Computation-guided optimization of split protein systems

Supplementary Information

Taylor B. Dolberg[1,2,9], Anthony T. Meger[3,4,9], Jonathan D. Boucher[2,5], William K. Corcoran[2,5], Elizabeth E. Schauer[1,2], Alexis N. Prybutok[1,2], Srivatsan Raman[3,4,8,*], Joshua N. Leonard[1,2,5,6,7,*]

[1]Department of Chemical and Biological Engineering, Northwestern University, Evanston, Illinois 60208, United States

[2]Center for Synthetic Biology, Northwestern University, Evanston, Illinois 60208, United States

[3]Department of Biochemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

[4]Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

[5]Interdisciplinary Biological Sciences Graduate Program, Northwestern University, Evanston, Illinois 60208, United States

[6]Chemistry of Life Processes Institute, Northwestern University, Evanston, Illinois 60208, United States

[7]Member, Robert H. Lurie Comprehensive Cancer Center, Northwestern University, Evanston, Illinois 60208, United States

[8]Department of Bacteriology, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

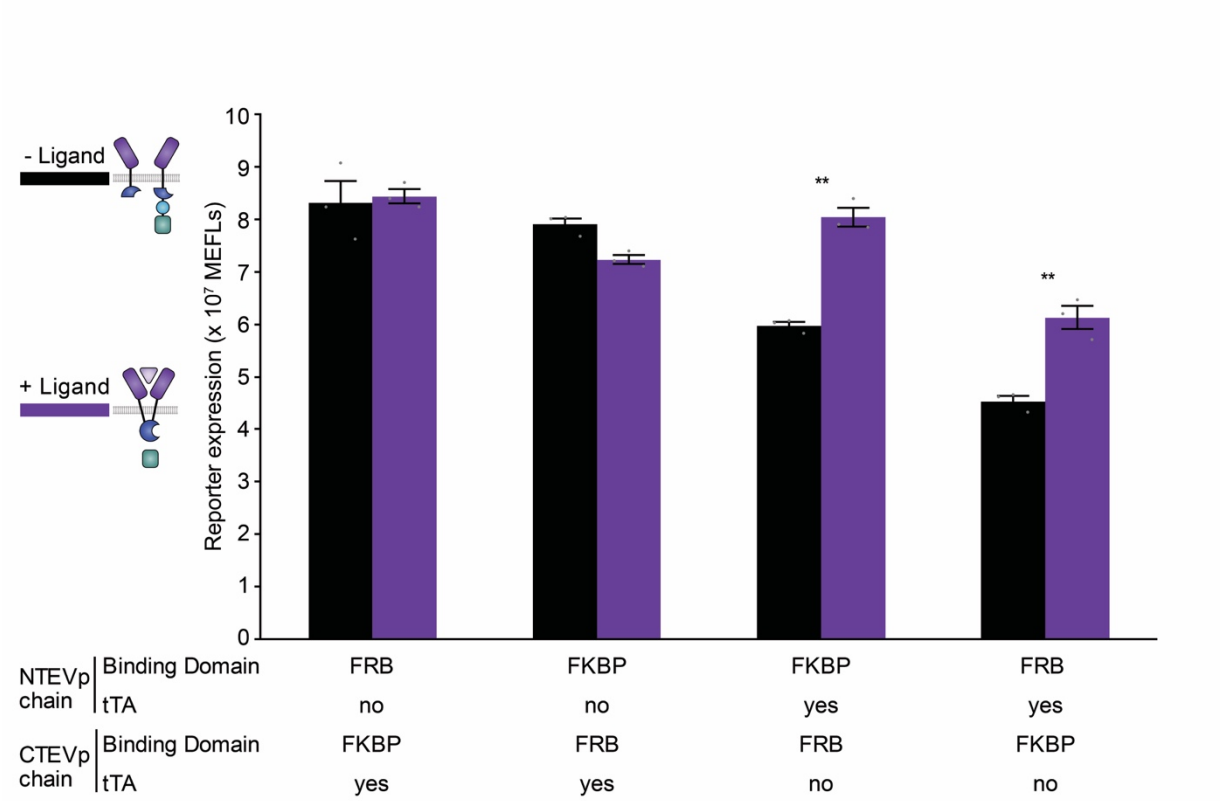
[9]These authors contributed equally to this work

Contact Information

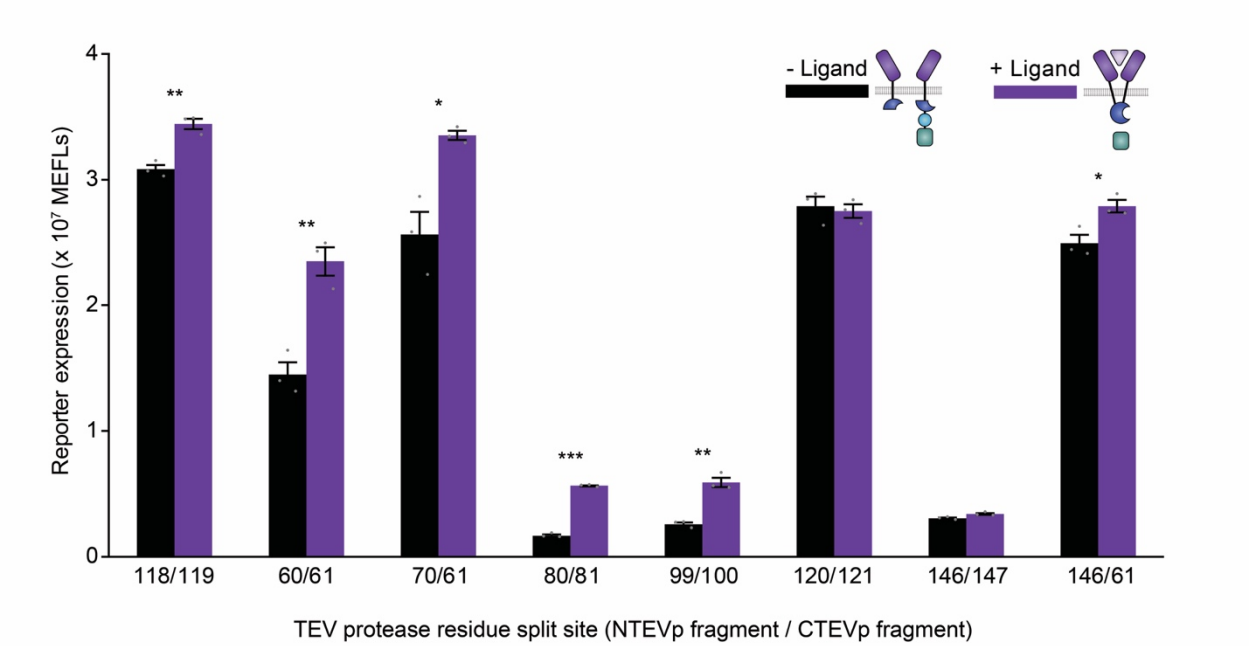
*Corresponding authors: Srivatsan Raman (sraman4@wisc.edu) and Joshua N. Leonard (j-leonard@northwestern.edu)

Contents:

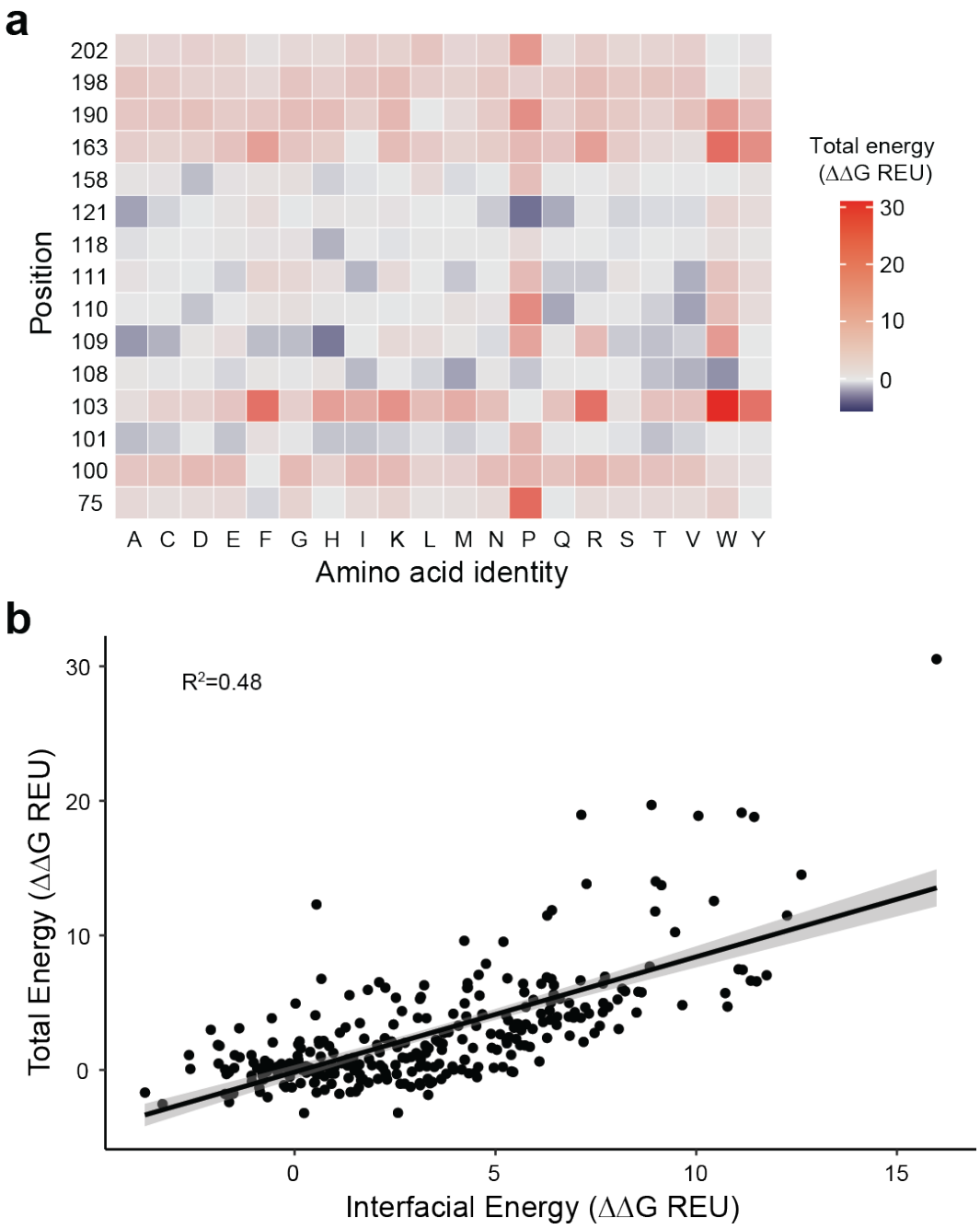
SUPPLEMENTARY FIGURES



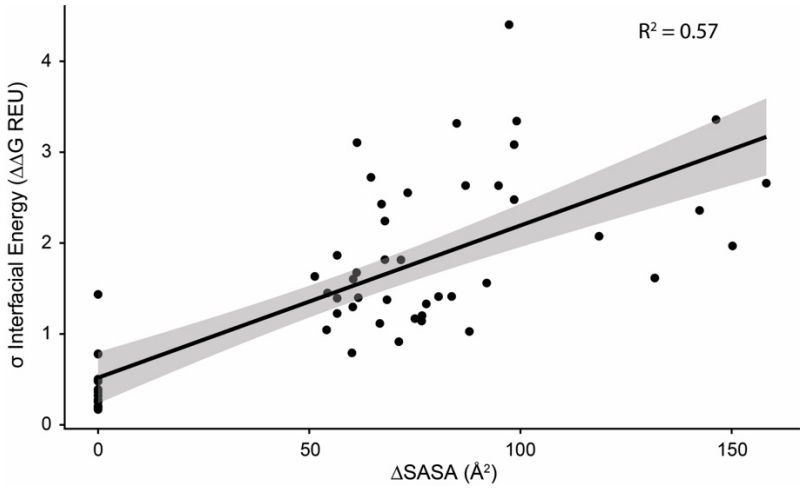
Supplementary Fig. 1 Comparative evaluation of architectures for the membrane-bound model system. Experimental analysis of various architectures of canonical 118/119 split TEVp MESA system, swapping the rapamycin binding domains (FRB, FKBP), TEVp fragment (NTEVp, CTEVp), and which chain contains the tTA. Reporter expression driven by reconstituted TEVp cleavage of tTA measured with flow cytometry. All four architectures were observed to have very high ligand independent signaling, suggesting that the 118/119 split of TEVp reconstitutes too easily without the addition of ligand, regardless of rapamycin binding domain and tTA chain architecture. The second architecture was carried forward for further analysis. Error bars depict S.E.M. Two-tailed Student's t-test (**$p \leq 0.01$). Experiments were performed in biological triplicate. Results are representative of two independent biological experiments.



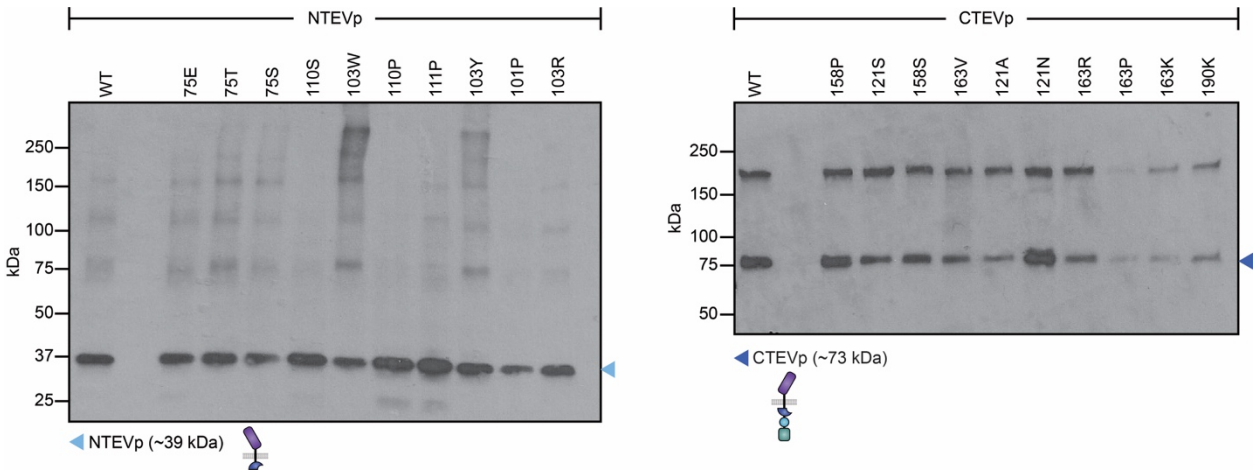
Supplementary Fig. 2 Evaluation of alternative TEVp split sites. Experimental analysis of TEVp splits other than 118/119 canonical split. The FKBP binding domain is on the NTEVp chain, and the FRB binding domain and tTA are on the CTEVp chain. Reporter expression driven by reconstituted TEVp cleavage of tTA measured with flow cytometry. Alternate splits were observed to have very high ligand independent signaling, or greatly diminished ligand independent and dependent signaling. This suggests that splits other than the canonical split do not have the optimal reconstitution propensity for the membrane bound environment. Error bars depict S.E.M. Two-tailed Student's t-test (*$p \leq 0.05$, **$p \leq 0.01$, *** $p \leq 0.001$). Experiments were performed in biological triplicate. Results are representative of two independent biological experiments.



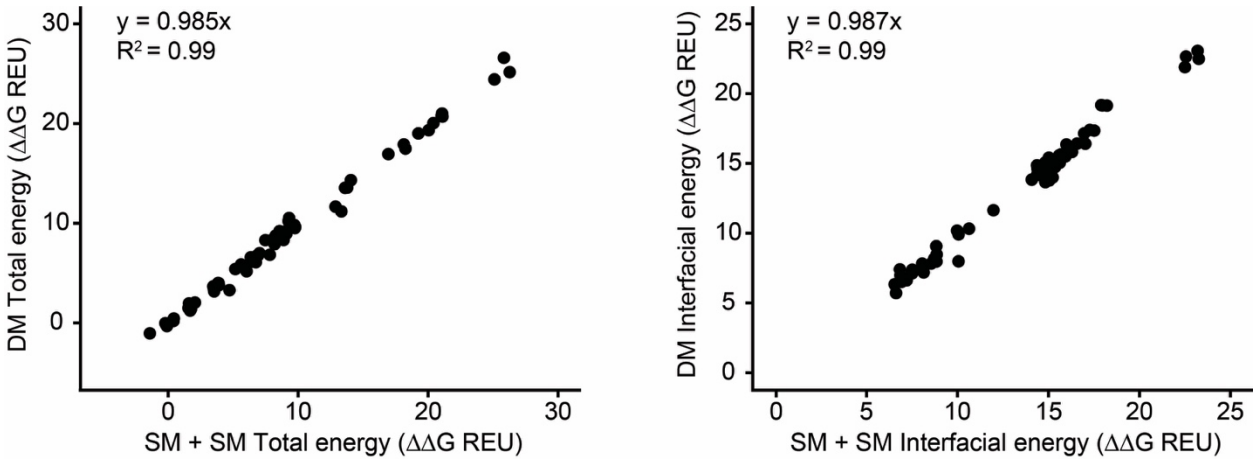
Supplementary Fig. 3 Mutational scanning of high ΔSASA residues using Rosetta. **a,** Energetic perturbations to total stability for all amino acid identities at each position are relative to the WT sequence. Change in interfacial energy for all amino acid identities at each position are shown in **Fig. 2b**. b, Energy landscape of all possible single amino acid substitutions at high ΔSASA residues. Error bands depict 95% C.I. of linear fit.



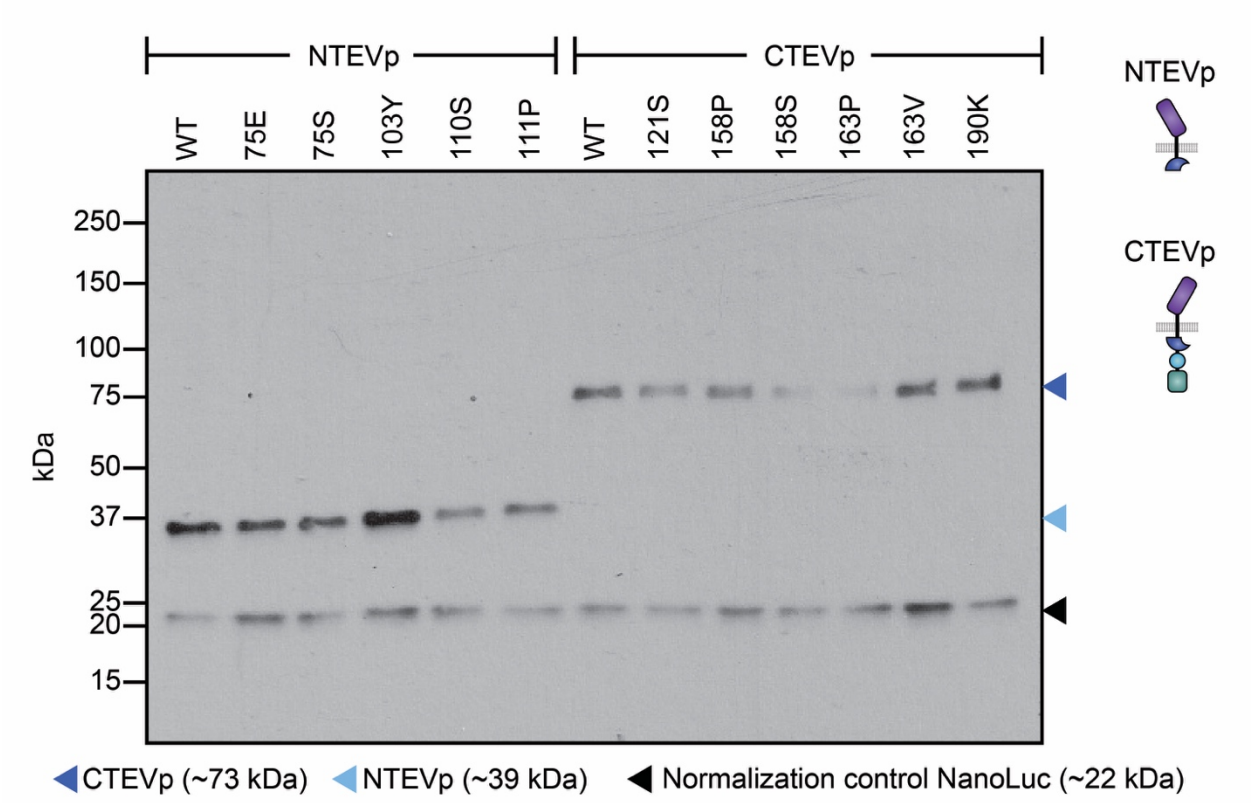
Supplementary Fig. 4 Reconstitution tuning potential of computed interface energies for positions with varying magnitudes of ΔSASA. Standard deviations (σ) were computed from Rosetta mutational scanning of 18 randomly selected positions with ΔSASA=0 and all positions with ΔSASA > 50 Å^2. Error bands depict 95% C.I. of linear fit.



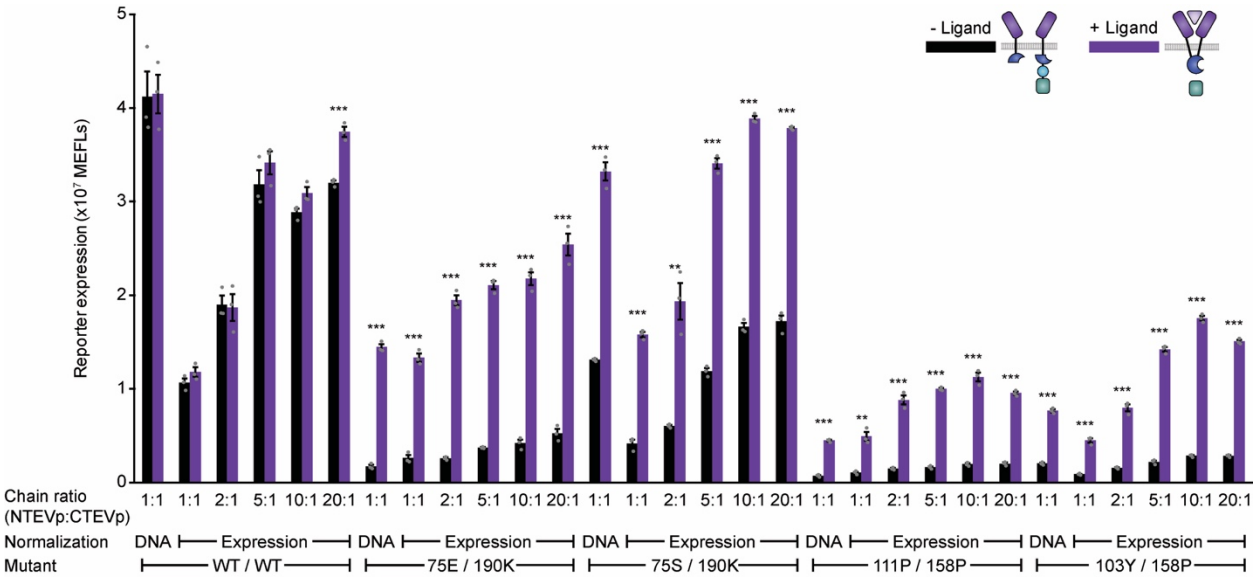
Supplementary Fig. 5 Western blot analysis of mutant split TEVp expression levels. Cells were transfected with single mutant chains which contain an N-terminal 3x-FLAG tag, the FKBP binding domain is on the NTEVp chain, and the FRB binding domain and tTA are on the CTEVp chain. Equal masses of protein (1 µg/lane) were loaded into each lane to investigate differences in chain expression level.



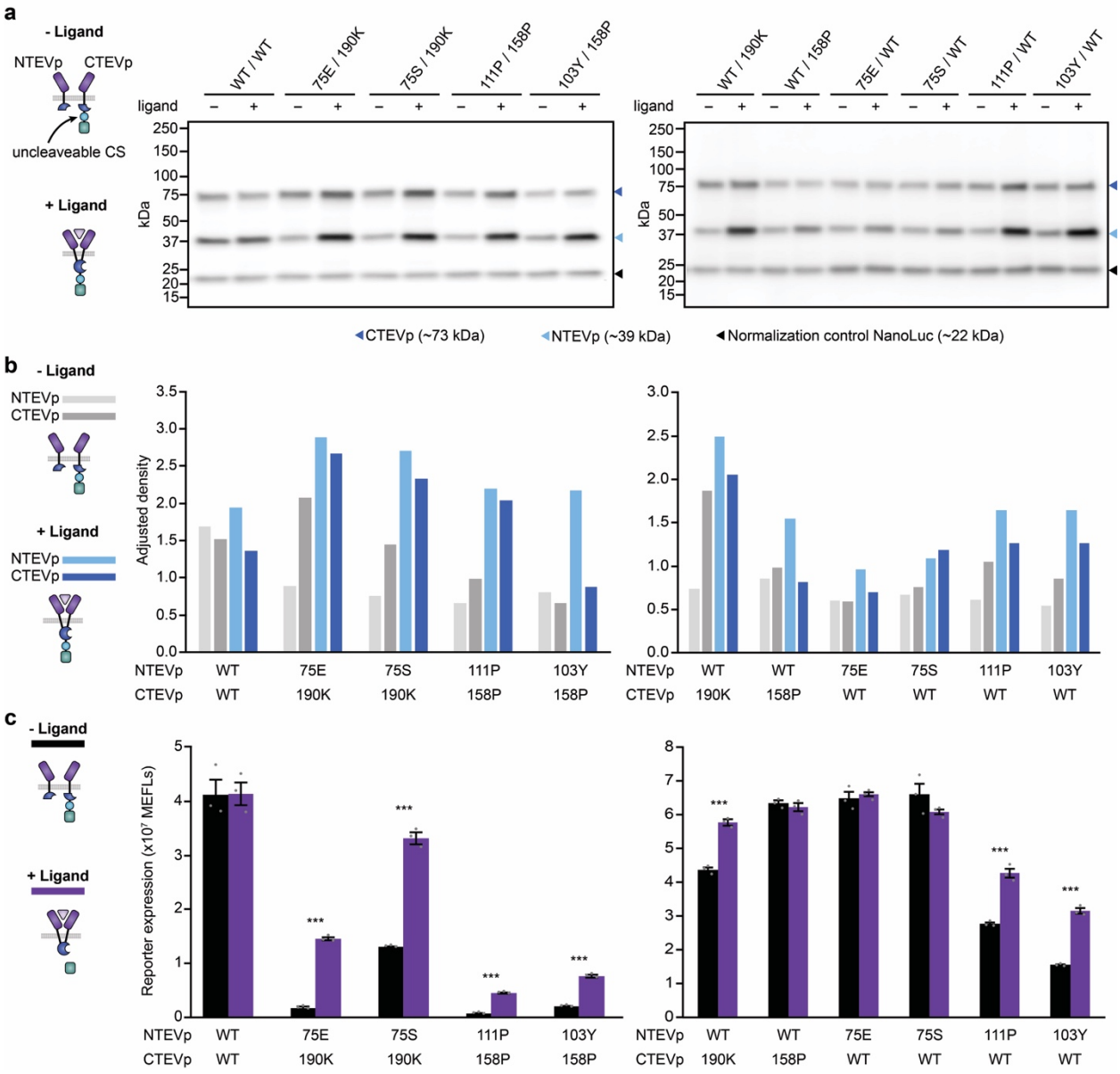
Supplementary Fig. 6 Additivity of energetic perturbations for mutations as computed by Rosetta. Data is shown for the sixty-seven double and paired mutants that can be created by combining the twenty single mutants evaluated in **Fig. 2c**.



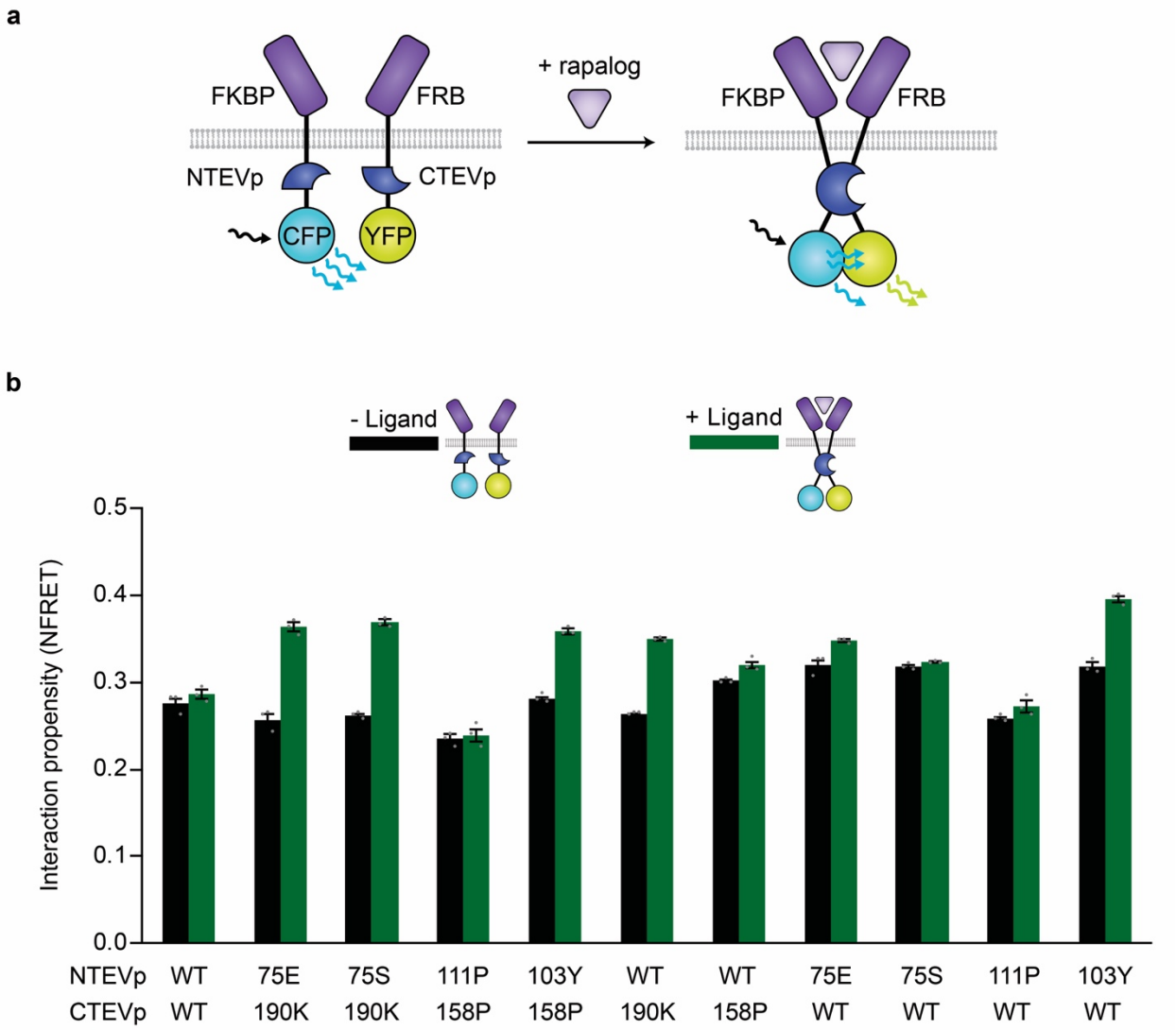
Supplementary Fig. 7 Western blot analysis of mutant split TEVp expression levels. Cells were transfected with single mutant chains which contain an N-terminal 3x-FLAG tag. The FKBP binding domain is on the NTEVp chain, and the FRB binding domain and tTA are on the CTEVp chain. Cells were co-transfected with a 3x-FLAG tagged NanoLuciferase as a normalization control (bottom band). Equal masses of protein (3 µg/lane) were loaded into each lane to investigate differences in chain expression level.



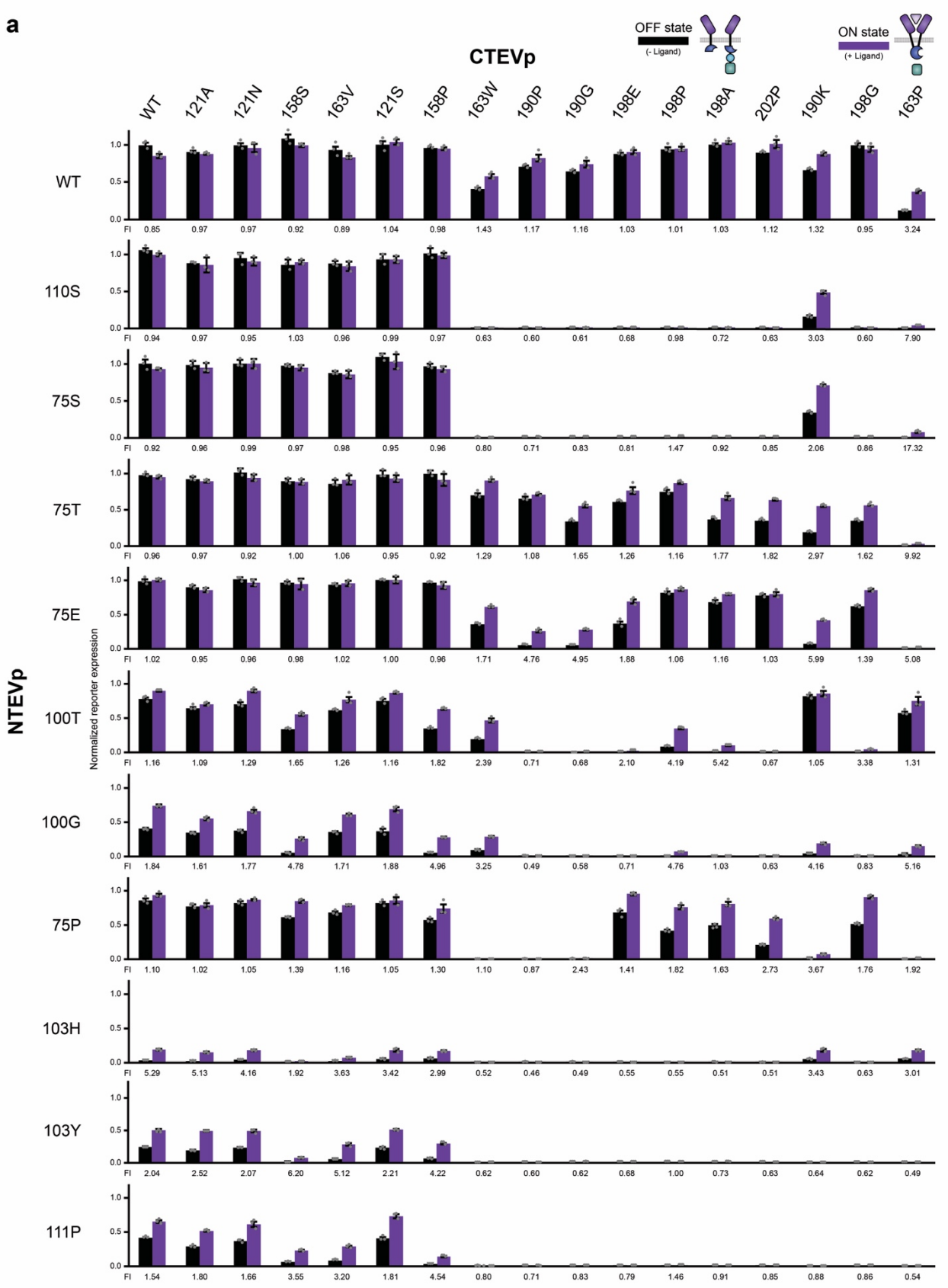
Supplementary Fig. 8 Optimized splits yield robust performance across a wide range of expression levels. Experimental analysis of DNA dose normalized and expression normalized select mutant pairs. The FKBP binding domain is on the NTEVp chain, and the FRB binding domain and tTA are on the CTEVp chain. Reconstituted TEVp cleavage of tTA measured by flow cytometry. Mutant pairs show inducible performance that is robust to the expression levels and ratios investigated, however WT performance is poor over this expression regime. Error bars depict S.E.M. Two-tailed Student's t-test (**p ≤ 0.01, ***p ≤ 0.001). Experiments were performed in biological triplicate. Results are representative of two independent biological experiments.



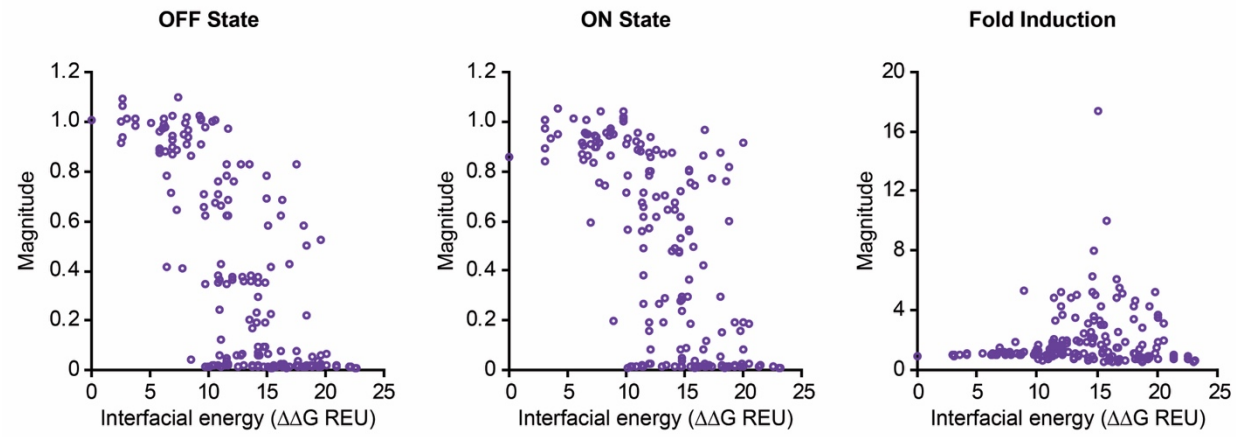
Supplementary Fig. 9 Effects of ligand addition on expression and accumulation of select split TEVp variants. a, Western blot analysis of split protein expression in the presence and absence of ligand. Cells were co-transfected with a WT or mutant NTEVp chain, a WT or mutant CTEVp chain, and 3x-FLAG tagged NanoLuciferase as a loading control. Arrows at the sides of blots indicate expected sizes for each species (FRB-CTEVp chain with uncleaveable cleavage sequence and tTA, FKBP-NTEVp chain, and NanoLuc loading control). Cells were exposed to rapalog ligand (+) or vehicle-only (-) for 24 hours before harvesting cell lysates. Equal masses of protein (8 µg/lane) were loaded for each case. **b,** Quantification of Western blot analyses in panel **a**. Band intensities were separately quantified for NTEVp and CTEVp chains; each measurement was normalized using the intensity of the NanoLuciferase loading control in its lane (to correct for variations in gel loading), and each loading-normalized metric was normalized to the corresponding value for the WT CTEVp chain to generate a scaled metric. **c,** To facilitate comparison, these functional data are reproduced from other figures (Left- subset of data from **Supplementary Fig. 8.**, Right- subset of data from **Fig. 2c.**) and describe the split TEVp pairs analyzed in panels **a** and **b**. Error bars depict S.E.M. Two-tailed Student's t-test (***p ≤ 0.001). Experiments were performed in biological triplicate. Results are representative of two independent biological experiments.



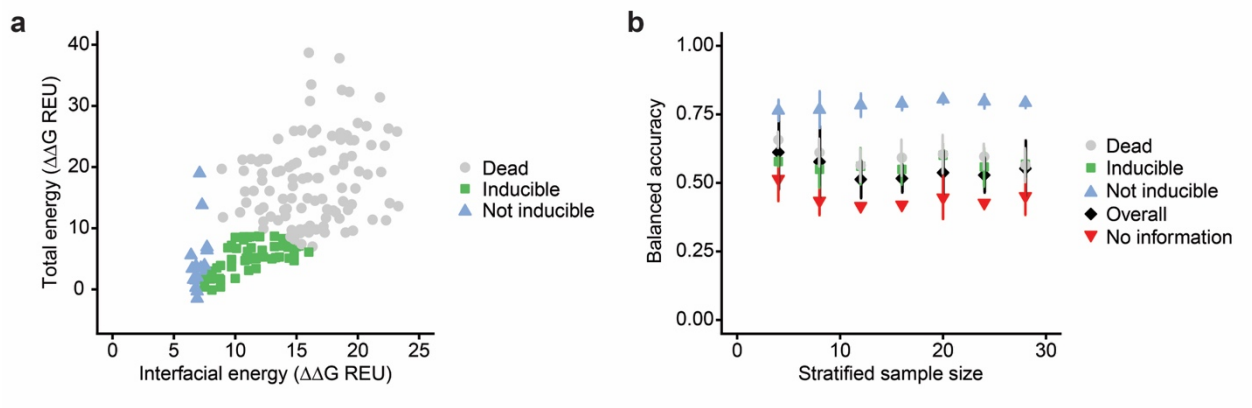
Supplementary Fig. 10 FRET analysis of select split TEVp paired chain variants. a, This cartoon illustrates the previously validated method[1] for quantifying ligand-independent (left) and ligand-mediated (right) receptor association using Förster resonance energy transfer (FRET). Fluorophores were appended to C-termini of the split TEVp receptor chains: mCerulean (donor) was added onto the FKBP-NTEVp chain, and mVenus (acceptor) was added onto the FRB-CTEVp chain, in place of the cleavage sequence and transcription factor. **b,** Interaction propensity measured by NFRET—a metric of FRET that is expression-normalized on a single cell basis[1]. Error bars depict S.E.M. Experiments were performed in biological triplicate. Results are representative of two independent biological experiments.


a
CTEVp
OFF state
(- Ligand)
ON state
(+ Ligand)
WT
121A
121N
158S
163V
121S
158P
163W
190P
190G
198E
198P
198A
202P
190K
198G
163P
NTEVp
WT
110S
75S
75T
75E
100T
100G
75P
103H
103Y
111P
Normalized reporter expression
FI
0.85 0.97 0.97 0.92 0.89 1.04 0.98 1.43 1.17 1.16 1.03 1.01 1.03 1.12 1.32 0.95 3.24
0.94 0.97 0.95 1.03 0.96 0.99 0.97 0.63 0.60 0.61 0.68 0.98 0.72 0.63 3.03 0.60 7.90
0.92 0.96 0.99 0.97 0.98 0.95 0.96 0.80 0.71 0.83 0.81 1.47 0.92 0.85 2.06 0.86 17.32
0.96 0.97 0.92 1.00 1.06 0.95 0.92 1.29 1.08 1.65 1.26 1.16 1.77 1.82 2.97 1.62 9.92
1.02 0.95 0.96 0.98 1.02 1.00 0.96 1.71 4.76 4.95 1.88 1.06 1.16 1.03 5.99 1.39 5.08
1.16 1.09 1.29 1.65 1.26 1.16 1.82 2.39 0.71 0.68 2.10 4.19 5.42 0.67 1.05 3.38 1.31
1.84 1.61 1.77 4.78 1.71 1.88 4.96 3.25 0.49 0.58 0.71 4.76 1.03 0.63 4.16 0.83 5.16
1.10 1.02 1.05 1.39 1.16 1.05 1.30 1.10 0.87 2.43 1.41 1.82 1.63 2.73 3.67 1.76 1.92
5.29 5.13 4.16 1.92 3.63 3.42 2.99 0.52 0.46 0.49 0.55 0.55 0.51 0.51 3.43 0.63 3.01
2.04 2.52 2.07 6.20 5.12 2.21 4.22 0.62 0.60 0.62 0.68 1.00 0.73 0.63 0.64 0.62 0.49
1.54 1.80 1.66 3.55 3.20 1.81 4.54 0.80 0.71 0.83 0.79 1.46 0.91 0.85 0.88 0.86 0.54

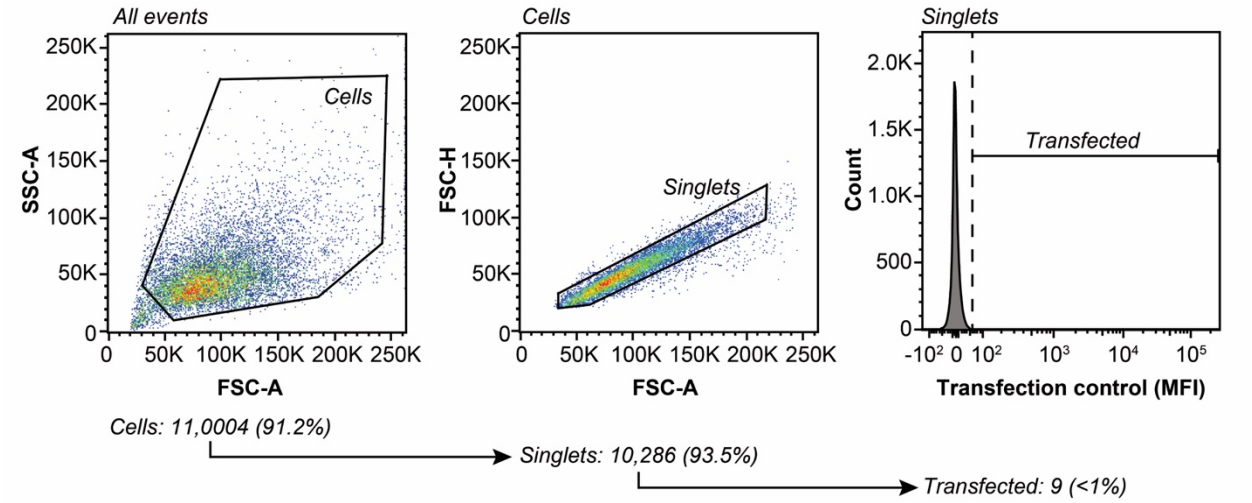
b



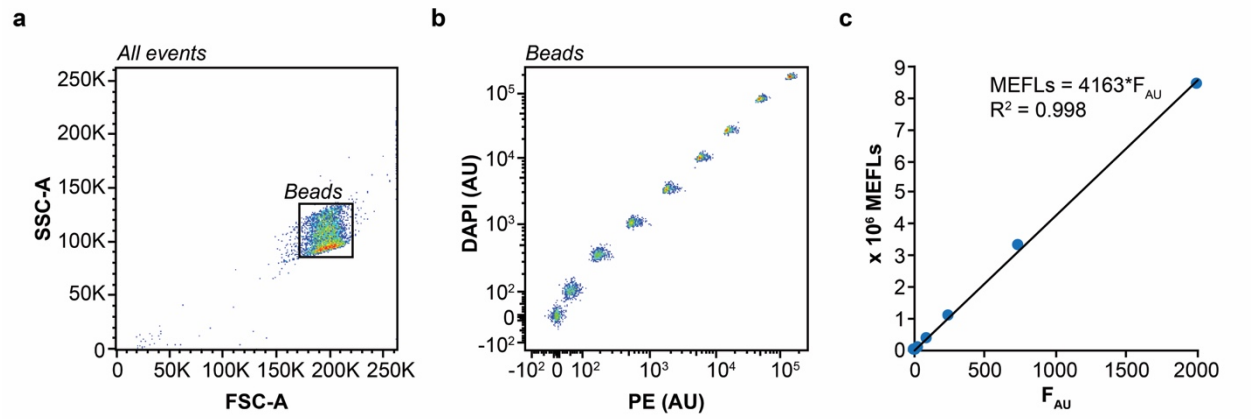
Supplementary Fig. 11 Supplemental analyses of all experimentally evaluated mutants. a, The complete data set of normalized reporter expression data reported in **Fig. 4a** as bubble plots is shown here in bar graph form. Fold induction (FI) is reported below each graph. Error bars depict S.E.M. Experiments were performed in biological triplicate. Results are representative of two independent biological experiments. **b**, All data reported in **Fig. 4a** are re-plotted here as a function of the interfacial energy calculated for each mutant/pair A table of all calculated interfacial energy values for each mutant/pair is provided in **Supplementary Table 4**.



Supplementary Fig. 12 Linear discriminant analysis (LDA) and retrospective bootstrapping. **a,** Phenotype predictions for experimentally characterized variants (omitting the calibration set) using an LDA model trained on our actual 20-variant calibration set. Note that the phenotype boundaries determined by the LDA analysis are similar to those that were subjectively inferred (**Fig. 2e**). **b**, Accuracy of phenotype predictions as a function of sample size using a retrospective bootstrapping approach of all 193 experimentally characterized variants. Classification accuracies were generated using LDA models trained from 10-fold stratified bootstrap partitioning of experimentally characterized variants into training sets**.** Accuracy was computed on variants not included in each training set. Error bars indicate the standard deviations across the 10-fold bootstrapping. Details and code for the analysis presented here are included in **Supplementary Software 1**.



Supplementary Fig. 13 Flow cytometry gating workflow. The plots show cells transfected with pcDNA vector only to illustrate the flow cytometry gating strategy used to identify single transfected cells. HEK293FT cells were first identified by the FSC-A vs. SSC-A profile. From this population, singlets were identified by the FSC-A vs. FSC-H profile. The transfected population was defined as all single cells with a transfection control signal greater than the sample of single cells transfected with pcDNA only, encompassing no more than 1% of the pcDNA-only transfected population.



Supplementary Fig. 14 Calibration of flow cytometry fluorescence intensities to absolute units with UltraRainbow beads. a, For each flow cytometry-based experiment, the bead population was identified based on the FSC-A vs. SSC-A profile. **b**, Two fluorescent channels, other than the channel of interest, were used to identity the 9 bead populations. **c**, The mean fluorescence intensity (MFI) of each population in the FITC channel (used for EYFP reporter expression) in arbitrary units (F_{AU}) was recorded and plotted against manufacturer provided values for fluorophores per bead for each populations (Molecules of Equivalent Fluorescein MEFLs). A calibration curve was generated, using a linear regression with y-intercept set to zero. This curve was used to convert exported MFI values to absolute units using the multiplier obtained from the regression.

SUPPLEMENTARY TABLES

Supplementary Table 1. Plasmid doses used in each figure

Figure 2c	
Plasmid doses:	
EBFP2 transfection control	200 ng
Reporter plasmid (EYFP)	200 ng
NTEVp plasmid	25 ng
CTEVp plasmid	25 ng
pcDNA	150 ng
Plasmid numbers:	
EBFP2: pPD193	
Reporter: pPD463	
NTEVp: pTD798, pTD825, pTD827, pTD828, pTD829, pTD834, pTD835, pTD837, pTD839, pTD840, pTD841	
CTEVp: pTD799, pTD824, pTD826, pTD830, pTD831, pTD832, pTD833, pTD836, pTD838, pTD842, pTD843	

Figure 3b	
Plasmid doses:	
EBFP2 transfection control	200 ng
Reporter plasmid (EYFP)	200 ng
NTEVp plasmid	25 ng
CTEVp plasmid	25 ng
pcDNA	150 ng
Plasmid numbers:	
EBFP2: pPD193	
Reporter: pPD463	
NTEVp: pTD825, pTD827, pTD828, pTD829, pTD837	
CTEVp: pTD824, pTD826, pTD830, pTD831, pTD832, pTD833, pTD838, pTD843	

Figure 3d	
Plasmid doses:	
EBFP2 transfection control	200 ng
Reporter plasmid (EYFP)	200 ng
NTEVp plasmid	0.4 ng, 1.4 ng, 1 ng (order corresponding to plasmid numbers below)
CTEVp plasmid	5 ng, 12 ng (order corresponding to plasmid numbers below)
pcDNA	188 to 195 ng
Plasmid numbers:	
EBFP2: pPD193	

Reporter: pPD463
NTEVp: pTD798, pTD825, pTD828
CTEVp: pTD799, pTD843

Figure 4a	
Plasmid doses:	
EBFP2 transfection control	200 ng
Reporter plasmid (EYFP)	200 ng
NTEVp plasmid	25 ng
CTEVp plasmid	25 ng
pcDNA	150 ng
Plasmid numbers:	
EBFP2: pPD193	
Reporter: pPD463	
NTEVp: pTD798, pTD825, pTD827, pTD828, pTD829, pTD837, pTD839, pTD863, pTD864, pTD866, pTD870	
CTEVp: pTD799, pTD824, pTD826, pTD830, pTD831, pTD832, pTD833, pTD838, pTD843, pTD872, pTD873, pTD874, pTD875, pTD876, pTD877, pTD878, pTD879	

Figure 5b	
Plasmid doses:	
EBFP2 transfection control	200 ng
Reporter plasmid (EYFP)	200 ng
TEVp plasmid	25 ng
tTA plasmid	50 ng
pcDNA	125 ng
Plasmid numbers:	
EBFP2: pPD193	
Reporter: pPD463	
TEVp: pPD801	
tTA: pTD618, pTD619, pTD620, pTD621, pTD622, pTD623, pTD624, pTD625, pTD626, pTD627, pTD628, pTD629, pTD630, pTD631, pTD632, pTD634, pTD635	

Figure 5c	
Plasmid doses:	
EBFP2 transfection control	200 ng
Reporter plasmid (EYFP)	200 ng
NTEVp plasmid	25 ng
CTEVp plasmid	25 ng

tTA plasmid	50 ng
pcDNA	100 ng
Plasmid numbers:	
EBFP2: pPD193	
Reporter: pPD463	
tTA: pTD634	
NTEVp: pTD811, pJB006, pJB007, pJB008, pJB009, pJB010	
CTEVp: pTD844, pJB011, pJB012, pJB013, pJB014, pJB015	

Supplementary Figure 1	
Plasmid doses:	
EBFP2 transfection control	200 ng
Reporter plasmid (EYFP)	200 ng
NTEVp plasmid	25 ng
CTEVp plasmid	25 ng
pcDNA	150 ng
Plasmid numbers:	
EBFP2: pPD193	
Reporter: pPD463	
NTEVp: pTD516, pTD518, pTD526, pTD527	
CTEVp: pTD524, pTD525, pTD528, pTD529	

Supplementary Figure 2	
Plasmid doses:	
EBFP2 transfection control	200 ng
Reporter plasmid (EYFP)	200 ng
NTEVp plasmid	25 ng
CTEVp plasmid	25 ng
pcDNA	150 ng
Plasmid numbers:	
EBFP2: pPD193	
Reporter: pPD463	
NTEVp: pTD798, pTD800, pTD801, pTD802, pTD803, pTD804, pTD805	
CTEVp: pTD799, pTD806, pTD807, pTD808, pTD809, pTD810	

Supplementary Figure 4	
Plasmid doses:	
EBFP2 transfection control	200 ng

NTEVp plasmid	100 ng
CTEVp plasmid	100 ng
Plasmid numbers:	
EBFP2: pPD193	
NTEVp: pTD798, pTD825, pTD827, pTD828, pTD829, pTD834, pTD835, pTD837, pTD839, pTD840, pTD841	
CTEVp: pTD799, pTD824, pTD826, pTD830, pTD831, pTD832, pTD833, pTD836, pTD838, pTD842, pTD843	

Supplementary Figure 6	
Plasmid doses:	
EBFP2 transfection control	200 ng
NanoLuc control	1 ng
NTEVp plasmid	5 ng
CTEVp plasmid	100 ng, 100 ng, 100 ng, 100 ng, 200 ng, 200 ng, 200 ng (order corresponding to plasmid numbers below)
pcDNA	100 to 395 ng
Plasmid numbers:	
EBFP2: pPD193	
NanoLuc: pPD463	
NTEVp: pTD798, pTD825, pTD827, pTD828, pTD829, pTD837, pTD839	
CTEVp: pTD799, pTD824, pTD826, pTD830, pTD831, pTD838, pTD843	

Supplementary Figure 7	
Plasmid doses:	
EBFP2 transfection control	200 ng
Reporter plasmid (EYFP)	200 ng
NTEVp plasmid (DNA 1:1)	25 ng
NTEVp plasmid (Expression 1:1)	0.4 ng, 1.4 ng, 1 ng, 1.2 ng, 0.7 ng (order corresponding to plasmid numbers below)
CTEVp plasmid (DNA 1:1)	25 ng
CTEVp plasmid (Expression 1:1)	25 ng, 53.8 ng, 59.6 ng (order corresponding to plasmid numbers below)
pcDNA	130 to 175 ng
Plasmid numbers:	
EBFP2: pPD193	
Reporter: pPD463	
NTEVp: pTD798, pTD825, pTD828, pTD837, pTD839	
CTEVp: pTD799, pTD824, pTD843	

Supplementary Figure 9a,b	
Plasmid doses:	
EBFP2 transfection control	200 ng
NanoLuc control	1 ng
NTEVp plasmid	5 ng
CTEVp plasmid	100 ng, 100 ng, 200 ng (order corresponding to plasmid numbers below)
pcDNA	195 to 295 ng
Plasmid numbers:	
EBFP2: pPD193	
NanoLuc: pPD463	
NTEVp: pTD798, pTD825, pTD827, pTD828, pTD829, pTD837, pTD839	
CTEVp: pTD962, pTD963, pTD964	

Supplementary Figure 9c	
Plasmid doses:	
EBFP2 transfection control	200 ng
Reporter plasmid (EYFP)	200 ng
NTEVp plasmid	25 ng
CTEVp plasmid	25 ng
pcDNA	150 ng
Plasmid numbers:	
EBFP2: pPD193	
Reporter: pPD463	
NTEVp: pTD798, pTD825, pTD828, pTD837, pTD839	
CTEVp: pTD799, pTD824, pTD843	

Supplementary Figure 10b	
Plasmid doses:	
mIRFP670 transfection control	200 ng
mCerulean-tagged NTEVp plasmid	25 ng
mVenus-tagged CTEVp plasmid	25 ng
pcDNA	350 ng
Plasmid numbers:	
mIRFP670: pPD861	
mCerulean-tagged NTEVp: pTD954, pTD956, pTD957, pTD958, pTD959	
mVenus-tagged CTEVp: pTD955, pTD960, pTD961	

Supplementary Table 2. BD LSR Fortessa instrument specifications for analytical flow cytometry to quantify reporter expression

Fluorescent protein	Parameter	Excitation laser	Filter set
EBFP2	Pacific Blue	Violet, 405 nm	450/50
EYFP	FITC	Blue, 488 nm	505LP, 530/30

Supplementary Table 3. BD LSR Fortessa instrument specifications for analytical flow cytometry to quantify FRET

Fluorescent protein	Parameter	Excitation laser	Filter set
mCerulean	Pacific Blue	Violet, 405 nm	450/50
mVenus	FITC	Blue, 488 nm	505LP, 530/30
FRET	AmCyan	Violet, 405 nm	505LP, 530/30
miRFP670	APC	Red, 640 nm	670/30

Supplementary Table 4. Calculated interfacial energy values for each mutant/pair

NTEVp	CTEVp	ΔΔG (REU)
WT	WT	0
	M121A	3.1
	M121N	3.1
	T158S	3.2
	I163V	3.2
	M121S	4.3
	T158P	6.6
	I163W	7
	L190P	7.3
	L190G	7.8
	W198E	7.8
	W198P	8.4
	W198A	9.9
	W202P	9.9
	L190K	11.1
	W198G	11.1
	I163P	11.6
H75T	WT	4.3
	M121A	7.4
	M121N	7.4
	T158S	7.5
	I163V	7.5
	M121S	8.6
	T158P	10.9
	I163W	11.3
	L190P	11.6
	L190G	12.1
	W198E	12.1
	W198P	12.7
	W198A	14.2
	W202P	14.2
	L190K	15.4
	W198G	15.4
	I163P	15.9

NTEVp	CTEVp	ΔΔG (REU)
C110S	WT	3.2
	M121A	6.3
	M121N	6.3
	T158S	6.4
	I163V	6.4
	M121S	7.5
	T158P	9.8
	I163W	10.2
	L190P	10.5
	L190G	11
	W198E	11
	W198P	11.6
	W198A	13.1
	W202P	13.1
	L190K	14.3
	W198G	14.3
	I163P	14.8
H75E	WT	5.6
	M121A	8.7
	M121N	8.7
	T158S	8.8
	I163V	8.8
	M121S	9.9
	T158P	12.2
	I163W	12.6
	L190P	12.9
	L190G	13.4
	W198E	13.4
	W198P	14
	W198A	15.5
	W202P	15.5
	L190K	16.7
	W198G	16.7
	I163P	17.2

NTEVp	CTEVp	ΔΔG (REU)
H75S	WT	3.6
	M121A	6.7
	M121N	6.7
	T158S	6.8
	I163V	6.8
	M121S	7.9
	T158P	10.2
	I163W	10.6
	L190P	10.9
	L190G	11.4
	W198E	11.4
	W198P	12
	W198A	13.5
	W202P	13.5
	L190K	14.7
	W198G	14.7
	I163P	15.2
F100T	WT	7
	M121A	10.1
	M121N	10.1
	T158S	10.2
	I163V	10.2
	M121S	11.3
	T158P	13.6
	I163W	14
	L190P	14.3
	L190G	14.8
	W198E	14.8
	W198P	15.4
	W198A	16.9
	W202P	16.9
	L190K	18.1
	W198G	18.1
	I163P	18.6

F100G	WT	8.3
	M121A	11.4
	M121N	11.4
	T158S	11.5
	I163V	11.5
	M121S	12.6
	T158P	14.9
	I163W	15.3
	L190P	15.6
	L190G	16.1
	W198E	16.1
	W198P	16.7
	W198A	18.2
	W202P	18.2
	L190K	19.4
	W198G	19.4
	I163P	19.9

H75P	WT	9
	M121A	12.1
	M121N	12.1
	T158S	12.2
	I163V	12.2
	M121S	13.3
	T158P	15.6
	I163W	16
	L190P	16.3
	L190G	16.8
	W198E	16.8
	W198P	17.4
	W198A	18.9
	W202P	18.9
	L190K	20.1
	W198G	20.1
	I163P	20.6

P103H	WT	9
	M121A	12.1
	M121N	12.1
	T158S	12.2
	I163V	12.2
	M121S	13.3
	T158P	15.6
	I163W	16
	L190P	16.3
	L190G	16.8
	W198E	16.8
	W198P	17.4
	W198A	18.9
	W202P	18.9
	L190K	20.1
	W198G	20.1
	I163P	20.6

P103Y	WT	11.5
	M121A	14.6
	M121N	14.6
	T158S	14.7
	I163V	14.7
	M121S	15.8
	T158P	18.1
	I163W	18.5
	L190P	18.8
	L190G	19.3
	W198E	19.3
	W198P	19.9
	W198A	21.4
	W202P	21.4
	L190K	22.6
	W198G	22.6
	I163P	23.1

L111P	WT	11.6
	M121A	14.7
	M121N	14.7
	T158S	14.8
	I163V	14.8
	M121S	15.9
	T158P	18.2
	I163W	18.6
	L190P	18.9
	L190G	19.3
	W198E	19.4
	W198P	20
	W198A	21.5
	W202P	21.5
	L190K	22.7
	W198G	22.7
	I163P	23.2

SUPPLEMENTARY NOTE

Supplementary Note 1. Rosetta calculations used in this manuscript

Computational Interface Scanning

Local Rosetta runs were performed with release version 3.9. The ref2015 scoring function was used for all calculations.

Protein Preparation

The structure of TEVp was obtained from the Research Crystallography for Structural Bioinformatics (RCSB) PDB (ID code: 1LVM). Chain designations A and B were used for residues 4-118 (NTEVp) and 119-221 (CTEVp), respectively. The structure was idealized, which sets bond lengths and angles to their ideal values and then minimizes the structure in the presence of coordinate constraints.

Commands:

```
${path_to_Rosetta}/main/source/bin/idealize_jd2.linuxgccrelease
        -database ${path_to_Rosetta}/main/database
        -in::file::fullatom
        -s 1LVM.pdb
        -no_optH false
        -flip_HNQ
```

Interface Mutagenesis

The idealized structure of TEVp was relaxed (100 iterations) using the default constrained relax script, and position constraints were added to backbone heavy atoms based on the crystal structure. During this relax procedure, Rosetta resfiles were used to incorporate single (or double) mutations.

Commands:

```
${path_to_Rosetta}/main/source/bin/relax.linuxgccrelease
        -database ${path_to_Rosetta}/main/database
        -relax::sequence_file always_constrained_relax.script
        -constrain_relax_to_start_coords
        -relax::coord_cst_width 0.25 -relax::coord_cst_stdev 0.25
        -s 1LVM.pdb
        -in::file::fullatom
        -no_optH false
        -flip_HNQ
        -nstruct 100
        -packing:resfile ${mutant}.resfile
        -relax:respect_resfile
```

always_constrained_relax.script:

```
        repeat 5
                ramp_repack_min 0.02  0.01     1.0
                ramp_repack_min 0.250 0.01     1.0
                ramp_repack_min 0.550 0.01     1.0
                ramp_repack_min 1     0.00001  1.0
                accept_to_best
        endrepeat
```

example ${mutant}.resfile:

```
        AUTO
        NATAA
        start
        75  A PIKAA A
```

Interface Scoring

The interfacial energy was computed for all relaxed structures of each variant using the rosetta_scripts application with the InterfaceAnalyzerMover. This mover calculates the total interaction energy between all residues in chain A (nTEV) with residues in chain B (cTEV).

Commands:

```
${path_to_Rosetta}/main/source/bin/rosetta_scripts.linuxgccrelease
        -database ${path_to_Rosetta}/main/database
```

```
        -parser:protocol interface_score.script
        -file:s ${mutant}.pdb
```

interface_score.script

```
<ROSETTASCRIPTS>
        <TASKOPERATIONS>
        </TASKOPERATIONS>
        <SCOREFXNS>
        </SCOREFXNS>
        <FILTERS>
        </FILTERS>
        <MOVERS>
                <InterfaceAnalyzerMover name="score_int" pack_separated="false" pack_input="false"
        packstat="false" interface_sc="true" ligandchain="B"/>
        </MOVERS>
        <PROTOCOLS>
        <Add mover_name="score_int"/>
        </PROTOCOLS>
</ROSETTASCRIPTS>
```

REFERENCES CITED IN THIS DOCUMENT

1. Edelstein, H.I. et al. Elucidation And Refinement Of Synthetic Receptor Mechanisms. *Synthetic Biology* (2020).