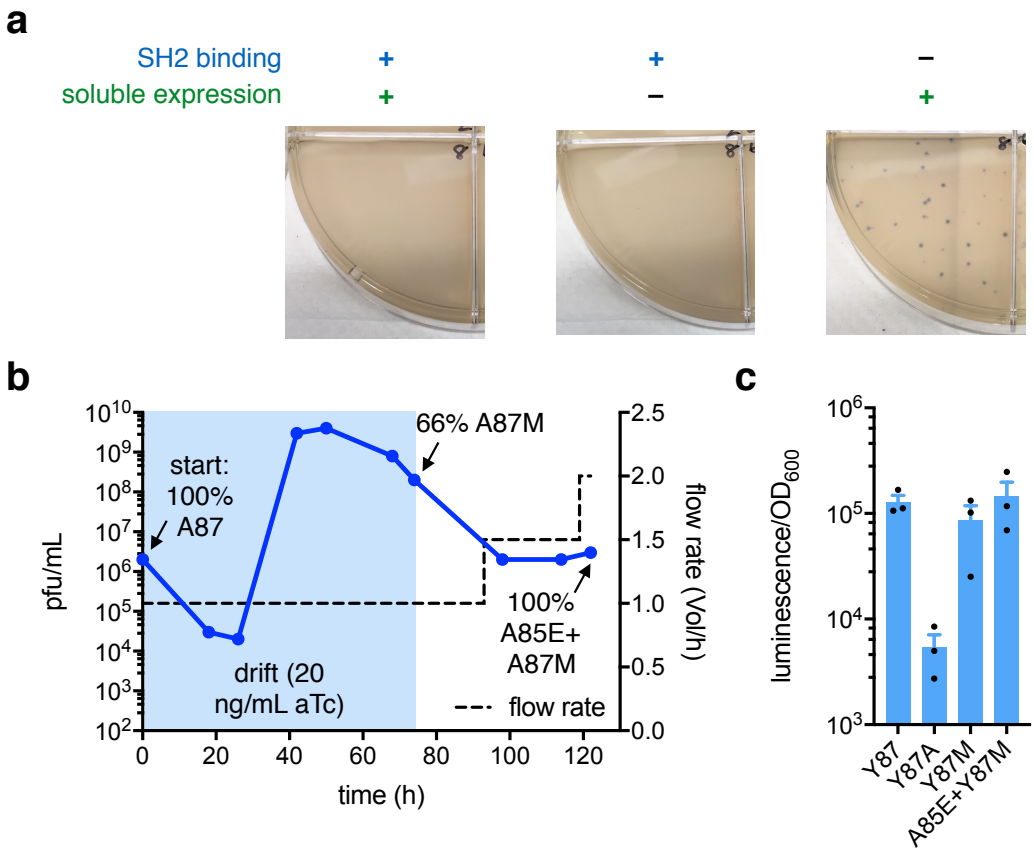
Supplementary Figure 4. Increasing protein soluble expression selection stringency. (**a**) Left: transcriptional activity of split T7 RNAPs harboring inactivating mutations in T7c. Right: Mutations in T7c do not affect the ability of the selection to differentiate between MBP and the MBP(G32D+I33P) folding mutant. Data reflect two technical replicates (unique clones). (**b**) Formation of premature stop codons allows translational re-initiation at adjacent start codons to produce T7n in a manner uncoupled from POI expression.





construct name	insertion pos.	signal peptide sequence	plaque strength
S2208 (pJC175e)	wt	MKKLLFAIPLVVPFYSHS (wt)	+++++
pTW026a1	L_5	MKKLLCFNIPLVVPFYSHS	-
pTW026a2	I_8	MKKLLFAICFNPLVVPFYSHS	++++
pTW026a3	L_{10}	MKKLLFAIPLCFNVVPFYSHS	+++++
pTW026a4	V_{11}	MKKLLFAIPLVCFNVPFYSHS	-
pTW026a5	V_{12}	MKKLLFAIPLVVCFNPFYSHS	-

Supplementary Figure 5. Determination of a split intein insertion site in the signal peptide of pIII. SP containing kan^r in place of gIII were allowed to form plaques on host cells transformed with APs encoding a CFN motif inserted at various positions in the pIII signal peptide sequence. This experiment was repeated once with similar results.

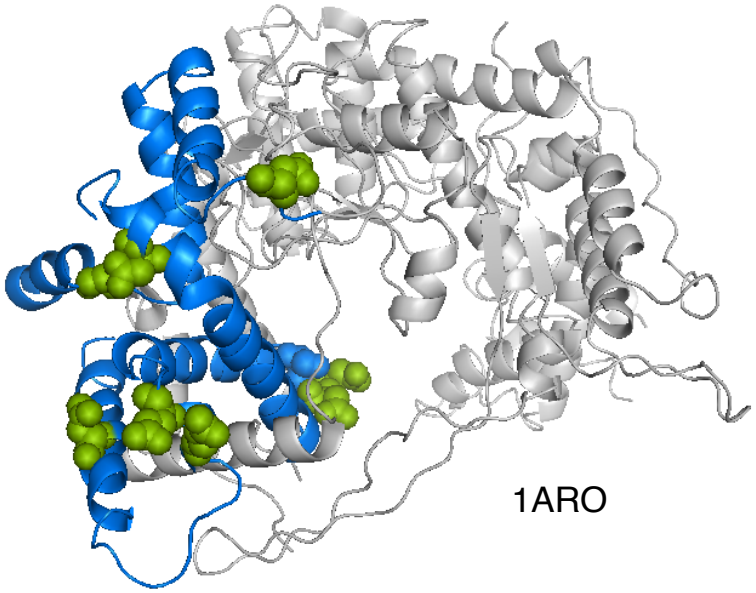


Supplementary Figure 6. Split-intein pIII allows dual positive selections during PACE. (**a**) Plaque activity of SP encoding mutant HA4(Y87A) monobodies that cannot bind the target SH2 domain on host cells transformed with various split-intein pIII APs. These SP pass the protein expression selection but not the protein binding selection, and therefore do not propagate on cells that contain a SH2 binding AP. This experiment was repeated once with similar results. (**b**) SE-PACE of HA4 Y87A enriches for SP encoding Met at position 87. (**c**) Apparent binding activity of evolved HA4 variants to the SH2 domain measured by transcriptional activation by the protein binding PACE selection. Data reflect the mean and s.d. of three technical replicates (unique clones).

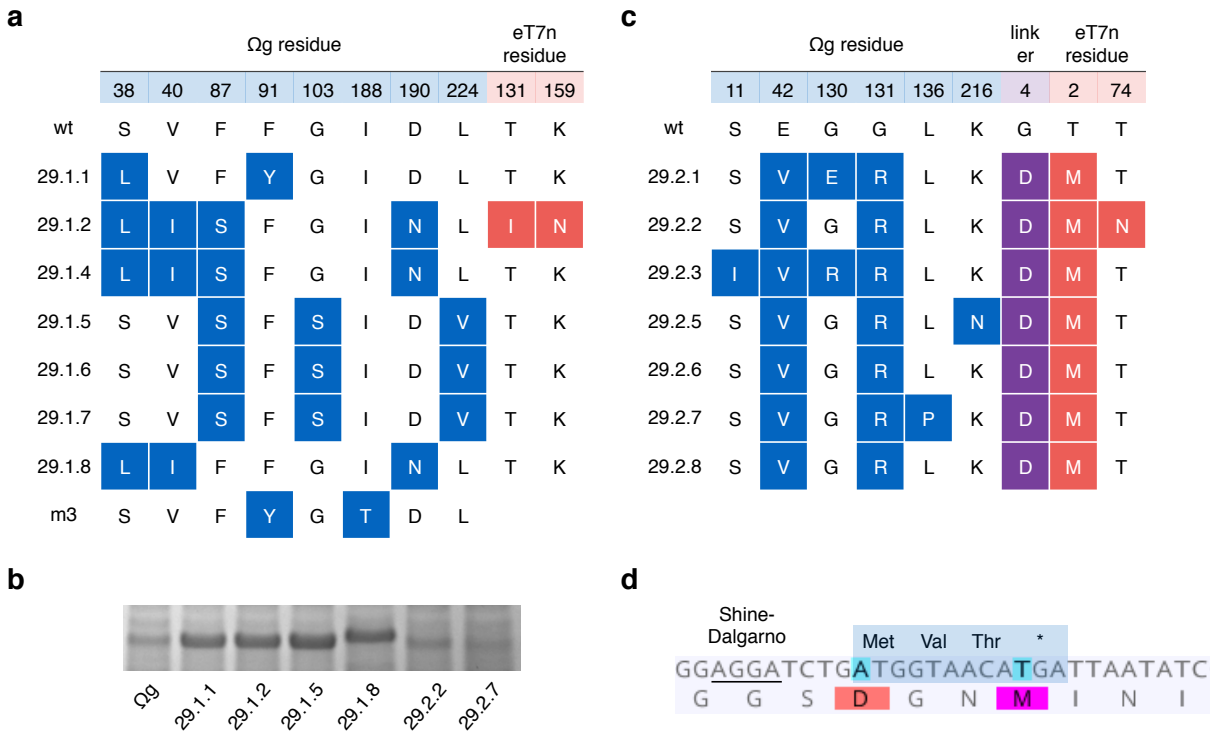
a

	residue number									
	9	46	67	70	85	103	106	129	130	155
wt	D	G	A	K	N	Q	Q	D	N	D
P23.2	D	D	T	K	N	Q	K	D	K	D
P24.2	D	D	T	K	K	Q	K	D	K	D
P27.2 ("eT7n")	G	D	T	N	K	H	K	N	K	Y

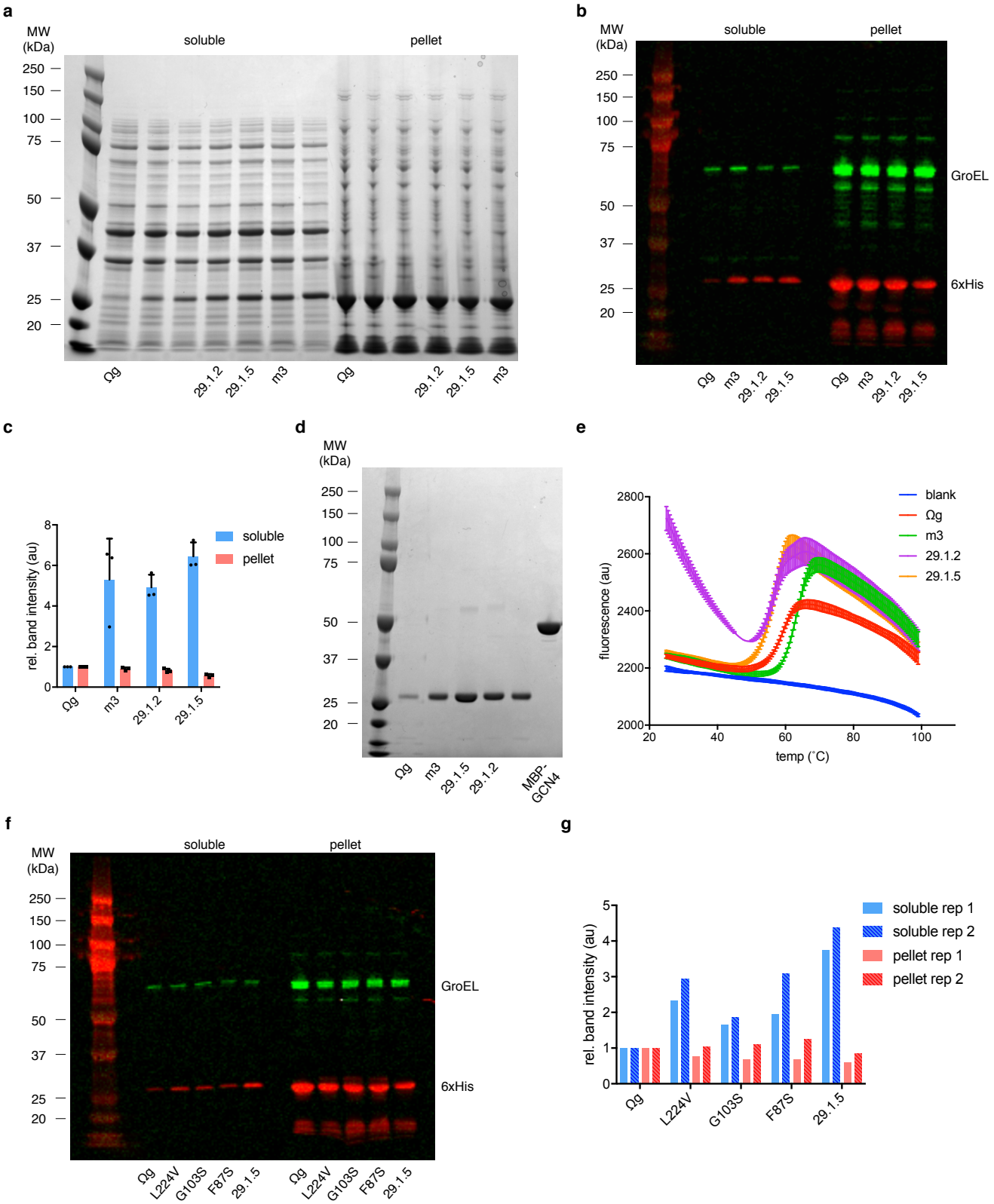
b



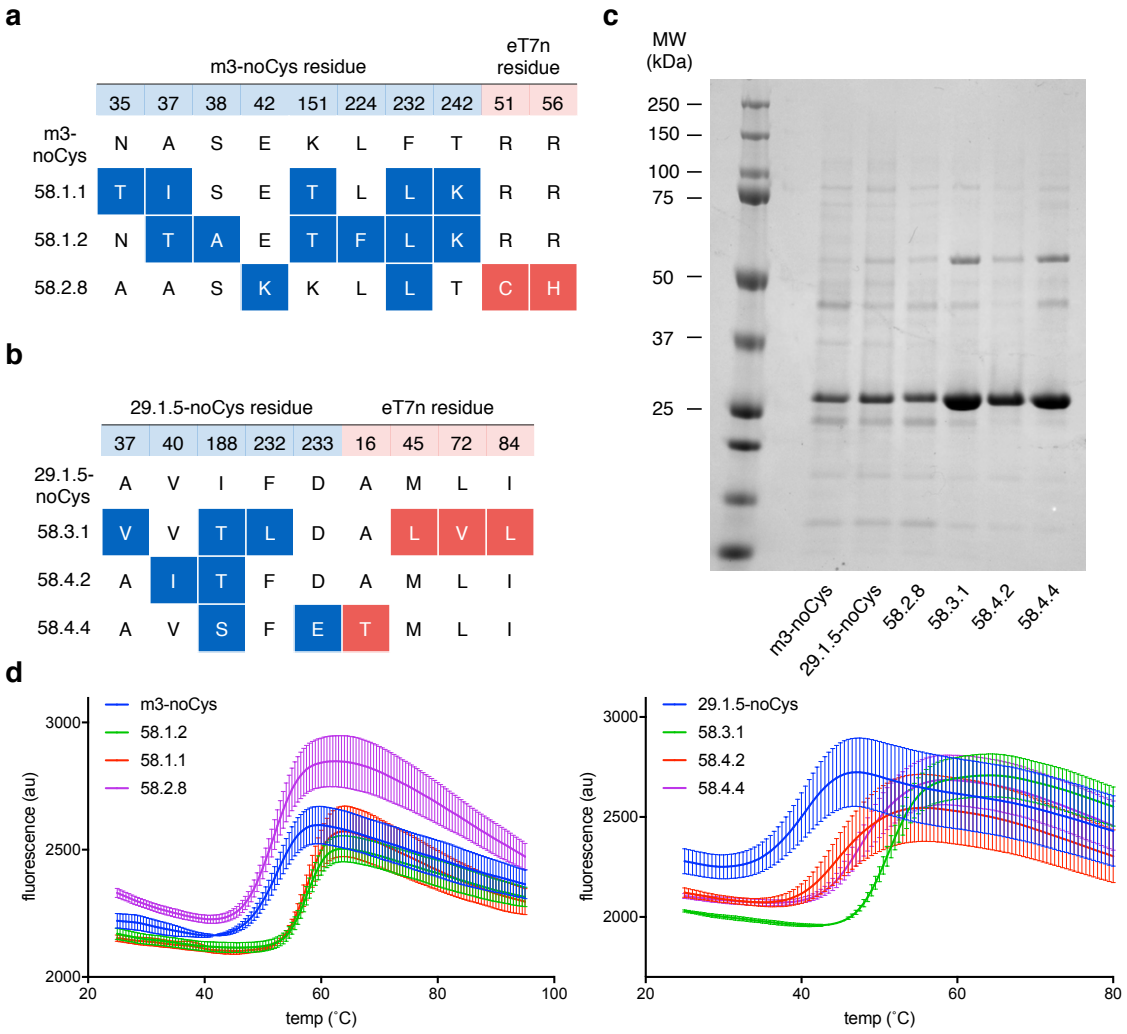
Supplementary Figure 7. Evolution of eT7n. (**a**) T7n was evolved in PACE with increasingly stringent soluble expression APs to yield eT7n. (**b**) Positions mutated in eT7n are shown in green spheres on the structure of T7 RNA polymerase. The region split into T7n is shown in blue. T7c is shown in gray.



Supplementary Figure 8. SE-PACE of anti-GCN4 scFv Ωg. (**a**) SP isolated from lagoon 1 of PACE with anti-GCN4 scFv Ωg at 72 h. (**b**) Soluble expression of Ωg variants in BL21 DE3 cells at 37 °C. This experiment was repeated at least once with similar results. (**c**) SP isolated from lagoon 2 of PACE with anti-GCN4 scFv Ωg at 72 h. (**d**) Evolution of an alternate translation initiation site in lagoon 2.


a
MW (kDa)
soluble
pellet
250
150
100
75
50
37
25
20
Ωg
29.1.2
29.1.5
m3
b
GroEL
6xHis
c
rel. band intensity (au)
d
MBP-GCN4
e
blank
fluorescence (au)
temp (˚C)
f
g
soluble rep 1
soluble rep 2
pellet rep 1
pellet rep 2
L224V
G103S
F87S

Supplementary Figure 9. Expression, purification, and melting temperature determination of anti-GCN4 scFv Ωg variants. (**a**) Expression of anti-GCN4 scFv Ωg variants in BL21 DE3 cells at 37 °C shown in crude lysates separated by SDS-PAGE. (**b**) Western blot of his-tagged variants in BL21 DE3 cells at 37 °C. GroEL is used as a loading control. (**c**) Relative expression levels of anti-GCN4 scFv variants as determined by densitometry of western blotting as shown in (**b**). Data reflect the mean and s.d. of three biological replicates (experiments conducted on different days). (**d**) SDS-PAGE of purified anti-GCN4 scFv Ωg variants and MBP-TEV-GCN4 expressed in BL21 DE3 cells at 16 °C. (**e**) Melting temperature curves of purified anti-GCN4 scFv Ωg variants. Data reflect the mean and s.d. of four technical replicates (four melt curve samples using protein from the same preparation). (**f**) Western blot and (**g**) densitometry of his-tagged variants bearing single point mutations from 29.1.5. Data reflect two biological replicates (experiments conducted on different days). All experiments in this figure were repeated at least once with similar results.



Supplementary Figure 10. Purification and melting temperature determination of disulfide-free anti-GCN4 scFv variants. (**a, b**) SP isolated from PACE with disulfide-free anti-GCN4 scFv variants m3 and 29.1.5 (denoted m3-noCys and 29.1.5-noCys, respectively). (**c**) SDS-PAGE of purified disulfide-free scFvs expressed in BL21 DE3 cells at 16 °C. This purification was performed a single time. (**d**) Melting temperature curves of purified variants. Data reflect the mean and s.d. of four technical replicates (four melt curve samples using protein from the same preparation). This experiment was repeated once with similar results.

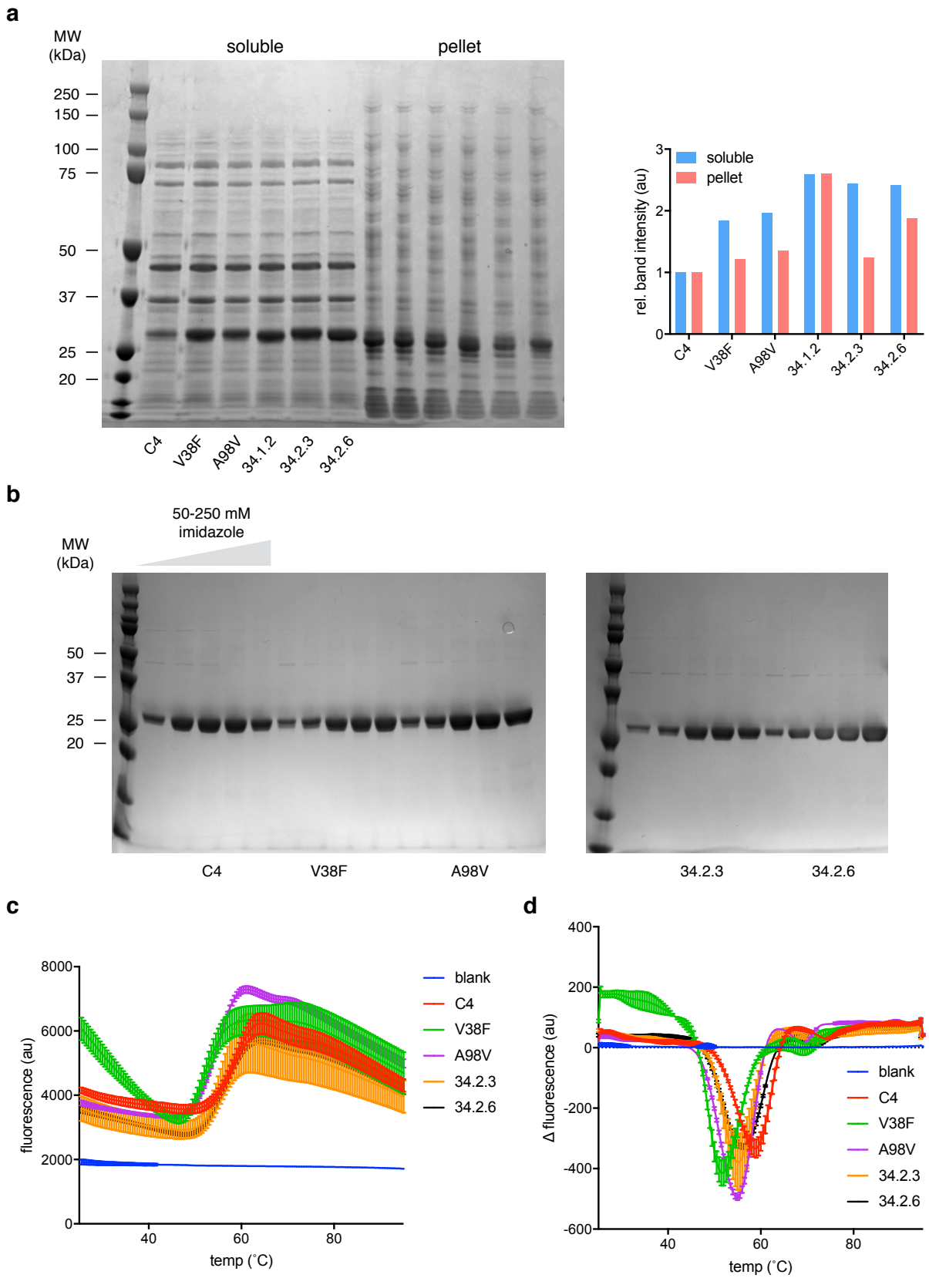
a

	C4 residue number										eT7n residue number					
	1	2	6	30	38	98	126	144	163	223	3	55	97	131	135	159
wt	A	Q	Q	F	V	A	S	S	Y	S	I	E	K	T	A	K
30.1.2	V	Q	Q	F	F	A	S	S	Y	S	M	E	K	T	A	K
30.1.3	A	Q	Q	F	V	T	S	S	Y	S	I	E	K	T	A	K
30.1.4	T	Q	K	F	V	A	S	L	Y	S	I	E	K	T	A	K
30.2.1	A	Q	Q	L	F	A	S	S	Y	S	I	E	R	T	A	K
30.2.2	A	Q	Q	F	V	V	S	S	Y	A	I	E	R	T	A	K
30.2.3	A	Q	Q	F	V	V	P	S	Y	S	I	E	R	T	A	K
30.2.4	A	Q	Q	F	F	A	S	S	Y	S	I	E	R	T	A	K
30.4.4	A	K	Q	F	V	V	S	S	Y	S	I	E	K	T	T	E
30.4.5	A	K	Q	F	V	V	S	S	Y	S	I	D	K	I	A	K
30.4.6	T	Q	Q	F	F	A	S	S	C	S	I	E	K	T	A	K

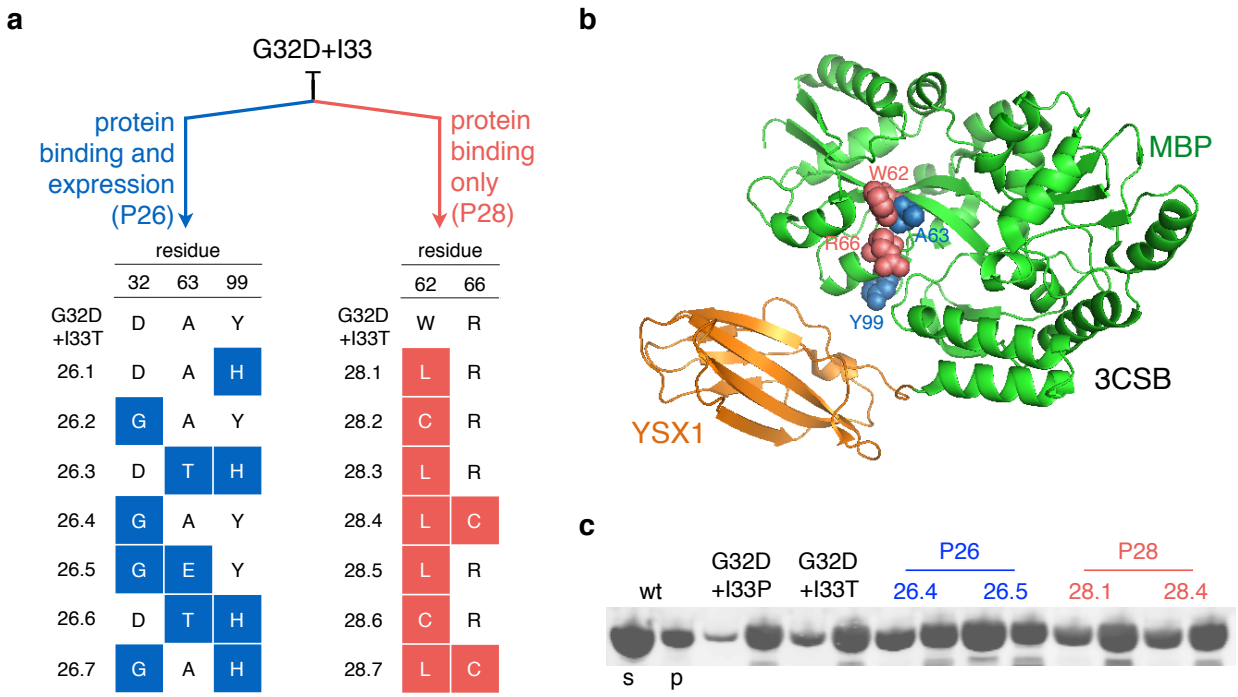
b

	C4 residue number													eT7n residue number					
	1	2	15	24	29	38	51	98	108	123	136	160	163	166	48	64	82	97	131	163
wt	A	Q	P	A	T	V	V	A	R	G	T	I	Y	V	A	A	A	K	T	K
34.1.1	T	Q	P	A	T	F	V	A	R	G	T	I	Y	V	A	V	A	K	I	K
34.1.2	A	K	P	A	P	V	V	V	R	G	I	I	Y	V	A	A	A	R	T	T
34.1.3	A	K	P	A	P	V	V	V	R	G	I	I	Y	V	A	A	A	R	T	T
34.1.4	A	K	P	A	T	V	A	A	L	G	I	V	Y	V	A	A	A	R	I	K
34.1.5	T	Q	P	A	T	F	V	A	R	G	T	I	Y	V	A	V	A	K	I	K
34.1.6	T	Q	P	A	T	F	V	A	R	G	T	I	Y	V	A	V	A	K	I	K
34.1.7	T	Q	P	A	T	F	V	A	R	G	T	I	Y	V	A	A	A	K	I	K
34.1.8	A	K	P	A	P	V	V	V	R	G	I	I	Y	V	A	A	A	R	T	T
34.2.1	A	Q	P	A	T	V	V	V	R	G	T	I	Y	V	S	A	A	K	I	K
34.2.2	A	Q	P	A	T	V	V	V	R	G	T	I	C	I	A	A	S	K	V	K
34.2.3	A	K	P	T	T	V	V	V	R	G	T	I	Y	V	A	A	A	K	I	K
34.2.4	A	Q	P	A	T	V	V	V	R	R	T	I	C	I	A	A	A	K	V	K
34.2.5	A	K	P	T	T	V	V	V	R	G	T	I	Y	V	A	A	A	K	I	K
34.2.6	A	Q	P	A	T	V	V	V	R	G	T	I	C	I	A	A	A	K	I	K
34.2.7	A	Q	P	A	T	V	V	V	R	G	T	I	C	I	A	A	A	K	I	K
34.2.8	T	Q	L	A	T	F	V	A	R	W	T	I	Y	V	A	A	A	K	I	K

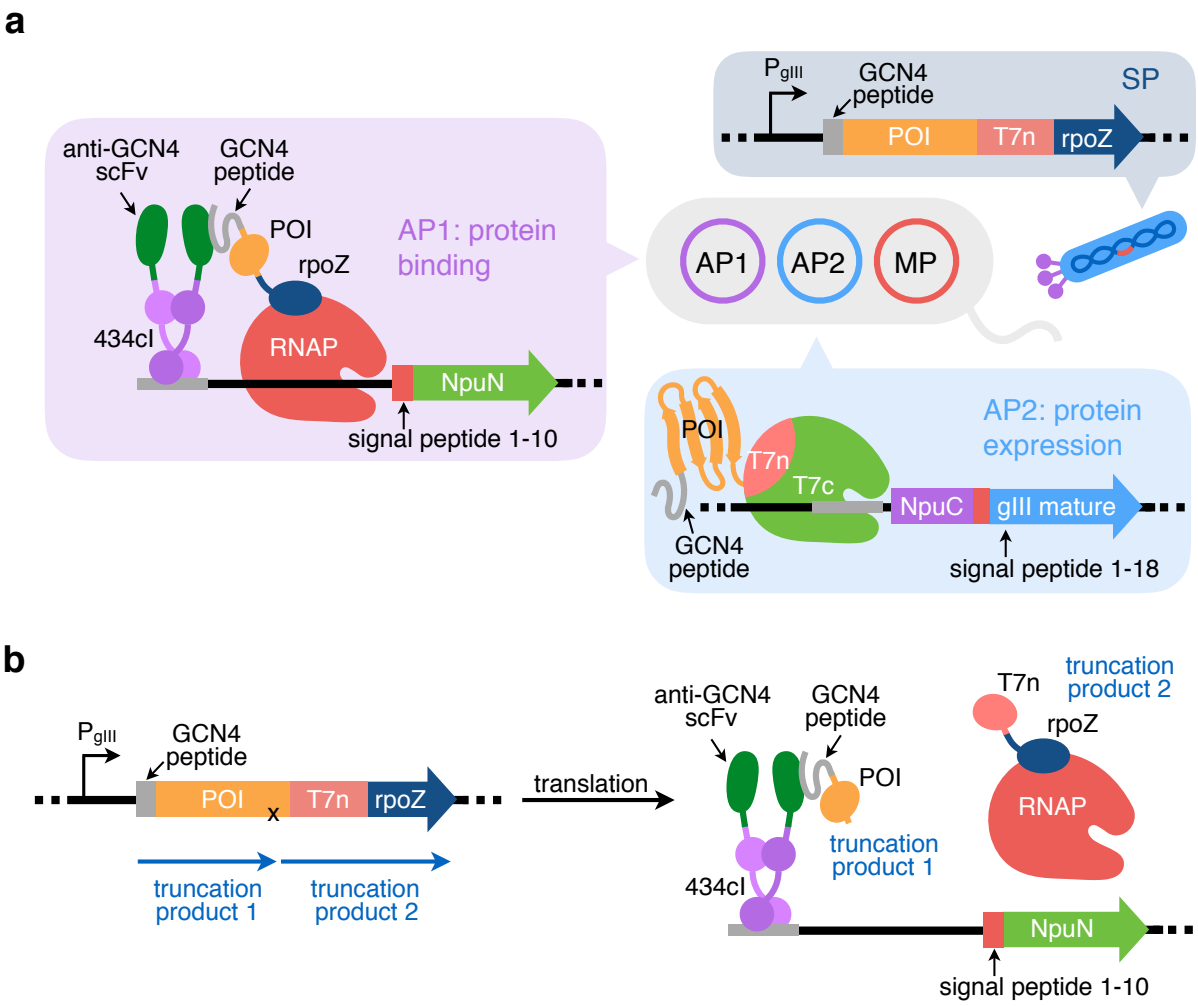
Supplementary Figure 11. Evolution of anti-htt scFv C4. SP isolated from PACE with a moderately stringent (**a**) and (**b**) highly stringent AP for protein expression.


a
MW
(kDa)
soluble
pellet
250
150
100
75
50
37
25
20
C4
V38F
A98V
34.1.2
34.2.3
34.2.6
soluble
pellet
rel. band intensity (au)
3
2
1
0
C4
V38F
A98V
34.1.2
34.2.3
34.2.6
b
50-250 mM
imidazole
MW
(kDa)
50
37
25
20
C4
V38F
A98V
34.2.3
34.2.6
c
fluorescence (au)
8000
6000
4000
2000
0
40
60
80
temp (˚C)
blank
C4
V38F
A98V
34.2.3
34.2.6
d
Δ fluorescence (au)
400
200
0
-200
-400
-600
40
60
80
temp (˚C)
blank
C4
V38F
A98V
34.2.3
34.2.6

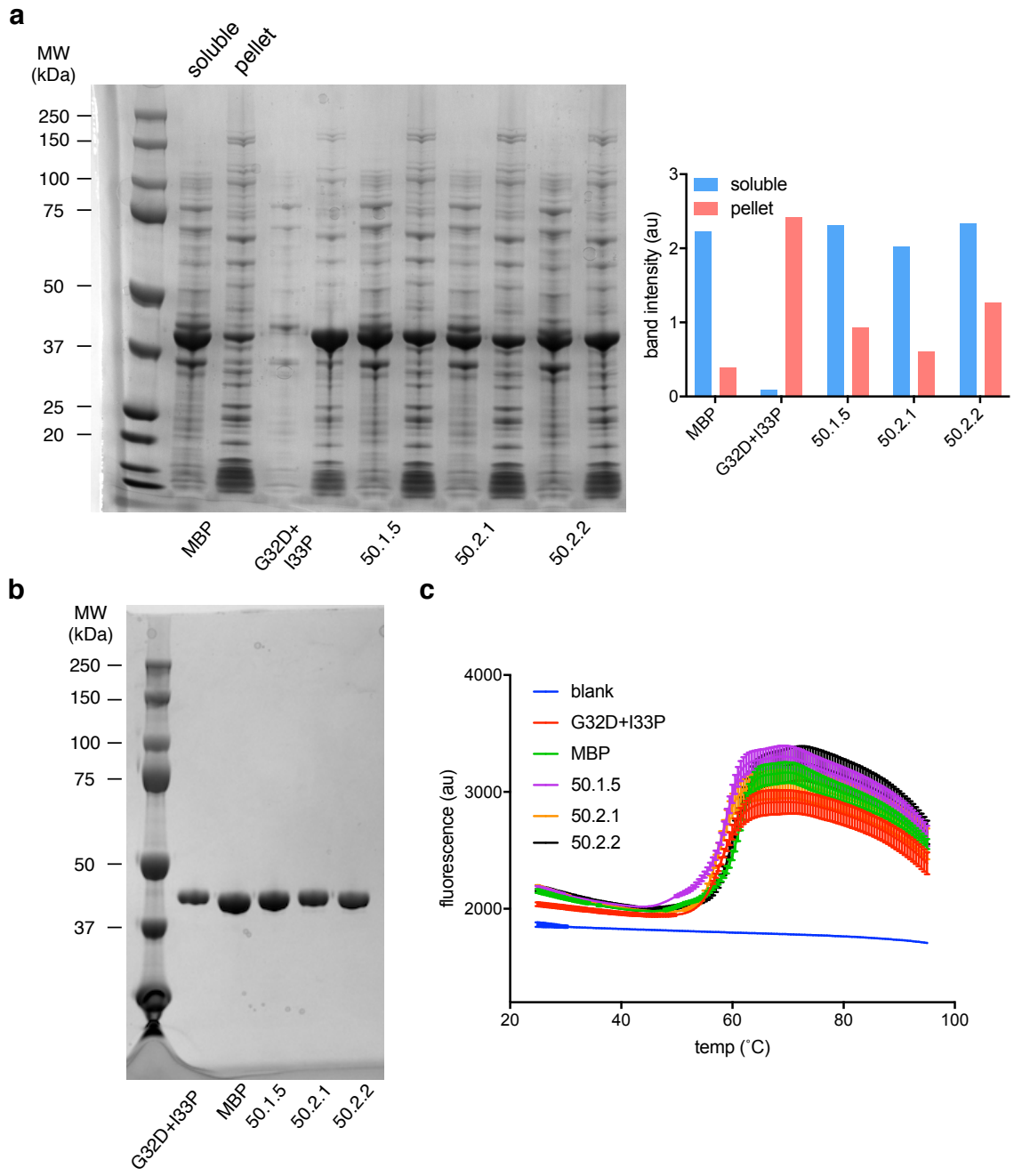
Supplementary Figure 12. Expression, purification, and melting temperature determination of anti-Htt scFv C4 variants. (**a**) Expression of anti-Htt scFv C4 variants in BL21 DE3 cells at 37 °C shown in crude lysates separated by SDS-PAGE. (**b**) SDS-PAGE of purified anti-htt scFv C4 variant fractions expressed in BL21 DE3 cells at 16 °C. (**c**) Melting temperature curves and first derivatives (**d**) for anti-htt scFv C4 variants. Data reflect the mean and s.d. of four technical replicates (four melt curve samples using protein from the same preparation). All experiments in this figure were repeated at least once with similar results.



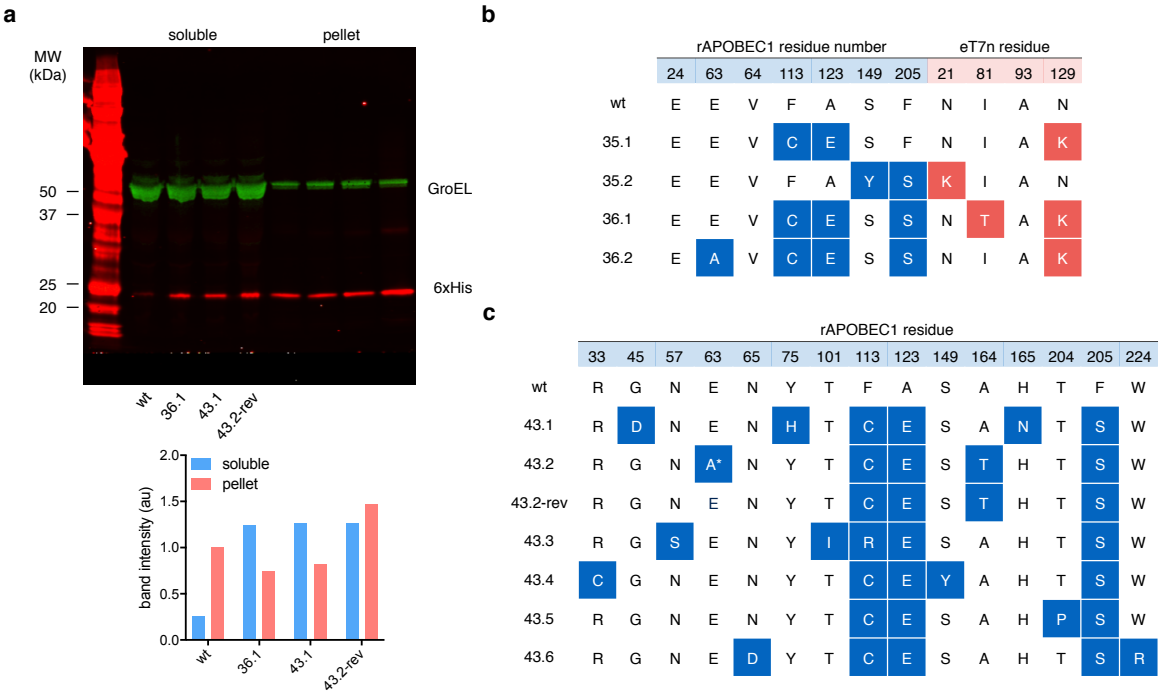
Supplementary Figure 13. Contribution of protein binding selection to improvement of protein activity. (**a**) Parallel PACE experiments with MBP G32D+I33T, where binding to anti-MBP monobody YSX1 was selected in the presence or absence of a simultaneous selection for protein expression. (**b**) Mutations from PACEs P26 and P28, as depicted as blue and red spheres, respectively, map to the YSX1-MBP binding interface. (**c**) Expression of evolved variants in BL21 DE3 cells at 37 °C. This experiment was repeated once with similar results.



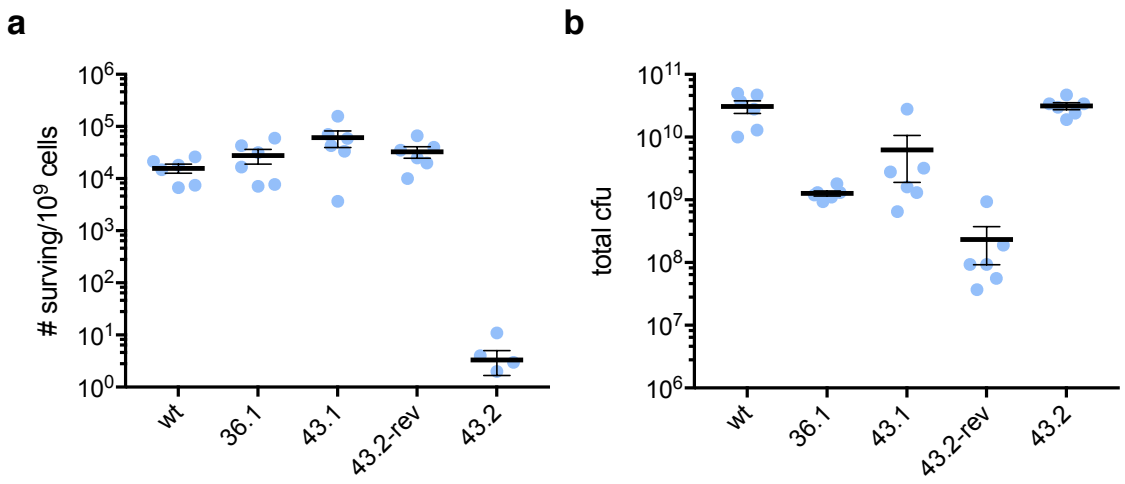
Supplementary Figure 14. Schematic of cheat-resistant activity-independent selection for protein expression. (**a**) The protein of interest (POI) is fused between an affinity tag (in this case, the GCN4 peptide) and T7n–rpoZ, allowing both the protein expression and protein binding selections to operate. As the SP carry the GCN4 peptide linked to rpoZ, they begin PACE with the ability to pass the protein binding selection. (**b**) Any SP that cheat by forming truncations in the POI lose the ability to propagate on the protein binding selection, as the GCN4 peptide is no longer connected to rpoZ. Therefore, these phage are expected to wash out of the lagoon.



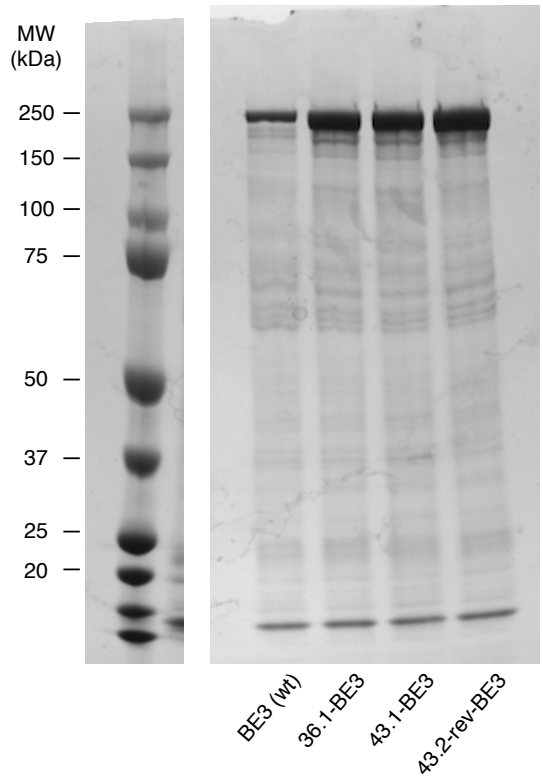
Supplementary Figure 15. Expression, purification, and melting temperature determination of MBP variants. (**a**) Expression of MBP variants in in BL21 DE3 cells at 37 °C shown in crude lysates separated by SDS-PAGE. This expression was performed once. (**b**) SDS-PAGE of purified MBP variants expressed in BL21 DE3 cells at 20 °C. This purification was performed once. (**c**) Melting temperature curves of purified MBP variants. Data reflect the mean and s.d. of four technical replicates (four melt curve samples using protein from the same preparation).



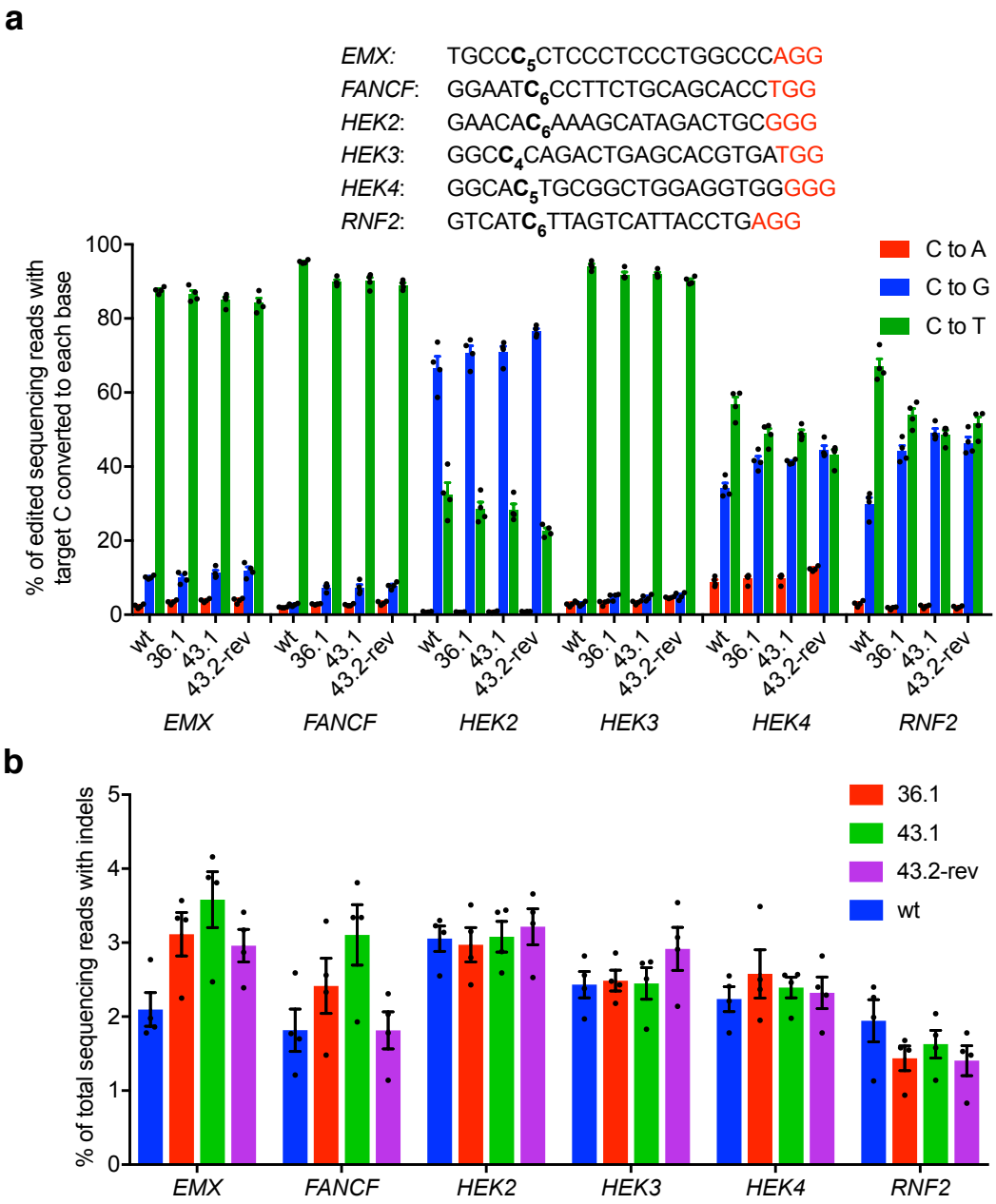
Supplementary Figure 16. Activity-independent protein expression PACE of rAPBOEC1. (**a**) Expression in in BL21 DE3 cells at 37 °C of rAPOBEC1 variants (6xHis) or GroEL (load control). This experiment was repeated once with similar results. (**b**) Consensus mutations from SP isolated from lagoons 1 and 2 of first two PACE experiments. (**c**) Mutations in rAPOBEC1 after 370 h of PACE.



Supplementary Figure 17. Activity of rAPOBEC1 variants in BL21 cells. (**a**) rAPOBEC1 variants from PACE retain deaminase activity in *E coli*, as measured by cells surviving selection on rifampin. (**b**)Total cfu from overnight cultures of BL21 cells after induction of rAPOBEC1 variants. All data in this figure reflect the mean and s.e.m. of six technical replicates (unique clones).



Supplementary Figure 18. SDS-PAGE of purified BE3 variants expressed in BL21 DE3 cells at 16 °C. BE3 variants were isolated after affinity purification using cobalt resin, then buffer-exchanged and concentrated using spin filtration. This purification was repeated once with similar results.



Supplementary Figure 19. Product purity and indel formation resulting from base editing with evolved rAPOBEC1 variants at six genomic loci in HEK293T cells. (**a**) Product distribution of BE3 mediated cytidine editing in HEK293T cells using wild-type rAPOBEC1 and evolved variants at six genomic loci, measured by the percentage of sequencing reads where the target C is converted to A, G, or T. The target C at each site (*EMX-RNF2*) is bolded in the sequences of the six genomic sites shown above the graph. (**b**) Indel formation resulting from editing with BE3 variants in HEK293T cells at the same sites shown in (**a**). All bars in this figure depict the mean and error bars s.e.m. of four biological replicates (experiments performed on different days).

Supplementary Note 1. DNA sequences.

a. MBP. Codons in blue (Gly 32 and Ile 33) were mutated to GAT and CCG in MBP(G32D+I33P). The codon highlighted in red (Tyr 293) was mutated to GAC in Y283D.

AAGATCGAAGAAGGTAAGTTAGTGATTTGGATCAACGGCGATAAAGGCTATAACGG
TTTAGCGGAAGTAGGCAAGAAGTTTGAGAAGGATACCGGAATCAAAGTGACGGTC
GAACACCCGGATAAGTTAGAGGAGAAGTTCCCTCAAGTAGCGGCAACGGGCGAC
GGTCCTGACATCATCTTTTGGGCACACGACCGCTTCGGTGGTTACGCTCAGTCAG
GTTTATTAGCGGAGATTACTCCAGACAAAGCGTTCCAGGACAAGTTATATCCATTTA
CGTGGGACGCAGTGCGTTATAACGGCAAGTTAATCGCATATCCCATTGCGGTAGA
AGCCCTGAGCCTGATTTACAACAAGGACCTGTTACCCAACCCTCCGAAGACGTGG
GAGGAGATTCCGGCTTTAGATAAGGAACTGAAAGCGAAAGGCAAGTCTGCCCTGA
TGTTTAACCTGCAAGAACCCTACTTCACCTGGCCTCTGATCGCGGCGGATGGCGG
TTATGCGTTTAAGTACGAAAACGGCAAATACGATATTAAGGATGTAGGCGTCGATA
ACGCCGGTGCCAAAGCCGGTCTGACCTTTTTAGTTGATCTGATTAAGAATAAGCAC
ATGAACGCAGACACCGACTACAGCATTGCGGAAGCTGCGTTTAATAAAGGTGAGA
CCGCCATGACGATTAACGGACCTTGGGCGTGGTCGAACATTGATACCAGTAAAGT
CAATTACGGCGTTACAGTCCTGCCGACCTTCAAAGGTCAGCCGTCAAAACCGTTC
GTAGGTGTCTTATCCGCCGGTATCAACGCGGCGTCCCCAAATAAAGAGTTGGCTA
AAGAGTTCCTGGAAAACTATCTGCTGACAGACGAAGGACTGGAGGCTGTGAACAA
GGACAAGCCACTGGGCGCTGTTGCGCTGAAAAGTTATGAGGAAGAACTGGCGAAA
GATCCCCGCATCGCGGCGACGATGGAAAACGCCCAAAAAGGCGAAATCATGCCG
AACATTCCGCAGATGTCAGCTTTTTGGTATGCCGTACGCACGGCTGTTATTAACGC
CGCGTCGGGCCGCCAAACCGTTGATGAGGCACTGAAGGACGCGCAGACTCGTAT
CACCAAG

b. gIII-N. NpuN is shown in blue.

ATGAAAAAATTATTATTCGCAATTCCTTTATGTCTCAGCTACGAAACCGAAATCTTG
ACCGTCGAATATGGTCTGCTGCCAATCGGCAAGATTGTTGAAAAACGTATTGAATG
TACGGTCTACTCAGTGGATAACAACGGCAATATCTACACCCAGCCGGTGGCCCAG
TGGCATGACCGTGGTGAACAGGAAGTGTTCGAATATTGTCTGGAAGACGGATCTTT
AATCCGTGCCACAAAGGATCACAAATTTATGACTGTAGATGGTCAGATGCTCCCAA
TCGACGAAATTTTTGAACGCGAATTAGACCTGATGCGCGTGGATAATCTCCCGAAT
TAA

c. gIII-C. NpuC is highlighted in blue. The +1, +2, and +3 exteins are highlighted in red.

ATGATCAAAATTGCCACGCGTAAATATTTAGGCAAACAGAATGTTTATGATATCGGT
GTCGAGCGCGATCATAATTTCGCGCTGAAAAACGGCTTTATCGCCAGCAATTGTTT
TAATGTTGTTCCTTTCTATTCTCACTCCGCTGAAACTGTTGAAAGTTGTTTAGCAAAA
CCCCATACAGAAAATTCATTTACTAACGTCTGGAAAGACGACAAAACTTTAGATCGT
TACGCTAACTATGAGGGCTGTCTGTGGAATGCTACAGGCGTTGTAGTTTGTACTGG
TGACGAAACTCAGTGTTACGGTACATGGGTTCCTATTGGGCTTGCTATCCCTGAAA

ATGAGGGTGGTGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTTCTGAGGGTG
GCGGTACTAAACCTCCTGAGTACGGTGATACACCTATTCCGGGCTATACTTATATC
AACCCTCTCGACGGCACTTATCCGCCTGGTACTGAGCAAAACCCCGCTAATCCTAA
TCCTTCTCTTGAGGAGTCTCAGCCTCTTAATACTTTCATGTTTCAGAATAATAGGTT
CCGAAATAGGCAGGGGGCATTAACTGTTTATACGGGCACTGTTACTCAAGGCACT
GACCCCGTTAAAACTTATTACCAGTACACTCCTGTATCATCAAAAGCCATGTATGAC
GCTTACTGGAACGGTAAATTCAGAGACTGCGCTTTCCATTCTGGCTTTAATGAGGA
TCCATTCGTTTGTGAATATCAAGGCCAATCGTCTGACCTGCCTCAACCTCCTGTCA
ATGCTGGCGGCGGCTCTGGTGGTGGTTCTGGTGGCGGCTCTGAGGGTGGTGGCT
CTGAGGGTGGCGGTTCTGAGGGTGGCGGCTCTGAGGGAGGCGGTTCCGGTGGT
GGCTCTGGTTCCGGTGATTTTGATTATGAAAAGATGGCAAACGCTAATAAGGGGGC
TATGACCGAAAATGCCGATGAAAACGCGCTACAGTCTGACGCTAAAGGCAAACTTG
ATTCTGTCGCTACTGATTACGGTGCTGCTATCGATGGTTTCATTGGTGACGTTTCC
GGCCTTGCTAATGGTAATGGTGCTACTGGTGATTTTGCTGGCTCTAATTCCCAAAT
GGCTCAAGTCGGTGACGGTGATAATTCACCTTTAATGAATAATTTCCGTCAATATTT
ACCTTCCCTCCCTCAATCGGTTGAATGTCGCCCTTTTGTCTTTGGCGCTGGTAAAC
CTTACGAGTTCAGTATCGACTGCGATAAGATCAACCTGTTCCGCGGTGTCTTTGCG
TTTCTTTTATATGTTGCCACCTTTATGTATGTATTTTCTACGTTTGCTAACATACTGC
GTAATAAGGAGTCTTAA

d. T7n.

AACACGATTAACATCGCTAAGAACGACTTCTCTGACATCGAACTGGCTGCTATCCC
GTTCAACACTCTGGCTGACCATTACGGTGAGCGTTTAGCTCGCGAACAGTTGGCC
CTTGAGCATGAGTCTTACGAGATGGGTGAAGCACGCTTCCGCAAGATGTTTGAGC
GTCAACTTAAAGCTGGTGAGGTTGCGGATAACGCTGCCGCCAAGCCTCTCATCAC
TACCCTACTCCCTAAGATGATTGCACGCATCAACGACTGGTTTGAGGAAGTGAAAG
CTAAGCGCGGCAAGCGCCCGACAGCCTTCCAGTTCCTGCAAGAAATCAAGCCGGA
AGCCGTAGCGTACATCACCATTAAGACCACTCTGGCTTGCCTAACCAGTGCTGACA
ATACAACCGTTCAGGCTGTAGCAAGCGCAATCGGTCGGGCCATTGAGGACGAGGC
TCGCTTCGGTCGTATCCGTGACCTTGAAGCTAAGCACTTCAAGAAAAACGTTGAGG
AACAACTCAACAAGCGCGTAGGGCACGTCTACAAG

e. eT7n. Nucleotides in blue represent mutations from T7n (see also **Supplementary Fig. 6a**).

AACACCATTAATATCGCTAAGAACGGCTTCTCTGATATCGAACTGGCTGCTATCCC
GTTCAACACTCTGGCTGACCATTACGGTGAGCGTTTAGCTCGCGAACAGTTGGCC
CTTGAGCATGAGTCTTACGAGATGGATGAAGCACGCTTCCGCAAGATGTTTGAGC
GTCAACTTAAAGCTGGTGAGGTTGCGGATAACACTGCCGCCAATCCTCTCATCACT
ACCCTACTCCTTAAGATGATTGCACGCATCAAGGACTGGTTTGAGGAAGTGAAAGC
TAAGCGCGGCAAGCGCCCGACAGCCTTCCATTTCCTGAAAGAAATCAAGCCGGAA
GCCGTAGCGTACATCACCATTAAGGCCACTCTGGCTTGCCTAACCAGTGCTAACAA
AACAACCGTTCAGGCTGTAGCAAGCGCAATCGGTCGGGCCATTGAGGACGAGGCT
CGCTTCGGTCGTATCCGTTACCTTGAAGCTAAGCACTTCAAGAAAAACGTTGAGGA
ACAACTCAACAAGCGCGTAGGGCACGTCCACAAG

f. GCN4 7P14P peptide.

TTGCAAAGAATGAAACAACTTGAACCGAAGGTTGAAGAATTGCTTCCGAAAAATTAT
CACTTGGAAAATGAGGTTGCCAGATTAAAGAAATTAGTTGGCGAACGC

g. scFv Ωg. Codons in blue (Phe 91 and Ile 188) were mutated to TAT and ACC in m3.

CGCGACATTGTTATGACGCAGTCGCCATCAAGCTTATCAGCGTCAGTGGGAGATC
GCGTTACAATTACATGCCGTTCGAGCACTGGGGCAGTCACAACCAGTAATTACGCT
TCGTGGGTCCAGGAAAAACCGGGTAAGTTGTTCAAGGGTTTGATTGGCGGTACTA
ATAACCGCGCACCGGGCGTCCCTAGCCGTTTTTCGGGGAGTTTGATTGGTGACAA
GGCCACACTTACTATCAGCAGTCTGCAACCAGAGGATTTCGCTACATACTTTTGTG
CATTGTGGTACTCCAACCATTGGGTCTTCGGTCAGGGCACGAAGGTTGAACTTAAA
CGCGGGGGTGGTGGCTCCGGAGGTGGGGGTTCAGGCGGCGGAGGGTCTTCGGG
TGGAGGGAGTGAGGTTAAGCTTCTTGAAAGTGGTGGTGGTCTTGTGCAGCCTGGA
GGCTCGTTAAAGCTGAGCTGCGCTGTGAGTGGTTTCTCGTTGACGGATTATGGGG
TCAATTGGGTACGCCAGGCACCGGGGCGTGGCTTGGAGTGGATTGGCGTCATCT
GGGGCGACGGAATCACTGATTATAACAGTGCCTTGAAGGATCGCTTTATCATCAGC
AAAGACGATTGCGAAAACACTGTCTATTTGCAAATGAGCAAAGTTCGCTCGGATGA
TACGGCGTTATACTACTGTGTCACCGGACTTTTTGACTACTGGGGGCAGGGCACTC
TTGTCACGGTCTCCAGC

h. 29.1.2. Nucleotides in blue represent mutations from scFv Ωg (see also **Supplementary Fig. 7a**).

CGCGACATTGTTATGACGCAGTCGCCATCAAGCTTATCAGCGTCAGTGGGAGATC
GCGTTACAATTACATGCCGTTCGAGCACTGGGGCAGTCACAACCAGTAATTACGCT
TTGTGGATCCAGGAAAAACCGGGTAAGTTGTTCAAGGGTTTGATTGGCGGTACTAA
TAACCGCGCACCGGGCGTCCCTAGCCGTTTTTCGGGGAGTTTGATTGGTGACAAG
GCCACACTTACTATCAGCAGTCTGCAACCAGAGGATTCCGCTACATACTTTTGTGC
ATTGTGGTACTCCAACCATTGGGTCTTCGGTCAGGGAACGAAGGTTGAACTTAAAC
GCGGGGGTGGTGGCTCCGGAGGTGGGGGTTCAGGCGGCGGAGGGTCTTCGGGT
GGAGGGAGTGAGGTTAAGCTTCTTGAAAGTGGTGGTGGTCTTGTGCAGCCTGGAG
GCTCGTTAAAGCTGAGCTGCGCTGTGAGTGGTTTCTCGTTGACGGATTATGGGGT
CAATTGGGTACGCCAGGCACCGGGGCGTGGCTTGGAGTGGATTGGCGTCATCTG
GGGCGACGGAATCACTAATTATAACAGTGCCTTGAAGGATCGCTTTATCATCAGCA
AAGACGATTGCGAAAACACTGTCTATTTGCAAATGAGCAAAGTTCGCTCGGATGAT
ACGGCGTTATACTACTGTGTCACCGGACTTTTTGACTACTGGGGGCAGGGCACTCT
TGTCACGGTCTCCAGC

i. 29.1.5. Nucleotides in blue represent mutations from scFv Ωg (see also **Supplementary Fig. 7a**).

CGCGACATTGTTATGACGCAGTCGCCATCAAGCTTATCAGCGTCAGTGGGAGATC
GCGTTACAATTACATGCCGTTCGAGCACTGGGGCAGTCACAACCAGTAATTACGCT

TCGTGGGTCCAGGAAAAACCGGGTAAGTTGTTCAAGGGTTTGATTGGCGGTACTA
ATAACCGCGCACCGGGCGTCCCTAGCCGTTTTTCGGGGAGTTTGATTGGTGACAA
GGCCACACTTACTATCAGCAGTCTGCAACCAGAGGATTCCGCTACATACTTTTGTG
CATTGTGGTACTCCAACCATTGGGTCTTCAGTCAGGGCACGAAGGTTGAACTTAAA
CGCGGGGGGTGGTGGCTCCGGAGGTGGGGGTTCAGGCGGCGGAGGGTCTTCGGG
TGGAGGGAGTGAGGTTAAGCTTCTTGAAAGTGGTGGTGGTCTTGTGCAGCCTGGA
GGCTCGTTAAAGCTGAGCTGCGCTGTGAGTGGTTTCTCGTTGACGGATTATGGGG
TCAATTGGGTACGCCAGGCACCGGGGCGTGGCTTGGAGTGGATTGGCGTCATCT
GGGGCGACGGAATCACTGATTATAACAGTGCCTTGAAGGATCGCTTTATCATCAGC
AAAGACGATTGCGAAAACACTGTCTATTTGCAAATGAGCAAAGTTCGCTCGGATGA
TACGGCGGTATACTACTGTGTCACCGGACTTTTTGACTACTGGGGGCAGGGCACT
CTTGTCACGGTCTCCAGC

j. m3-noCys. Nucleotides in blue represent mutations from the disulfide-forming cysteines in m3.

CGCGACATTGTTATGACGCAGTCGCCATCAAGCTTATCAGCGTCAGTGGGAGATC
GCGTTACAATTACAAGCCGTTCGAGCACTGGGGCAGTCACAACCAGTAATTACGCT
TCGTGGGTCCAGGAAAAACCGGGTAAGTTGTTCAAGGGTTTGATTGGCGGTACTA
ATAACCGCGCACCGGGCGTCCCTAGCCGTTTTTCGGGGAGTTTGATTGGTGACAA
GGCCACACTTACTATCAGCAGTCTGCAACCAGAGGATTTCGCTACATACTATAGTG
CATTGTGGTACTCCAACCATTGGGTCTTCGGTCAGGGCACGAAGGTTGAACTTAAA
CGCGGGGGGTGGTGGCTCCGGAGGTGGGGGTTCAGGCGGCGGAGGGTCTTCGGG
TGGAGGGAGTGAGGTTAAGCTTCTTGAAAGTGGTGGTGGTCTTGTGCAGCCTGGA
GGCTCGTTAAAGCTGAGCAGCGCTGTGAGTGGTTTCTCGTTGACGGATTATGGGG
TCAATTGGGTACGCCAGGCACCGGGGCGTGGCTTGGAGTGGATTGGCGTCATCT
GGGGCGACGGAACCACTGATTATAACAGTGCCTTGAAGGATCGCTTTATCATCAG
CAAAGACGATTGCGAAAACACTGTCTATTTGCAAATGAGCAAAGTTCGCTCGGATG
ATACGGCGTTATACTACAGTGTCACCGGACTTTTTGACTACTGGGGGCAGGGCACT
CTTGTCACGGTCTCCAGC

k. 29.1.5-noCys. Nucleotides in blue represent mutations from the disulfide-forming cysteines in 29.1.5.

CGCGACATTGTTATGACGCAGTCGCCATCAAGCTTATCAGCGTCAGTGGGAGATC
GCGTTACAATTACAAGCCGTTCGAGCACTGGGGCAGTCACAACCAGTAATTACGCT
TCGTGGGTCCAGGAAAAACCGGGTAAGTTGTTCAAGGGTTTGATTGGCGGTACTA
ATAACCGCGCACCGGGCGTCCCTAGCCGTTTTTCGGGGAGTTTGATTGGTGACAA
GGCCACACTTACTATCAGCAGTCTGCAACCAGAGGATTCCGCTACATACTTTAGTG
CATTGTGGTACTCCAACCATTGGGTCTTCAGTCAGGGCACGAAGGTTGAACTTAAA
CGCGGGGGGTGGTGGCTCCGGAGGTGGGGGTTCAGGCGGCGGAGGGTCTTCGGG
TGGAGGGAGTGAGGTTAAGCTTCTTGAAAGTGGTGGTGGTCTTGTGCAGCCTGGA
GGCTCGTTAAAGCTGAGCAGCGCTGTGAGTGGTTTCTCGTTGACGGATTATGGGG
TCAATTGGGTACGCCAGGCACCGGGGCGTGGCTTGGAGTGGATTGGCGTCATCT
GGGGCGACGGAATCACTGATTATAACAGTGCCTTGAAGGATCGCTTTATCATCAGC
AAAGACGATTGCGAAAACACTGTCTATTTGCAAATGAGCAAAGTTCGCTCGGATGA

TACGGCGGTATACTACAGTGTCACCGGACTTTTTGACTACTGGGGGCAGGGCACT
CTTGTCACGGTCTCCAGC

l. 58.1.1. Nucleotides in blue represent mutations from m3-noCys (see also **Supplementary Fig. 15a**).

CGCGACATTGTTATGACGCAGTCGCCATCAAGCTTATCAGCGTCAGTGGGAGATC
GCGTTACAATTACAAGCCGTTCGAGCACTGGGGCAGTCACAACCAGTACTTACATT
TCGTGGGTCCAGGAAAAACCGGGTAAGTTGTTCAAGGGTTTGATTGGCGGTACTA
ATAACCGCGCACCGGGCGTCCCTAGCCGTTTTTCGGGGAGTTTGATTGGTGACAA
GGCCACACTTACTATCAGCAGTCTGCAACCAGAGGATTTCGCTACATACTATAGTG
CATTGTGGTACTCCAACCATTGGGTCTTCGGTCAGGGCACGAAGGTTGAACTTAAA
CGCGGGGGTGGTGGCTCCGGAGGTGGGGGTTCAGGCGGCGGAGGGTCTTCGGG
TGGAGGGAGTGAGGTTAAGCTTCTTGAAAGTGGTGGTGGTCTTGTGCAACCTGGA
GGCTCGTTAACGCTGAGCAGCGCTGTGAGTGGTTTCTCGTTGACGGATTATGGGG
TCAATTGGGTACGCCAGGCACCGGGGCGTGGCTTGGAGTGGATTGGCGTCATCT
GGGGCGACGGAACCACTGATTATAACAGTGCCTTGAAGGATCGCTTTATCATCAG
CAAAGACGATTGCGAAAACACTGTCTATTTGCAAATGAGCAAAGTTCGCTCGGATG
ATACGGCGTTATACTACAGTGTCACCGGACTTCTTGACTACTGGGGGCAGGGCAC
TCTTGTCAAGGTCTCCAGC

m. 58.1.2. Nucleotides in blue represent mutations from m3-noCys (see also **Supplementary Fig. 15a**).

CGCGACATTGTTATGACGCAGTCGCCATCAAGCTTATCAGCGTCAGTGGGAGATC
GCGTTACAATTACAAGCCGTTCGAGCACTGGGGCAGTCACAACCAGTAATTACACT
GCGTGGGTCCAGGAAAAACCGGGTAAGTTGTTCAAGGGTTTGATTGGCGGTACTA
ATAACCGCGCACCGGGCGTCCCTAGCCGTTTTTCGGGGAGTTTGATTGGTGACAA
GGCCACACTTACTATCAGCAGTCTGCAACCAGAGGATTTCGCTACATACTATAGTG
CATTGTGGTACTCCAACCATTGGGTCTTCGGTCAGGGCACGAAGGTTGAACTTAAA
CGCGGGGGTGGTGGCTCCGGAGGTGGGGGTTCAGGCGGCGGAGGGTCTTCGGG
TGGAGGGAGTGAAGTTAAGCTTCTTGAAAGTGGTGGTGGTCTTGTGCAACCTGGA
GGCTCGTTAACGCTGAGCAGCGCTGTGAGTGGTTTCTCGTTGACGGATTATGGGG
TCAATTGGGTACGCCAGGCACCGGGGCGTGGCTTGGAGTGGATTGGCGTCATCT
GGGGCGACGGAACCACTGATTATAACAGTGCCTTGAAGGATCGCTTTATCATCAG
CAAAGACGATTGCGAAAACACTGTCTATTTGCAAATGAGCAAAGTTCGCTCGGATG
ATACGGCGTTCTACTACAGTGTCACCGGACTTCTTGACTACTGGGGGCAGGGCAC
TCTTGTCAAGGTCTCCAGC

n. 58.2.8. Nucleotides in blue represent mutations from m3-noCys (see also **Supplementary Fig. 15a**).

CGCGACATTGTTATGACGCAGTCGCCATCAAGCTTATCAGCGTCAGTGGGAGATC
GCGTTACAATTACAAGCCGTTCGAGCACTGGGGCAGTCACAACCAGTAATTACGCT
TCGTGGGTCCAGAAAAAACCGGGTAAGTTGTTCAAGGGTTTGATTGGCGGTACTAA
TAACCGCGCACCGGGCGTCCCTAGCCGTTTTTCGGGGAGTTTGATTGGTGACAAG

GCCACACTTACTATCAGCAGTCTGCAACCAGAGGATTTCGCTACATACTATAGTGC
ATTGTGGTACTCCAACCATTGGGTCTTCGGTCAGGGCACGAAGGTTGAACTTAAAC
GCGGGGGTGGTGGCTCCGGAGGTGGGGGTTCAGGCGGCGGAGGGTCTTCGGGT
GGAGGGAGTGAGGTTAAGCTTCTTGAAAGTGGTGGTGGTCTTGTGCAGCCTGGAG
GCTCGTTAAAGCTGAGCAGCGCTGTGAGTGGTTTCTCGTTGACGGATTATGGGGT
CAATTGGGTACGCCAGGCACCGGGGCGTGGCTTGGAGTGGATTGGCGTCATCTG
GGGCGACGGAACCACTGATTATAACAGTGCCTTGAAGGATCGCTTTATCATCAGCA
AAGACGATTGCGAAAACACTGTCTATTTGCAAATGAGCAAAGTTCGCTCGGATGAT
ACGGCGTTATACTACAGTGTCACCGGACTTCTTGACTACTGGGGGCAGGGCACTC
TTGTCACGGTCTCCAGC

o. 58.3.1. Nucleotides in blue represent mutations from 29.1.5-noCys (see also **Supplementary Fig. 15b**).

CGCGACATTGTTATGACGCAGTCGCCATCAAGCTTATCAGCGTCAGTGGGAGATC
GCGTTACAATTACAAGCCGTTCGAGCACTGGGGCAGTCACAACCAGTAATTACGTT
TCGTGGGTCCAGGAAAAACCGGGTAAGTTGTTCAAGGGTTTGATTGGCGGTACTA
ATAACCGCGCACCGGGCGTCCCTAGCCGTTTTTCGGGGAGTTTGATTGGTGACAA
GGCCACACTTACTATCAGCAGTCTGCAACCAGAGGATTCCGCTACATACTTTAGTG
CATTGTGGTACTCCAACCATTGGGTCTTCAGTCAGGGCACGAAGGTTGAACTTAAA
CGCGGGGGTGGTGGCTCCGGAGGTGGGGGTTCAGGCGGCGGAGGGTCTTCGGG
TGGAGGGAGTGAGGTTAAGCTTCTTGAAAGTGGTGGTGGTCTTGTGCAGCCTGGA
GGCTCGTTAAAGCTGAGCAGCGCTGTGAGTGGTTTCTCGTTGACGGATTATGGGG
TCAATTGGGTACGCCAGGCACCGGGGCGTGGCTTGGAGTGGATTGGCGTCATCT
GGGGCGACGGAACCACTGATTATAACAGTGCCTTGAAGGATCGCTTTATCATCAG
CAAAGACGATTGCGAAAACACTGTCTATTTGCAAATGAGCAAAGTTCGCTCGGATG
ATACGGCGGTATACTACAGTGTCACCGGACTTCTTGACTACTGGGGGCAGGGCAC
TCTTGTCACGGTCTCCAGC

p. 58.4.2. Nucleotides in blue represent mutations from 29.1.5-noCys (see also **Supplementary Fig. 15b**).

CGCGACATTGTTATGACGCAGTCGCCATCAAGCTTATCAGCGTCAGTGGGAGATC
GCGTTACAATTACAAGCCGTTCGAGCACTGGGGCAGTCACAACCAGTAATTACGCT
TCGTGGATCCAGGAAAAACCGGGTAAGTTGTTCAAGGGTTTGATTGGCGGTACTAA
TAACCGCGCACCGGGCGTCCCTAGCCGTTTTTCGGGGAGTTTGATTGGTGACAAG
GCCACACTTACTATCAGCAGTCTGCAACCAGAGGATTCCGCTACATACTTTAGTGC
ATTGTGGTACTCCAACCATTGGGTCTTCAGTCAGGGCACGAAGGTTGAACTTAAAC
GCGGGGGTGGTGGCTCCGGAGGTGGGGGTTCAGGCGGCGGAGGGTCTTCGGGT
GGAGGGAGTGAGGTTAAGCTTCTTGAAAGTGGTGGTGGTCTTGTGCAGCCTGGAG
GCTCGTTAAAGCTGAGCAGCGCTGTGAGTGGTTTCTCGTTGACGGATTATGGGGT
CAATTGGGTACGCCAGGCACCGGGGCGTGGCTTGGAGTGGATTGGCGTCATCTG
GGGCGACGGAACCACTGATTATAACAGTGCCTTGAAGGATCGCTTTATCATCAGCA
AAGACGATTGCGAAAACACTGTCTATTTGCAAATGAGCAAAGTTCGCTCGGATGAT
ACGGCGGTATACTACAGTGTCACCGGACTTTTTGACTACTGGGGGCAGGGCACTC
TTGTCACGGTCTCCAGC

q. 58.4.4. Nucleotides in blue represent mutations from 29.1.5-noCys (see also **Supplementary Fig. 15b**).

CGCGACATTGTTATGACGCAGTCGCCATCAAGCTTATCAGCGTCAGTGGGAGATC
GCGTTACAATTACAAGCCGTTCGAGCACTGGGGCAGTCACAACCAGTAATTATGCT
TCGTGGGTCCAGGAAAAACCGGGTAAGTTGTTCAAGGGTTTGATTGGCGGTACTA
ATAACCGCGCACCGGGCGTCCCTAGCCGTTTTTCGGGGAGTTTGATTGGGGACAA
GGCCACACTTACTATCAGCAGTCTGCAACCAGAGGATTCCGCTACATACTTTAGTG
CATTGTGGTACTCCAACCATTGGGTCTTCAGTCAGGGCACGAAGGTTGAACTTAAA
CGCGGGGGGTGGTGGCTCCGGAGGTGGGGGTTCAGGCGGCGGAGGGTCTTCGGG
TGGAGGGAGTGAGGTTAAGCTTCTTGAAAGTGGTGGTGGTCTTGTGCAGCCTGGA
GGCTCGTTAAAGCTGAGCAGCGCTGTGAGTGGTTTCTCGTTGACGGATTATGGGG
TCAATTGGGTACGCCAGGCACCGGGGCGTGGCTTGGAGTGGATTGGCGTCATCT
GGGGCGACGGAAGCACTGATTATAACAGTGCCTTGAAGGATCGCTTTATCATCAG
CAAAGACGATTGCGAAAACACTGTCTATTTGCAAATGAGCAAAGTTCGCTCGGATG
ATACGGCGGTATACTACAGTGTCACCGGACTTTTTGAGTACTGGGGGCAGGGCAC
TCTTGTCACGGTCTCCAGC

r. htt peptide.

ATGGCGACCTTAGAAAAACTTATGAAGGCATTTGAGTCGCTGAAATCTTTC

s. scFv C4.

GCCCAAGTTCAGCTTCAAGAGAGCGGGGGAGGTTTAGTTCAGCCCGGTGGCAGTT
TACGCTTATCTTGTGCCGCTTCTGGGTTCACCTTTTCATCATACTCAATGTCCTGGG
TCCGCCAAGCCCCCGGAAAGGGACTGGAGTGGGTCGCAGTAATCTCCTATGATGG
AAGCAACAAATATTACGCCGACAGTGTGAAGGGACGTTTTACCATTAGTCGCGACA
ATTCAAAAAATACTTTGTACTTGCAAATGAACTCATTACGTGCGGAAGACACAGCG
GTATATTATTGCGCGCGTGACCGTTACTTTGACTTATGGGGACGCGGCACATTGGT
GACTGTCTCAAGCGGTGGTGGCGGATCAGGGGGGGGCGGTTCAGGTGGAGGTG
GGTCACAATCAGCTCTGACACAACCTGCGTCCGTAAGCGGTTCACCAGGACAGAG
TATTACAATCTCTTGTACAGGTACTAGCTCCGATATTGGCGCTTACAACTATGTCTC
CTGGTATCAACAGTATCCGGGCAAGGCCCCAAAATTATTGATCTACGATGTTAGCA
ATCGTCCAAGTGGAATCAGTAACCGCTTCAGCGGCAGCAAATCCGGAGACACCGC
TTCATTGACTATCAGCGGGCTGCAAGCCGAGGATGAAGCTGATTATTACTGTTCGT
CCTTCGCAAATAGCGGGCCTTTGTTTGGAGGTGGAACCAAAGTTACCGTGTTGGG
C

t. 34.1.2. Nucleotides in blue represent mutations from scFv C4 (see also **Supplementary Fig. 9**).

GCCAAAGTTCAGCTTCAAGAGAGCGGGGGAGGTTTAGTTCAGCCCGGTGGCAGTT
TACGCTTATCTTGTGCCGCTTCTGGGTTCCCCTTTTCATCATACTCAATGTCCTGGG
TCCGCCAAGCCCCCGGAAAGGGACTGGAGTGGGTCGCAGTAATCTCCTATGATGG

AAGCAACAAATATTACGCCGACAGTGTGAAGGGACGTTTTACCATTAGTCGCGACA
ATTCAAAAAATACTTTGTACTTGCAAATGAACTCATTACGTGCGGAAGACACAGCG
GTATATTATTGCGTGCGTGACCGTTACTTTGACTTATGGGGACGCGGCACATTGGT
GACTGTCTCAAGCGGTGGTGGCGGATCAGGGGGGGGCGGTTCAGGTGGAGGTG
GGTCACAATCAGCTCTGATACAACCTGCGTCCGTAAGCGGTTCACCAGGACAGAG
TATTACAATCTCTTGTACAGGTACTAGCTCCGATATTGGCGCTTACAACTATGTCTC
CTGGTATCAACAGTATCCGGGCAAGGCCCCAAAATTATTGATCTACGATGTTAGCA
ATCGTCCAAGTGGAATCAGTAACCGCTTCAGCGGCAGCAAATCCGGAGACACCGC
TTCATTGACTATCAGCGGGCTGCAAGCCGAGGATGAAGCTGATTATTACTGTTCGT
CCTTCGCAAATAGCGGGCCTTTGTTTGGAGGTGGAACCAAAGTTACCGTGTTGGG
C

u. 34.2.3. Nucleotides in blue represent mutations from scFv C4 (see also **Supplementary Fig. 9**).

GCCAAAGTTCAGCTTCAAGAGAGCGGGGGAGGTTTAGTTCAGCCCGGTGGCAGTT
TACGCTTATCTTGTACCGCTTCTGGGTTCACCTTTTCATCATACTCAATGTCCTGGG
TCCGCCAAGCCCCCGGAAAGGGACTGGAGTGGGTCGCAGTAATCTCCTATGATGG
AAGCAACAAATATTACGCCGACAGTGTGAAGGGACGTTTTACCATTAGTCGCGACA
ATTCAAAAAATACTTTGTACTTGCAAATGAACTCATTACGTGCGGAAGACACAGCG
GTATATTATTGCGTGCGTGACCGTTACTTTGACTTATGGGGACGCGGCACATTGGT
GACTGTCTCAAGCGGTGGTGGCGGATCAGGGGGGGGCGGTTCAGGTGGAGGTG
GGTCACAATCAGCTCTGACACAACCTGCGTCCGTAAGCGGTTCACCAGGACAGAG
TATTACAATCTCTTGTACAGGTACTAGCTCCGATATTGGCGCTTACAACTATGTCTC
CTGGTATCAACAGTATCCGGGCAAGGCCCCAAAATTATTGATCTACGATGTTAGCA
ATCGTCCAAGTGGAATCAGTAACCGCTTCAGCGGCAGCAAATCCGGAGACACCGC
TTCATTGACTATCAGCGGGCTGCAAGCCGAGGATGAAGCTGATTATTACTGTTCGT
CCTTCGCAAATAGCGGGCCTTTGTTTGGAGGTGGAACCAAAGTTACCGTGTTGGG
C

v. 34.2.6. Nucleotides in blue represent mutations from scFv C4 (see also **Supplementary Fig. 9**).

GCCCAAGTTCAGCTTCAAGAGAGCGGGGGAGGTTTAGTTCAGCCCGGTGGCAGTT
TACGCTTATCTTGTGCCGCTTCTGGGTTCACCTTTTCATCATACTCAATGTCCTGGG
TCCGCCAAGCCCCCGGAAAGGGACTGGAGTGGGTCGCAGTAATCTCCTATGATGG
AAGCAACAAATATTACGCCGACAGTGTGAAGGGACGTTTTACCATTAGTCGCGACA
ATTCAAAAAATACTTTGTACTTGCAAATGAACTCATTACGTGCGGAAGACACAGCG
GTATATTATTGCGTGCGTGACCGTTACTTTGACTTATGGGGACGCGGCACATTGGT
AACTGTCTCAAGCGGTGGTGGCGGATCAGGGGGGGGCGGTTCAGGTGGAGGTG
GGTCACAATCAGCTCTGACACAACCTGCGTCCGTAAGCGGTTCACCAGGACAGAG
TATTACAATCTCTTGTACAGGTACTAGCTCCGATATTGGCGCTTGCAACTATATCTC
CTGGTATCAACAGTATCCGGGCAAGGCCCCAAAATTATTGATCTACGATGTTAGCA
ATCGTCCAAGTGGAATCAGTAACCGCTTCAGCGGCAGCAAATCCGGAGACACCGC
TTCATTGACTATCAGCGGGCTGCAAGCCGAGGATGAAGCTGATTATTACTGTTCGT

CCTTCGCAAATAGCGGGCCTTTGTTTGGAGGTGGAACCAAAGTTACCGTGTTGGG
C

w. 50.1.5. Nucleotides in blue represent mutations from MBP G32D+I33P.

AAGATCGAAGAAGGTAAGTTAGTGATTTGGATCAACGGCGATAAAGGCTATAACGG
TTTAGCGGAAGTAGGCAAGAAGTTTGAGAAGGATACCGGTACGAAAGTGACGGTC
GAACACCCGGATAAGTTAGAGGAGAAGTTCCCTCAAGTAGCGGCAACGGGCGAC
GGTCCTGACATCATCTTTTGGGCACACGACCGCTTCGGTGGTTACGCTCAGTCAG
GTTTATTAGCGGAGATTACTCCAGACAAAGCGTTCCAGGACAAGTTATATCCATTTA
CGTGGGACGCAGTGCGTTATAACGGCAAGTTAATCGCATATCCCATTGCGGTAGA
AGCCCTGAGCCTGATTTACAACAAGGACCTGTTACCCAACCCTCCGAAGACGTGG
GAGGAGATTCCGGCTTTAGATAAGGAACTGAAAGCGAAAGGCAAGTCTGCCCTGA
TGTTTAACCTGCAAGAACCCTACTTCACCTGGCCTCTGATCGCGGCGGATGGCGG
TTATGCGTTTAAGTACGAAAACGGCAAATACGATATTAAGGATGTAGGCGTCGATA
ACGCCGGTGCCAAAGCCGGTCTGACCTTTTTAGTTGATCTGATTAAGAATAAGCAC
ATGAACGCAGACACCGACTACAGCATTGCGGAAGCTGCGTTTAATAAAGGTGAGA
CCGCCATGACGATTAACGGACCTTGGGCGTGGTCGAACATTGATACCAGTAAAGT
CAATTACGGCGTTACAGTCCTGCCGACCTTCAAAGGTCAGCCGTCAAAACCGTAC
GTAGGTGTCTTATCCGCCGGTATCAACGCGGCGTCCCCAAATAAAGAGTTGGCTA
AAGAGTTCCTGGAAAACTATCTGCTGACAGACGAAGGACTGGAGGCTGTGAACAA
GGACAAGCCACTGGGCGCTGTTGCGCTGAAAAGTTATGAGGAAGAACTGGCGAAA
GATCCCCGCATCGCGGCGACGATGGAAAACGCCCAAAAAGGCGAAATCATGCCG
AACATTCCGCAGATGTCAGCTTTTTGGTATGCCGTACGCACGGCTGTTATTAACGC
CGCGTCGGGCCGCCAAACCGTTGATGAGGCACTGAAGGACGCGCAGACTCGTAT
CACCAAG

x. 50.2.1. Nucleotides in blue represent mutations from MBP G32D+I33P.

AAGATCGAAGAAGGTAAGTTAGTGATTTGGATCAACGGCGATAAAGGCTATAACGG
TTTAGCGGAAGTAGTCAAGAAGTTTGAGAAGGATACCGATTCGAAAGTGACGGTC
GAACACCCGGATAAGTTAGAGGAGAAGTTCCCTCAAGTAGCGGCAACGGGCGAC
GGTCCTGACATCATCTTTTGGGCACACGACCGCTTCGGTGGTTACGCTCAGTCAG
GTTTATTAGCGGAGATTACTCCAGACAAAGCGTTCCAGGACAAGTTATATCCATTTA
CGTGGGACGCAGTGCGTTATAACGGCAAGTTAATCGCATATCCCATTGCGGTAGA
AGCCCTGAGCCTGATTTACAACAAGGACCTGTTACCCAACCCTCCGAAGACGTGG
GAGGAGATTCCGGCTTTAGATAAGGAACTGAAAGCGAAAGGCAAGTCTGCCCTGA
TGTTTAACCTGCAAGAACCCTACTTCACCTGGCCTCTGATCGCGGCGGATGGCGG
TTATGCGTTTAAGTACGAAAACGGCAAATACGATATTAAGGATGTAGGCGTCGATA
ACGCCGGTGCCAAAGCCGGTCTGACCTTTTTAGTTGATCTGATTAAGAATAAGCAC
ATGAACGCAGACACCGACTACAGCATTGCGGAAGCTGCGTTTAATAAAGGTGAGA
CCGCCATGACGATTAACGGACCTTGGGCGTGGTCGAACATTGATACCAGTAAAGT
CAATTACGGCGTTACAGTCCTGCCGACCTTCAAAGGTCAGCCGTCAAAACCGTTC
GTAGGTGTCTTATCCGCCGGTATCAACGCGGCGTCCCCAAATAAAGAGTTGGCTA
AAGAGTTCCTGGAAAACTATCTGCTGACAGACGAAGGACTGGAGGCTGTGAACAA
GGACAAGCCACTGGGCGCTGTTGCGCTGAAAAGTTATGAGGAAGAACTGGCGAAA

GATCCCCGCATCGCGGCGACGATGGAAAACGCCCAAAAAGGCGAAATCATGCCG
AACATTCCGCAGATGTCAGCTTTTTGGTATGCCGTACGCACGGCTGTTATTAACGC
CGCGTCGGGCCGCCAAACCGTTGATGAGGCACTGAAGGACGCGCAGACTCGTAT
CACCAAG

y. 50.2.2. Nucleotides in blue represent mutations from MBP G32D+I33P.

AAGATCGAAGAAGGTAAGTTAGTGATTTGGATCAACGGCGATAAAGGCTATAACGG
TTTAGCGGAAGTAGGCAAGAAGTTTGAGAAGGATACCGGTCTGAAAGTGACGGTC
GAACACCCGGATAAGTTAGAGGAGAAGTTCCCTCAAGTAGCGGCAACGGGCGAC
GGTCCTGACATCATCTTTTGGGCACACGACCGCTTCGGTGGTTACGCTCAGTCAG
GTTTATTAGCGGAGATTACTCCAGACAAAGCGTTCCAGGACAAGTTATATCCATTTA
CGTGGGACGCAGTGCGTTATAACGGCAAGTTAATCGCATATCCCATTGCGGTAGA
AGCCCTGAGCCTGATTTACAACAAGGACCTGTTACCCAACCCTCCGAAGACGTGG
GAGGAGATTCCGGCTTTAGATAAGGAACTGAAAGCGAAAGGCAAGTCTGCCCTGA
TGTTTAACCTGCAAGAACCCTACTTCACCTGGCCTCTGATCGCGGCGGATGGCGG
TTATGCGTTTAAGTACGAAAACGGCAAATACGATATTAAGGATGTAGGCGTCGATA
ACGCCGGTGCCAAAGCCGGTCTGACCTTTTTAGTTGATCTGATTAAGAATAAGCAC
ATGAACGCAGACACCGACTACAGCATTGCGGAAGCTGCGTTTAATAAAGGTGAGA
CCGCCATGACGATTAACGGACCTTGGGCGTGGTCGAACATTGATACCAGTAAAGT
CAATTACGGCGTTACAGTCCTGCCGACCTTCAAAGGTCAGCCGTCAAAACCGTTC
GTAGGTGTCTTATCCGCCGGTATCAACGCGGCGTCCCCAAATAAAGAGTTGGCTA
AAGAGTTCCTGGAAAACTATCTGCTGACAGACGAAGGACTGGAGGCTGTGAACAA
GGACAAGCCACTGGGCGCTGTTGCGCTGAAAAGTTATGAGGAAGAACTGGCGAAA
GATCCCCGCATCGCGGCGACGATGGAAAACGCCCAAAAAGGCGAAATCATGCCG
AACATTCCGCAGATGTCAGCTTTTTGGTATGCCGTACGCACGGCTGTTATTAACGC
CGCGTCGGGCCGCCAAACCGTTGATGAGGCACTGAAGGACGCGCAGACTCGTAT
CACCAAG

z. rAPOBEC1.

TCTTCTGAAACCGGTCCGGTTGCGGTTGACCCGACCCTGCGTCGTCGTATCGAAC
CGCACGAATTCGAAGTTTTCTTCGACCCGCGTGAACTGCGTAAAGAAACCTGCCTG
CTGTACGAAATCAACTGGGGTGGTCGTCACTCTATCTGGCGTCACACCTCTCAGAA
CACCAACAAACACGTTGAAGTTAACTTCATCGAAAAATTCACCACCGAACGTTACTT
CTGCCCGAACACCCGTTGCTCTATCACCTGGTTCCTGTCTTGGTCTCCGTGCGGT
GAATGCTCTCGTGCGATCACCGAATTCCTGTCTCGTTACCCGCACGTTACCCTGTT
CATCTACATCGCGCGTCTGTACCACCACGCGGACCCGCGTAACCGTCAGGGTCTG
CGTGACCTGATCTCTTCTGGTGTTACCATCCAGATCATGACCGAACAGGAATCTGG
TTACTGCTGGCGTAACTTCGTTAACTACTCTCCGTCTAACGAAGCGCACTGGCCGC
GTTACCCGCACCTGTGGGTTCGTCTGTACGTTCTGGAACTGTACTGCATCATCCTG
GGTCTGCCGCCGTGCCTGAACATCCTGCGTCGTAAACAGCCGCAGCTGACCTTCT
TCACCATCGCGCTGCAGTCTTGCCACTACCAGCGTCTGCCGCCGCACATCCTGTG
GGCGACCGGTCTGAAA

aa. 36.1. Nucleotides in blue represent mutations from rAPOBEC1 (see also **Supplementary Fig. 13b**).

TCTTCTGAAACCGGTCCGGTTGCGGTTGACCCGACCCTGCGTCGTCGTATCGAAC
CGCACGAATTCGAAGTTTTCTTCGACCCGCGTGAACTGCGTAAAGAAACCTGCCTG
CTGTACGAAATCAACTGGGGTGGTCGTCACTCTATCTGGCGTCACACCTCTCAGAA
CACCAACAAACACGTTGAAGTTAACTTCATCGAAAAATTCACCACCGAACGTTACTT
CTGCCCGAACACCCGTTGCTCTATCACCTGGTTCCTGTCTTGGTCTCCGTGCGGT
GAATGCTCTCGTGCGATCACCGAATTCCTGTCTCGTTACCCGCACGTTACCCTGTG
CATCTACATCGCGCGTCTGTACCACCACGAGGACCCGCGTAACCGTCAGGGTCTG
CGTGACCTGATCTCTTCTGGTGTTACCATCCAGATCATGACCGAACAGGAATCTGG
TTACTGCTGGCGTAACTTCGTTAACTACTCTCCGTCTAACGAAGCGCACTGGCCGC
GTTACCCGCACCTGTGGGTTCGTCTGTACGTTCTGGAACTGTACTGCATCATCCTG
GGTCTGCCGCCGTGCCTGAACATCCTGCGTCGTAAACAGCCGCAGCTGACCTCCT
TCACCATCGCGCTGCAGTCTTGCCACTACCAGCGTCTGCCGCCGCACATCCTGTG
GGCGACCGGTCTGAAA

ab. 43.1. Nucleotides in blue represent mutations from rAPOBEC1 (see also **Supplementary Fig. 13c**).

TCTTCTGAAACCGGTCCGGTTGCGGTTGACCCGACCCTGCGTCGTCGTATCGAAC
CGCACGAATTCGAAGTTTTCTTCGACCCGCGTGAACTGCGTAAAGAAACCTGCCTG
CTGTACGAAATCAACTGGGATGGTCGTCACTCTATCTGGCGTCACACCTCTCAGAA
CACCAACAAACACGTTGAAGTTAACTTCATCGAAAAATTCACCACCGAACGTCACTT
CTGCCCTAACACCCGTTGCTCTATCACCTGGTTCCTGTCTTGGTCTCCGTGCGGTG
AATGCTCTCGTGCGATCACCGAATTCCTGTCTCGTTACCCGCACGTTACCCTGTGC
ATCTACATCGCGCGTCTGTACCACCACGAGGACCCGCGTAACCGTCAGGGTCTGC
GTGACCTGATCTCTTCTGGTGTTACCATCCAGATCATGACCGAACAGGAATCTGGT
TACTGCTGGCGTAACTTCGTTAACTACTCTCCGTCTAACGAAGCGAACTGGCCGCG
TTACCCGCACCTGTGGGTTCGTCTGTACGTTCTGGAACTGTACTGCATCATCCTGG
GTCTGCCGCCGTGCCTGAACATCCTGCGTCGTAAACAGCCGCAGCTGACCTCCTT
CACCATCGCGCTGCAGTCTTGCCACTACCAGCGTCTGCCGCCGCACATCCTGTGG
GCGCCCGGTCTGAAA

ac. 43.2. Nucleotides in blue and red represent mutations from rAPOBEC1 (see also **Supplementary Fig. 13c**). The nucleotide in red was reverted to an A in 43.2-rev.

TCTTCTGAAACCGGTCCGGTTGCGGTTGACCCGACCCTGCGTCGTCGTATCGAAC
CGCACGAATTCGAAGTTTTCTTCGACCCGCGTGAACTGCGTAAAGAAACCTGTCTG
CTGTACGAAATCAACTGGGGTGGTCGTCACTCTATCTGGCGTCACACCTCTCAGAA
CACCAACAAACACGTTGCAGTTAACTTCATCGAAAAATTCACCACCGAACGTTACTT
CTGCCCGAACACCCGTTGCTCTATCACCTGGTTCCTGTCTTGGTCTCCGTGCGGT
GAATGCTCTCGTGCGATCACCGAATTCCTGTCTCGTTACCCGCACGTTACCCTGTG
CATCTACATCGCGCGTCTGTACCACCACGAGGACCCGCGTAACCGTCAGGGTCTG
CGTGACCTGATCTCTTCTGGTGTTACCATCCAGATCATGACCGAACAGGAATCTGG
TTACTGCTGGCGTAACTTCGTTAACTACTCTCCGTCTAACGAAACGCACTGGCCGC

GTTACCCGCACCTGTGGGTTCGTCTGTACGTTCTGGAACTGTACTGCATCATCCTG
GGTCTGCCGCCGTGCCTGAACATCCTGCGTCGTAAACAGCCGCAGCTGACCTCCT
TCACCATCGCGCTGCAGTCTTGCCACTACCAGCGTCTGCCGCCGCACATCCTGTG
GGCGACCGGTCTGAAA

ad. rAPOBEC1 mammalian codon optimized.

AGCTCAGAGACTGGCCCAGTGGCTGTGGACCCCACATTGAGACGGCGGATCGAG
CCCCATGAGTTTGAGGTATTCTTCGATCCGAGAGAGCTCCGCAAGGAGACCTGCC
TGCTTTACGAAATTAATTGGGGGGGCCGGCACTCCATTTGGCGACATACATCACAG
AACACTAACAAGCACGTCGAAGTCAACTTCATCGAGAAGTTCACGACAGAAAGATA
TTTCTGTCCGAACACAAGGTGCAGCATTACCTGGTTTCTCAGCTGGAGCCCATGCG
GCGAATGTAGTAGGGCCATCACTGAATTCCTGTCAAGGTATCCCCACGTCACTCTG
TTTATTTACATCGCAAGGCTGTACCACCACGCTGACCCCCGCAATCGACAAGGCCT
GCGGGATTTGATCTCTTCAGGTGTGACTATCCAAATTATGACTGAGCAGGAGTCAG
GATACTGCTGGAGAAACTTTGTGAATTATAGCCCGAGTAATGAAGCCCACTGGCCT
AGGTATCCCCATCTGTGGGTACGACTGTACGTTCTTGAACTGTACTGCATCATACT
GGGCCTGCCTCCTTGTCTCAACATTCTGAGAAGGAAGCAGCCACAGCTGACATTC
TTTACCATCGCTCTTCAGTCTTGTCATTACCAGCGACTGCCCCCACACATTCTCTG
GGCCACCGGGTTGAAA

ae. 36.1 mammalian codon optimized.

AGCTCAGAGACTGGCCCAGTGGCTGTGGACCCCACATTGAGACGGCGGATCGAG
CCCCATGAGTTTGAGGTATTCTTCGATCCGAGAGAGCTCCGCAAGGAGACCTGCC
TGCTTTACGAAATTAATTGGGGGGGCCGGCACTCCATTTGGCGACATACATCACAG
AACACTAACAAGCACGTCGAAGTCAACTTCATCGAGAAGTTCACGACAGAAAGATA
TTTCTGTCCGAACACAAGGTGCAGCATTACCTGGTTTCTCAGCTGGAGCCCATGCG
GCGAATGTAGTAGGGCCATCACTGAATTCCTGTCAAGGTATCCCCACGTCACTCTG
TGTATTTACATCGCAAGGCTGTACCACCACGAAGACCCCCGCAATCGACAAGGCC
TGCGGGATTTGATCTCTTCAGGTGTGACTATCCAAATTATGACTGAGCAGGAGTCA
GGATACTGCTGGAGAAACTTTGTGAATTATAGCCCGAGTAATGAAGCCCACTGGCC
TAGGTATCCCCATCTGTGGGTACGACTGTACGTTCTTGAACTGTACTGCATCATAC
TGGGCCTGCCTCCTTGTCTCAACATTCTGAGAAGGAAGCAGCCACAGCTGACATC
CTTTACCATCGCTCTTCAGTCTTGTCATTACCAGCGACTGCCCCCACACATTCTCTG
GGCCACCGGGTTGAAA

af. 43.1 mammalian codon optimized.

AGCTCAGAGACTGGCCCAGTGGCTGTGGACCCCACATTGAGACGGCGGATCGAG
CCCCATGAGTTTGAGGTATTCTTCGATCCGAGAGAGCTCCGCAAGGAGACCTGCC
TGCTTTACGAAATTAATTGGGATGGCCGGCACTCCATTTGGCGACATACATCACAG
AACACTAACAAGCACGTCGAAGTCAACTTCATCGAGAAGTTCACGACAGAAAGACA
CTTCTGTCCGAACACAAGGTGCAGCATTACCTGGTTTCTCAGCTGGAGCCCATGC
GGCGAATGTAGTAGGGCCATCACTGAATTCCTGTCAAGGTATCCCCACGTCACTCT
GTGTATTTACATCGCAAGGCTGTACCACCACGAAGACCCCCGCAATCGACAAGGC

CTGCGGGATTTGATCTCTTCAGGTGTGACTATCCAAATTATGACTGAGCAGGAGTC
AGGATACTGCTGGAGAAACTTTGTGAATTATAGCCCGAGTAATGAAGCCAACTGGC
CTAGGTATCCCCATCTGTGGGTACGACTGTACGTTCTTGAACTGTACTGCATCATA
CTGGGCCTGCCTCCTTGTCTCAACATTCTGAGAAGGAAGCAGCCACAGCTGACAT
CCTTTACCATCGCTCTTCAGTCTTGTCATTACCAGCGACTGCCCCCACACATTCTCT
GGGCCACCGGGTTGAAA

ag. 43.2-rev mammalian codon optimized.

AGCTCAGAGACTGGCCCAGTGGCTGTGGACCCCACATTGAGACGGCGGATCGAG
CCCCATGAGTTTGAGGTATTCTTCGATCCGAGAGAGCTCCGCAAGGAGACCTGCC
TGCTTTACGAAATTAATTGGGGGGGCCGGCACTCCATTTGGCGACATACATCACAG
AACACTAACAAGCACGTCGAAGTCAACTTCATCGAGAAGTTCACGACAGAAAGATA
TTTCTGTCCGAACACAAGGTGCAGCATTACCTGGTTTCTCAGCTGGAGCCCATGCG
GCGAATGTAGTAGGGCCATCACTGAATTCCTGTCAAGGTATCCCCACGTCACTCTG
TGTATTTACATCGCAAGGCTGTACCACCACGAAGACCCCCGCAATCGACAAGGCC
TGCGGGATTTGATCTCTTCAGGTGTGACTATCCAAATTATGACTGAGCAGGAGTCA
GGATACTGCTGGAGAAACTTTGTGAATTATAGCCCGAGTAATGAAACCCACTGGCC
TAGGTATCCCCATCTGTGGGTACGACTGTACGTTCTTGAACTGTACTGCATCATAC
TGGGCCTGCCTCCTTGTCTCAACATTCTGAGAAGGAAGCAGCCACAGCTGACATC
CTTTACCATCGCTCTTCAGTCTTGTCATTACCAGCGACTGCCCCCACACATTCTCTG
GGCCACCGGGTTGAAA

Supplementary References

1. Ringquist, S. *et al*. Translation initiation in *Escherichia coli*: sequences within the ribosome-binding site. *Mol Microbiol* **6**, 1219-1229 (1992).
2. Carlson, J. C., Badran, A. H., Guggiana-Nilo, D. A. & Liu, D. R. Negative selection and stringency modulation in phage-assisted continuous evolution. *Nature chemical biology* **10**, 216-222 (2014).
3. Davis, J.H., Rubin, A.J., & Sauer, R.T. Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res* **39**, 1131-1141 (2011).
4. Rees, H.A. *et al.* Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nature Communications* **8**, 15790 doi: 10.1038/ncomms15790 (2017).
5. Gaudelli, N.M. *et al.* Programmable base editing of A• T to G• C in genomic DNA without DNA cleavage. *Nature* **551**, 464 (2017).
6. Komor, A.C. *et al.* Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420 (2016).
7. Badran, A. H. & Liu, D. R. Development of potent in vivo mutagenesis plasmids with broad mutational spectra. *Nature communications* **6**, 8425 (2015).
8. Badran, A. H. *et al.* Continuous evolution of Bacillus thuringiensis toxins overcomes insect resistance. *Nature* **533**, 58-63 (2016).