Supporting Information

Characterization of a sulfated anti-HIV antibody using an expanded genetic code

Xiang Li[1], Justin Hitomi[1], Chang C. Liu[1,2,3]

1. Department of Biomedical Engineering, University of California, Irvine, California 92617, United States
2. Department of Chemistry, University of California, Irvine, California 92617, United States
3. Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92617, United States

Materials and Methods

Construction of expression plasmids for E51 Fabs and small hairpin RNAs

For Fab expression in bacterial cells, E51 Fab cassette for sulfoform 1 was synthesized (IDT) and cloned into the pGLO expression vector via Gibson Assembly. The remaining E51 sulfoforms with different numbers of UAG codons were sequentially cloned via inverse PCR site-directed mutagenesis. The resulting pGLO-E51 plasmids were co-transformed with pULTRA-sY[1] (Addgene #82417) into C321.ΔA.exp cells[2] (Addgene #49018).

For expression in mammalian cells, the variable and constant region of E51 heavy chain fused to a C-terminal HA tag for sulfoform **1** was cloned into a pCMV expression vector from pGLO-E51 via Gibson Assembly while the variable and constant region of E51 light chain for sulfoform 1 was cloned into a separate pCMV following the same process. DNA fragments corresponding to small hairpin RNAs (shRNAs) that target tyrosylprotein sulfotransferase-1 (TPST-1) (5'-GGGCCATGCTGGACGCACATC-3') or tyrosylprotein sulfotransferase-2 (TPST-2) (5'-GGGACAGCAGGTGCTAGAGTG-3') were cloned into pSQT1313 (Addgene #53370) via inverse PCR.

Expression of E51 Fabs in bacterial cells with an expanded genetic code for sY

Thirty-two E51 sulfoforms were expressed in C321.ΔA.exp cells in M9T minimal media, which improves the permeability of sY into cells for high expression of sulfated proteins[3,4]. Briefly, single colonies of C321.ΔA.exp cells transformed with pULTRA-sY and pGLO-E51 were picked into Terrific Broth (TB). After growing the cells at 37°C and 200 rpm for 16 hours, the saturated cultures in TB were diluted into M9T (1X M9 salts, 10 g/L tryptone, 5 g/L NaCl, 1 mM MgSO4, and 80 μM biotin) at a ratio of 1:100 and grown to saturation at 37°C and 200 rpm for 12 hours. Two mL of the saturated M9T culture was added to 200 mL of M9T with 0, 3, 5, or 10 mM of sY for expressing E51 sulfoforms with 0, 1-3, 4, or 5 UAG codons, respectively, and grown to mid-log phase (OD600 = 0.6-0.8) at 37°C and 200 rpm. The expression of STyrRS and E51 Fabs was induced by adding 0.2% of L-arabinose and 1 mM of IPTG and incubating the cultures at 22°C and 200 rpm for 24 hours.

Fabs were extracted from the cultures using either of two methods. The first method involved extracting Fabs from the periplasm through cold osmotic shock. Cell pellets were resuspended in 16 mL of cold periplasmic lysis buffer (20% sucrose, 20 mM Tris, 2 mM EDTA, pH 8.0) with protease inhibitor cocktail (Sigma) and incubated at 22°C and 200 rpm for 4 hours. The supernatant of the lysis consisting of the first periplasmic fraction was stored at 4°C, and the cell pellet was resuspended in 16 mL of cold periplasmic lysis buffer with protease inhibitor cocktail and incubated overnight at 4°C. The second supernatant of the lysis consisting of the second periplasmic extraction was combined with the first periplasmic extraction. The combined periplasmic fractions were filtered using a 0.22 µm filter, concentrated using a 10 kDa molecular weight cut off (MWCO) spin-filter column (MilliporeSigma), and then purified using Protein G columns (Pierce) according to the manufacturer's protocols. The second method involved concentrating Fabs from the supernatant of the culture via ammonium sulfate precipitation. After the supernatant of Fab cultures was filtered using a 0.22 µm filter, 0.2 M of ammonium sulfate was gradually added to the supernatant with stirring, and the solution was incubated to fully mix at 22°C for 1 hour. The precipitated Fabs were pelleted through centrifugation at 4°C and 10,000 g for 15 min, resuspended in PBS, concentrated using a 10 kDa molecular weight cut off (MWCO) spin-filter column (MilliporeSigma), and then purified using Protein G columns (Pierce) according to the manufacturer's protocols. The final concentrations of Fabs were measured on a microvolume spectrophotometer (NanoDrop) using the appropriate molecular weight and ε value calculated by ExPASy ProtParam.

Expression of E51 Fabs in mammalian cells

E51 Fabs sulfated by TPSTs were expressed in HEK293T cells. To increase the sulfation of E51, plasmids encoding the variable and constant region of E51 light or heavy chain were co-transfected into HEK293T cells in serum media at ~90% confluency with plasmids encoding TPST-1 or TPST-2 (Addgene #11252 and 11253 respectively). To decrease the sulfation of E51, plasmids encoding the variable and constant region of E51 light or heavy chain were co-transformed into HEK293T cells in serum media at ~90% confluency with plasmids encoding shRNAs that target TPST-1 or TPST-2. After 12 hours, the media was swapped with serum-free media and subsequently incubated at 28°C for 5 days for protein expression. The supernatant of the culture was collected, concentrated using a 10 kDa MWCO spin-filter column, and purified using protein G columns.

Analysis of gp120 binding using ELISA

0.2 μg of soluble gp120 (ADA, Clade B) or gp120(Q422L) (Immune-tech) in 100 μL of PBS was coated onto the surface of a polystyrene high binding microplate well at 4°C for overnight. After blocking the wells with 200 μL of 2% milk at 37°C for 2 hours and then washing 3 times with 300 μL of PBST (PBS + 0.05% Tween-20), various concentrations of Fab samples with the addition of 0.2 μg of sCD4 (Progenics) in 100 μL of 2% Milk PBST were added and incubated at 37°C for 2 hours. The wells were washed 5 times with 200 μL of PBST, incubated with 100 μL of a rabbit anti-HA polyclonal antibody (ThermoFisher Scientific) at 37°C for 1 hour, washed 5 times with 200 μL of PBST, and then incubated with 100 μL of an HRP-conjugated anti-rabbit polyclonal antibody (Sigma) at 37°C for 1 hour, washed 5 times with 200 μL of PBST, and incubated with 100 μL of QuantaBlu solution (ThermoFisher Scientific) at 22°C for 30 minutes. Fluorescence (ex: 325 nm, em: 420 nm) was measured using a microplate reader (TECAN). ELISA binding curves were fitted using a 4PL model fit and the concentration at half of the maximum binding signal (EC50) was determined by interpolation. ELISA binding curves and EC50 fits were determined using Graphpad Prism 5.0.

Mass spectrometry of Fabs

Intact Fab samples were analyzed for intact mass by UPLC-MS. E51 Fabs were dialyzed in 0.1% formic acid and diluted to a concentration of 100 ng/μL. Two μL of the diluted samples were injected into a Xevo G2-XS QTof mass spectrometer equipped with a lockspray ion source (Waters). Intact Fabs were separated with an ACQUITY UPLC protein column (300Å, 2.1 x 50 mm, BEHC4 1.7 μm) using a 10 minute linear gradient at a flow rate of 0.2 mL/min. The instrument was operated in positive-ion mode with a capillary voltage of 3.0 kV. Solvent A was composed of 0.1% formic acid in water, and solvent B was composed of 0.1% formic acid in acetonitrile. The samples were calibrated against a lock mass composed of 50 ng/mL of leucine enkephalin, and the instrument was calibrated with sodium iodide. MS1 scan was performed at 0.5 seconds across the range of 400-4000 m/z. The desolvation gas and ion source temperatures were set to 300°C and 100°C, respectively.

The extracted charge state ladders of the intact Fabs were deconvolved using MaxEnt 1 on the Applied Waters Markerlynx software. The deconvolution was done with the following settings: mass range of 45,000–50,000 Da, resolution of 0.45 Da/channel, damage model set to Uniform Gaussian mode with width at half height of 0.8 Da, minimum intensity ratios at 33% each, and maximum 12 iterations.

References:

(1) Chatterjee, A., Sun, S. B., Furman, J. L., Xiao, H., and Schultz, P. G. (2013) A versatile platform for single- and multiple-unnatural amino acid mutagenesis in Escherichia coli. *Biochemistry 52*, 1828-1837.

(2) Lajoie, M. J., Rovner, A. J., Goodman, D. B., Aerni, H.-R., Haimovich, A. D., Kuznetsov, G., Mercer, J. A., Wang, H. H., Carr, P. A., Mosberg, J. A., Rohland, N., Schultz, P. G., Jacobson, J. M., Rinehart, J., Church, G. M., and Isaacs, F. J. (2013) Genomically Recoded Organisms Expand Biological Functions. *Science 342*, 357–360.

(3) Li, X., and Liu, C. C. (2018) Site-specific incorporation of sulfotyrosine using an expanded genetic code. *Methods Mol. Biol. 1728*, 191–200.

(4) Schwessinger, B., Li, X., Ellinghaus, T. L., Chan, L. J. G., Wei, T., Joe, A., Thomas, N., Pruitt, R., Adams, P. D., Chern, M. S., Petzold, C. J., Liu, C. C., and Ronald, P. C. (2016) A second-generation expression system for tyrosine-sulfated proteins and its application in crop protection. *Integr. Biol. 8*, 542-545.

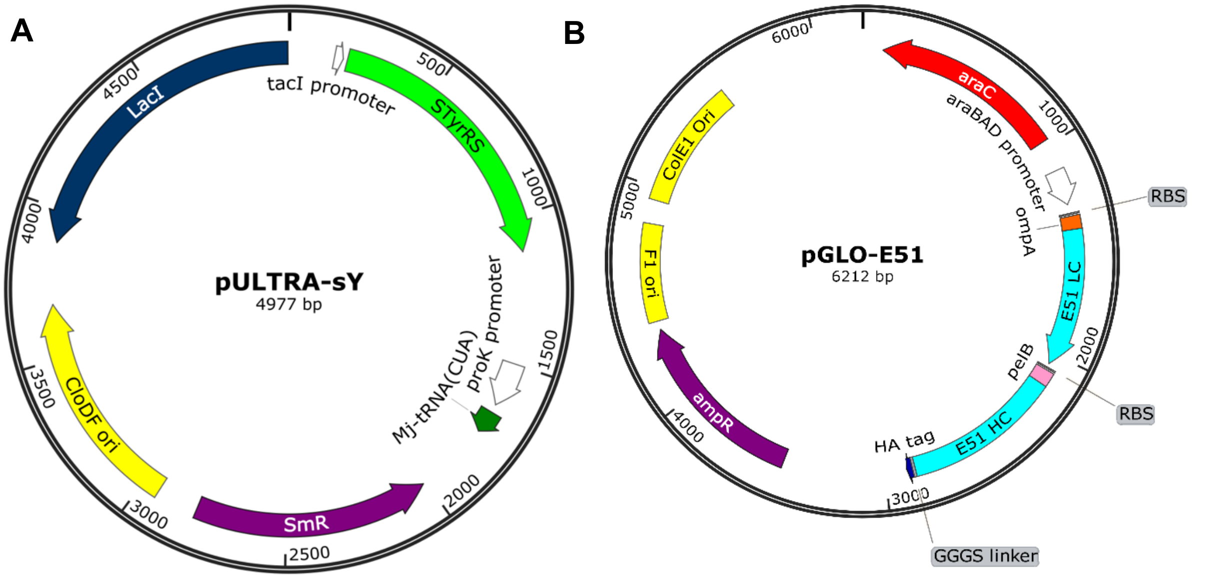


Figure S1. Plasmid maps. (A) pULTRA-sY constitutively expresses an engineered *Methanococcus jannaschii* tRNA that decodes UAG codons (Mj-tRNA$_{CUA}$) and a *M. jannaschii* aminoacyl-tRNA synthetase engineered to charge sulfotyrosine (STyrRS) under a tacI promoter inducible by IPTG. (B) pGLO-E51 contains an E51 Fab expression cassette under an araBAD promoter inducible by L-arabinose. The Fab cassette encodes the E51 light chain (LC) with an N-terminal OmpA signaling peptide and the E51 heavy chain (HC) with an N-terminal pelB signaling peptide and a C-terminal HA tag. Up to five tyrosine codons in the V$_H$CDR3 were replaced with UAG stop codons in each E51 sulfoform.

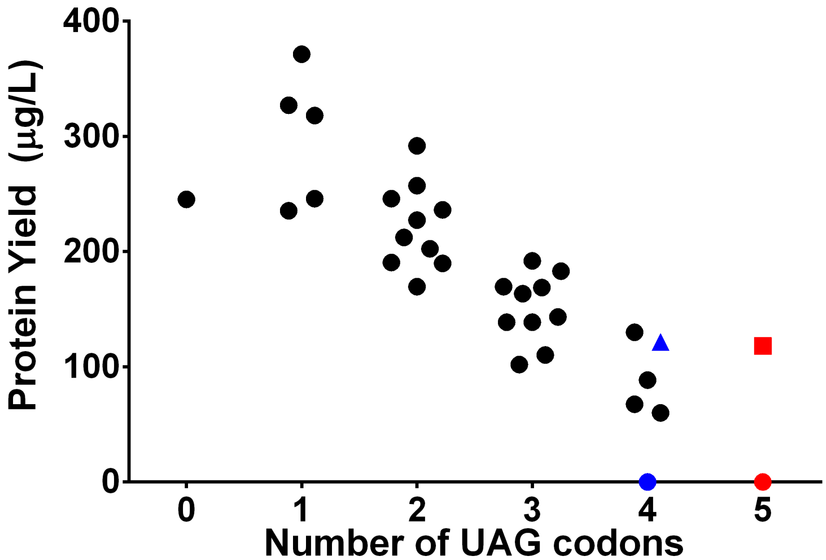


Figure S2. Yields of E51 sulfoforms **1-27** and **29-31** expressed in C321.ΔA.exp cells grown in media with 3 mM of sY (black), sulfoform **28** in 3 mM (blue dot) or 5 mM (blue triangle) of sY, and **32** in 3 mM (red dot) or 10 mM (red square) of sY.

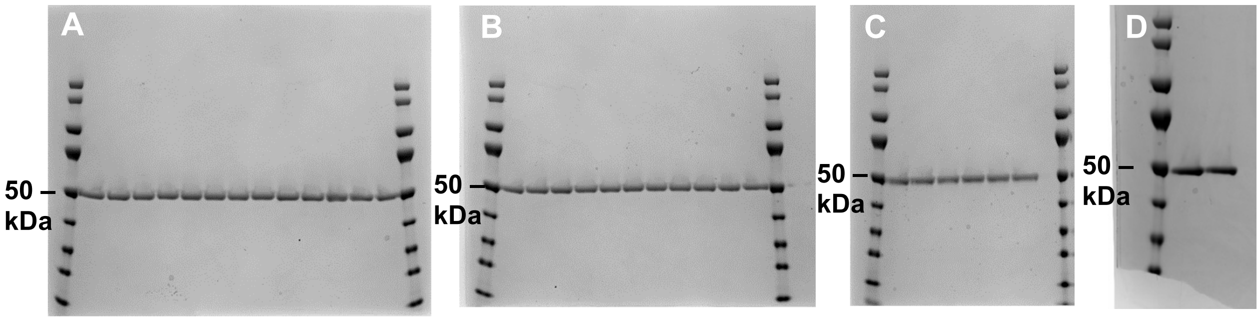
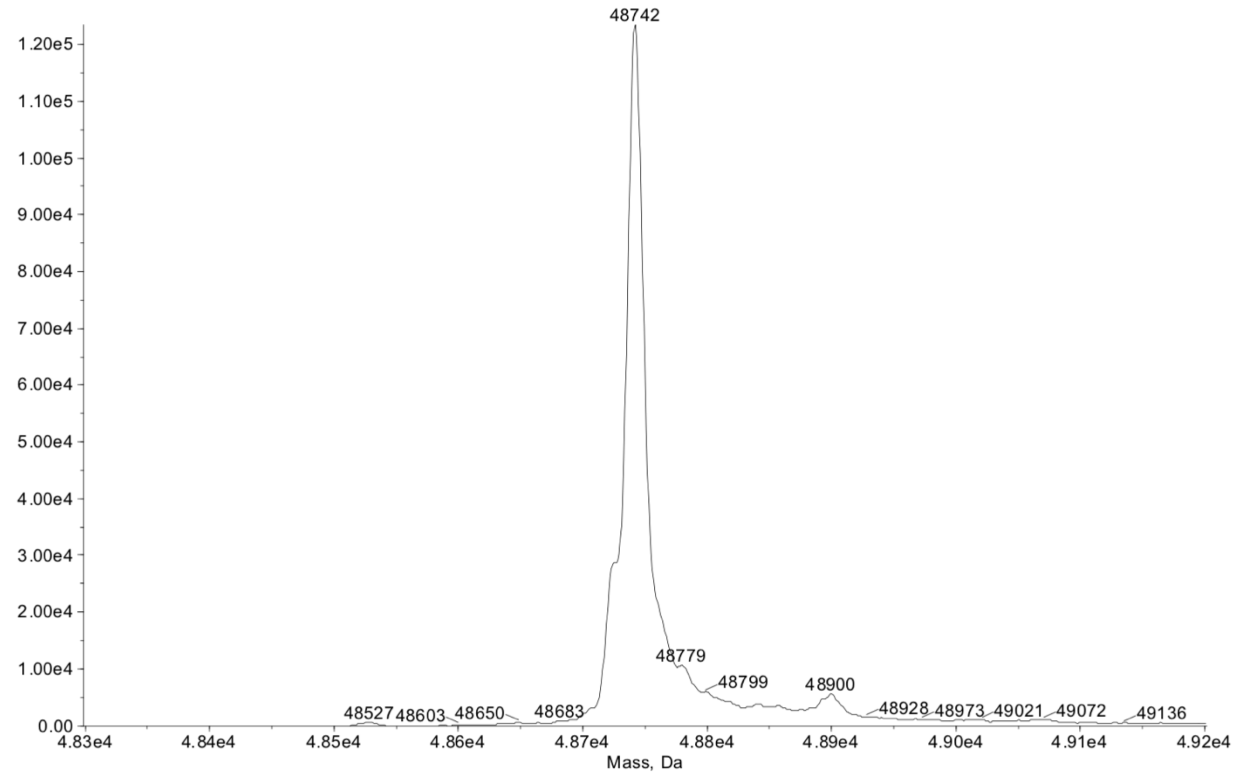
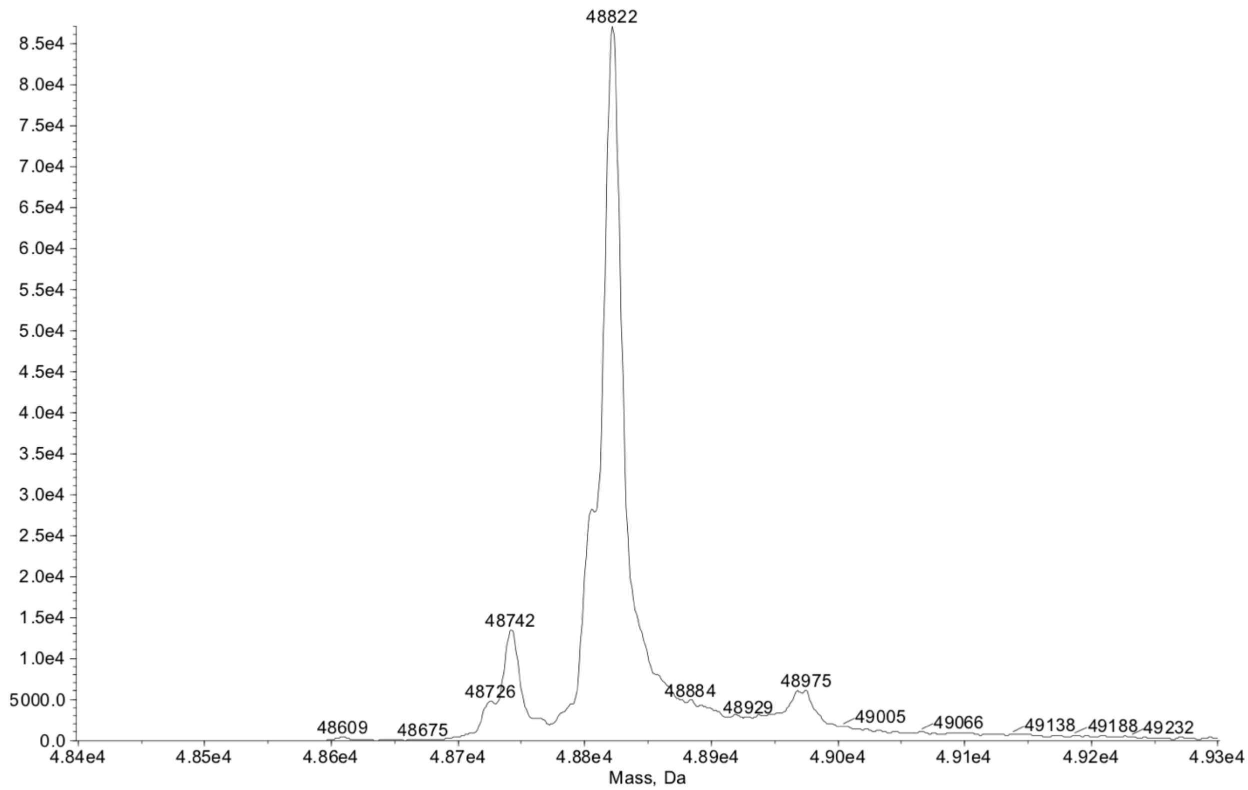


Figure S3. Visualization of 32 E51 sulfoform Fabs on SDS-PAGE gels. (A) Sulfoforms **1-13**. (B) Sulfoforms **14-24**. (C) Sulfoforms **25-27** and **29-31**. (D) Sulfoforms **28** and **32**.

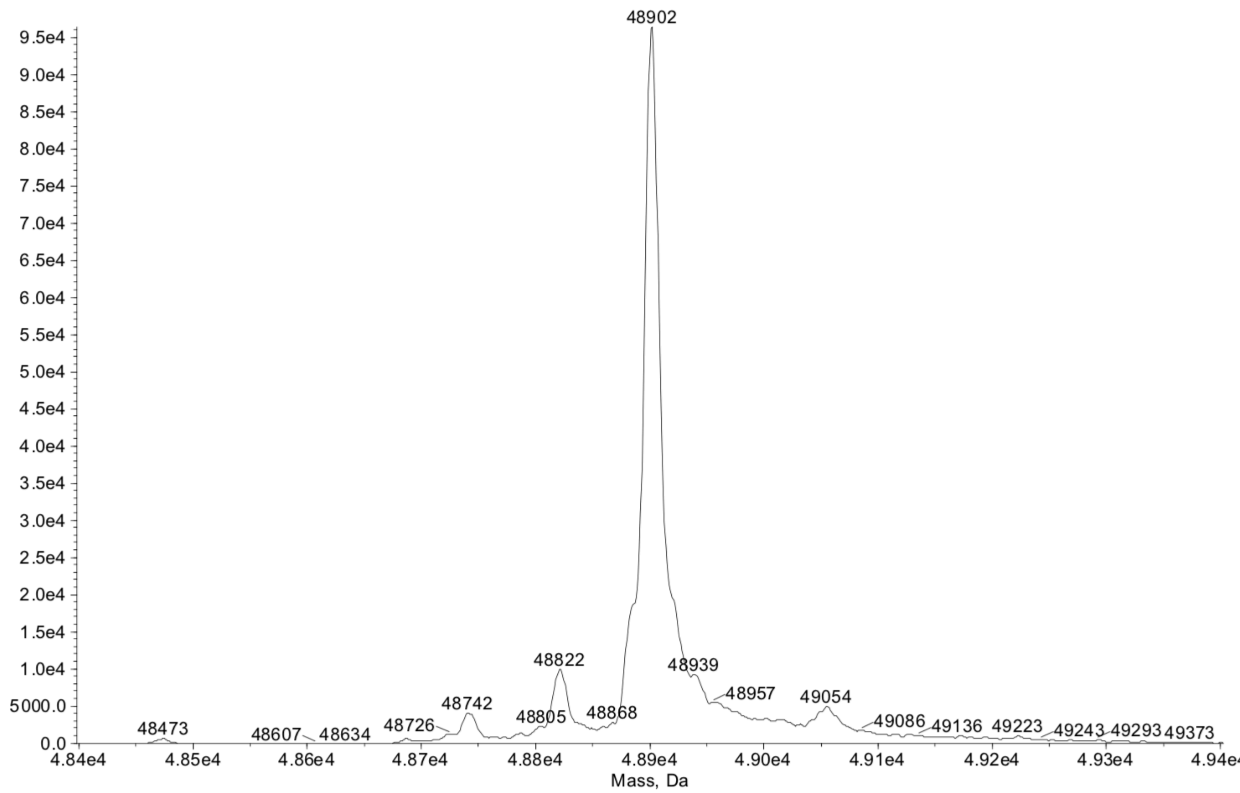
Mass spectrum of E51 sulfoform **1** (0 sY)



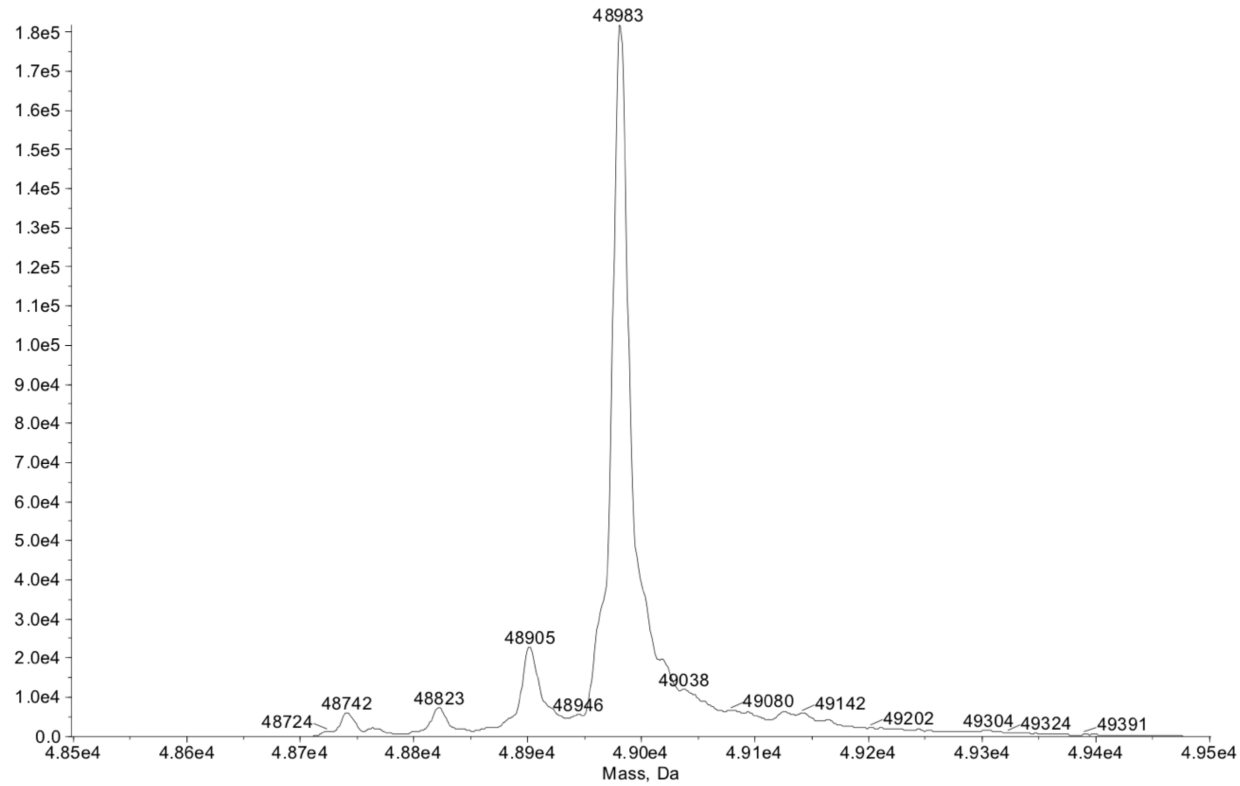
Mass spectrum of E51 sulfoform **2** (1 sY)



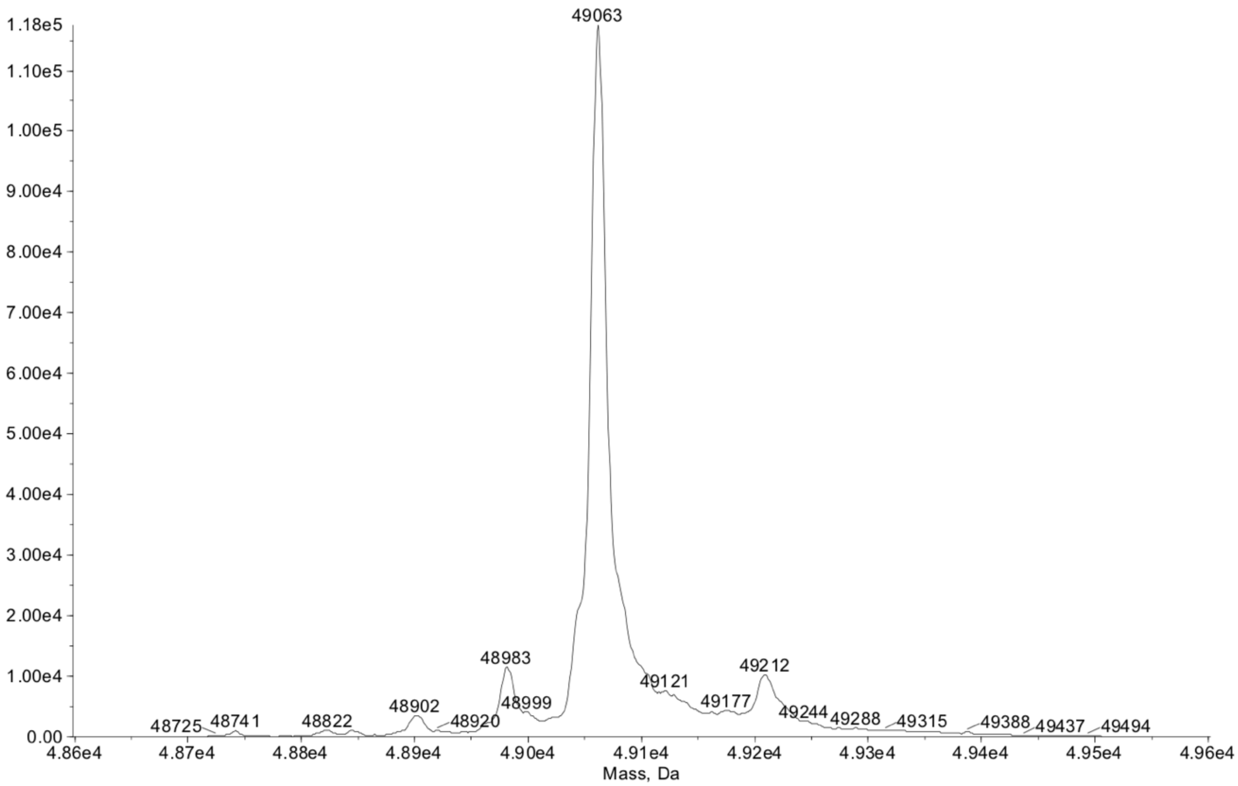
Mass spectrum of E51 sulfoform **7** (2 sYs)



Mass spectrum of E51 sulfoform **17** (3 sYs)



Mass spectrum of E51 sulfoform **27** (4 sYs)



Mass spectrum of E51 sulfoform **32** (5 sYs)

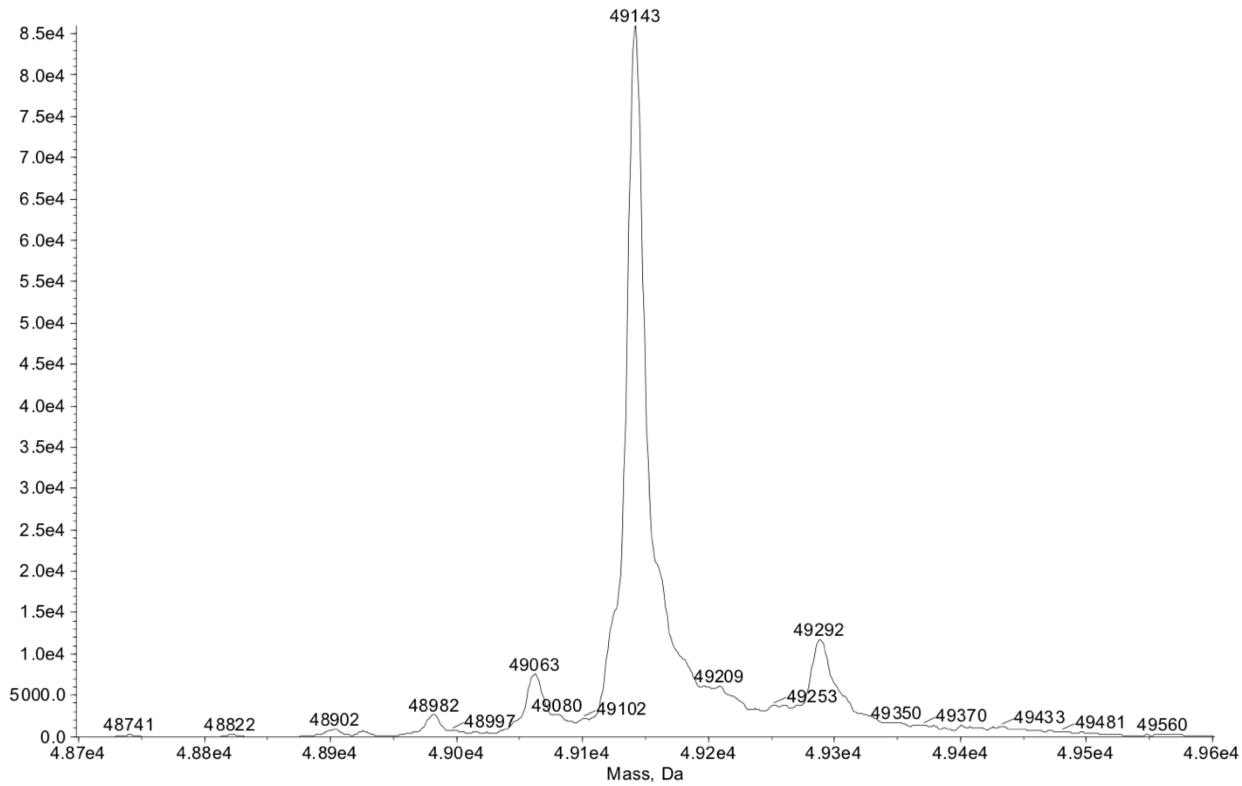


Figure S4. Representative mass spectra of E51 Fabs containing 0-5 sYs.

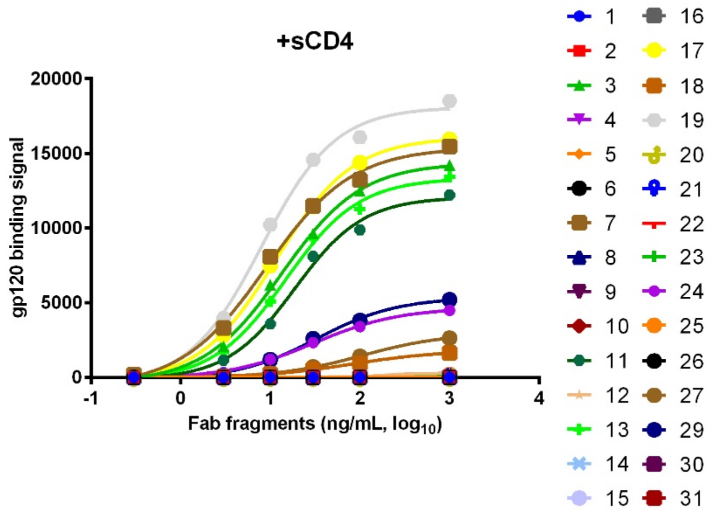


Figure S5. ELISA on 30 out of 32 E51 sulfoforms against gp120 with sCD4. E51 sulfoforms that did not exhibit detectable binding were not tested in Figure 3B.

pGLO-E51:

atcgatgcataatgtgcctgtcaaatggacgaagcagggattctgcaaaccctatgctactccgtcaagccgtcaattgtctgattcgttacca
attatgacaacttgacggctacatcattcactttttcttcacaaccggcacggaactcgctcgggctggccccggtgcattttttaaatacccgc
gagaaatagagttgatcgtcaaaaccaacattgcgaccgacggtggcgataggcatccgggtggtgctcaaaagcagcttcgcctggctg
atacgttggtcctcgcgccagcttaagacgctaatccctaactgctggcggaaaagatgtgacagacgcgacggcgacaagcaaacatgc
tgtgcgacgctggcgatatcaaaattgctgtctgccaggtgatcgctgatgtactgacaagcctcgcgtacccgattatccatcggtggatgg
agcgactcgttaatcgcttccatgcgccgcagtaacaattgctcaagcagatttatcgccagcagctccgaatagcgcccttccccttgcccg
gcgttaatgatttgcccaaacaggtcgctgaaatgcggctggtgcgcttcatccgggcgaaagaaccccgtattggcaaatattgacggcca
gttaagccattcatgccagtaggcgcgcggacgaaagtaaacccactggtgataccattcgcgagcctccggatgacgaccgtagtgatg
aatctctcctggcgggaacagcaaaatatcacccggtcggcaaacaaattctcgtccctgatttttcaccaccccctgaccgcgaatggtga
gattgagaatataacctttcattcccagcggtcggtcgataaaaaaatcgagataaccgttggcctcaatcggcgttaaacccgccaccagat
gggcattaaacgagtatcccggcagcaggggatcattttgcgcttcagccatacttttcatactcccgccattcagagaagaaaccaattgtc
catattgcatcagacattgccgtcactgcgtcttttactggctcttctcgctaaccaaaccggtaaccccgcttattaaaagcattctgtaacaaa
gcgggaccaaagccatgacaaaaacgcgtaacaaaagtgtctataatcacggcagaaaagtccacattgattatttgcacggcgtcacactt
tgctatgccatagcatttttatccataagattagcggatcctacctgacgctttttatcgcaactctctactgtttctccatacccgtttttttgggcta
gaaataattttgtttaactttaagaaggagatatacatatgaagaagaccgcgattgcgattgcggtggcgctggcgggctttgcgaccgtgg
cgcaggcgcagtctgtgttgacgcagccgccctcagtgtctgcggccccaggacagaaggtcaccatctcctgctctggaagcagctcca
acattgggaataatgatgtttcctggtaccagcagttcccaggaacagtccccaaactcgtcatctatgaaaataatcagcgaccctcaggga
ttcctgaccgcttctctggctccaagtctggcacgtcagccaccctgggcatcaccggactccagactggggacgaggccgattattactgc
ggcacatgggatagcagcctgagtgctgtggtcttcggcggagggtccaaggtgaccgtcctaggacagccgaaagcggcgccgagcg
tgaccctgtttccgccgagcagcgaagaactgcaggcgaacaaagcgaccctggtgtgcctgattagcgatttttatccgggcgcggtgac
cgtggcgtggaaagcggatagcagcccggtgaaagcgggcgtggaaaccaccaccccgagcaaacagagcaacaacaaatatgcgg
cgagcagctatctgagcctgaccccggaacagtggaaaagccatcgcagctatagctgccaggtgacccatgaaggcagcaccgtggaa
aaaaccgtggcgccgaccgaatgctgatattttaataattaaaaggagatataccatgaaatacctattgcctacggcagccgctggcttgct
gctgctggcagctcagccggccatggctcaggtgcagctggtgcagtctggggctgaggtgaacaagcctgggtcctcggtgaaggtctc
ctgtcaggcttctggagccaccttaaacagtcatgctttcagctgggtgcgacaggcccctggacaagggcttgagtggatggcagggatc
atccctatatttggttcatcacactacgcacaaaagttccagggcagagtcacgattaccgcggacgagtccacgcgcacagtgtatttgcat
ctgagaggcctgagatctgatgacacggccgtttattactgtgcgagtaattctattgccggagtagcagctgccggagactacgctgactac
gatgggggctactactacgatatggatgtctggggccaagggacaacggtcaccgtctcctcagcctcgaccaagggcccatcggtcttcc
ccctggcaccctcctccaagagcacctctgggggcacagcggccctgggctgcctggtcaaggactacttccccgaaccggtgacggtgt
cgtggaactcaggcgccctgaccagcggcgtgcacaccttcccggctgtcctacagtcctcaggactctactccctcagcagcgtggtgac
cgtgccctccagcagcttgggcacccagacctacatctgcaacgtgaatcacaagcccagcaacaccaaggtggacaagaaagttgagc
ccaaatcttgtggtggcggctcctatccgtatgatgtgccggattatgcgtgatgaattcgagctcggtacccggggatcctctagagtcgac
ctgcaggcatgcaagcttggctgttttggcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctgataaa
acagaatttgcctggcggcagtagcgcggtggtcccacctgaccccatgccgaactcagaagtgaaacgccgtagcgccgatggtagtgt
ggggtccccccatgcgagagtagggaactgccaggcatcaaataaaacgaaaggctcagtgcaaagactgggcctttcgttttatctgttgttt
gtcggtgaacgctctcctgagtaggacaaatccgccgggagcggatttgaacgttgcgaagcaacggcccggagggtggcgggcagga
cgcccgccataaactgccaggcatcaaattaagcagaaggccatcctgacggatggcctttttgcgtttctacaaactctttgtttatttttctaa
atacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgt
gtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtg
cacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaa
agttctgctatgtggcgcggtattatcccgtgttgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttga
gtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcgg
ccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttggg
aaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgcagcaatggcaacaacgttgcgcaaactattaactg

gcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccg
gctggctggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctcccg
tatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcatt
ggtaactgtcagaccaagtttactcatatatactttagattgatttacgcgccctgtagcggcgcattaagcgcggcgggtgtggtggttacgc
gcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaa
gctctaaatcgggggctccctttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgatttgggtgatggttcacgtagtgg
gccatcgccctgatagacggtttttcgccctttgacgttggagtccacgttctttaatagtggactcttgttccaaactggaacaacactcaacc
ctatctcgggctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttta
acaaaatattaacgtttacaatttaaaaggatctaggtgaagatcctttttgataatctcatgaccaaaatcccttaacgtgagttttcgttccactg
agcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgcta
ccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttc
tagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccag
tggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacggggggttcgtgcac
acagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaagcgccacgcttcccgaagggaga
aaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttata
gtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcg
gcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagt
gagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgcggtattttct
ccttacgcatctgtgcggtatttcacaccgcatatggtgcactctcagtacaatctgctctgatgccgcatagttaagccagtatacactccgct
atcgctacgtgactgggtcatggctgcgccccgacacccgccaacacccgctgacgcgccctgacgggcttgtctgctcccggcatccgc
ttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcgaggcagcaaggagatgg
cgcccaacagtcccccggccacggggcctgccaccatacccacgccgaaacaagcgctcatgagcccgaagtggcgagcccgatcttc
cccatcggtgatgtcggcgatataggcgccagcaaccgcacctgtggcgccggtgatgccggccacgatgcgtccggcgtagaggatct
aattctcatgtttgacagcttatc

pULTRA-sY:

accaccctgaattgactctcttccgggcgctatcatgccataccgcgaaaggttttgcgccattcgatggtgtccgggatctcgacgctctccc
ttatgcgactcctgcattagggagctgttgacaattaatcatcggctcgtataatgtgtggaattgtgagcggataacaatttcacaaaggaggt
gcggccgcatggacgaatttgaaatgataaagagaaacacatctgaaattatcagcgaggaagagttaagagaggttttaaaaaaagatga
aaaatctgctctgataggttttgaaccaagtggtaaaatacatttagggcattatctccaaataaaaaagatgattgatttacaaaatgctggattt
gatataattatacctttggctgatttacatgcctatttaaaccagaaaggagagttggatgagattagaaaaataggagattataacaaaaaagt
ttttgaagcaatggggttaaaggcaaaatatgtttatggaagtgaattccagcttgataaggattatacactgaatgtctatagattggctttaaaa
actaccttaaaaagagcaagaaggagtatggaacttatagcaagagaggatgaaaatccaaaggttgctgaagttatctatccaataatgca
ggttaatggttgtcattataagggcgttgatgttgcagttggagggatggagcagagaaaaatacacatgttagcaagggagcttttaccaaa
aaaggttgtttgtattcacaaccctgtcttaacgggtttggatggagaaggaaagatgagttcttcaaaagggaattttatagctgttgatgactc
tccagaagagattagggctaagataaagaaagcatactgcccagctggagttgttgaaggaaatccaataatggagatagctaaatacttcc
ttgaatatcctttaaccataaaaaggccagaaaaatttggtggagatttgacagttaatagctatgaggagttagagagtttatttaaaaataagg
aattgcatccaatggatttaaaaaatgctgtagctgaagaacttataaagattttagagccaattagaaagagattataagcggccgcgtttaaa
cggtctccagcttggctgttttggcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctgataaaacaga
atttgcctggcggcagtagcgcggtggtcccacctgaccccatgccgaactcagaagtgaaacgccgtagcgccgatggtagtgtggggt
ctccccatgcgagagtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggccttgtttgtgagctcccggtcat
caatcatccccataatccttgttagcctgcaggtaattccgcttcgcaacatgtgagcaccggtttattgactaccggaagcagtgtgaccgtg
tgcttctcaaatgcctgaggccagtttgctcaggctctccccgtggaggtaataattgacgatatgatcagtgcacggctaactaagcggcct
gctgactttctcgccgatcaaaaggcattttgctattaagggattgacgagggcgtatctgcgcagtaagatgcgccccgcattccggcggta
gttcagcagggcagaacggcggactctaaatccgcatggcaggggttcaaatcccctccgccggaccaaattcgaaaagcctgctcaacg

agcaggctttttgcatgctcgagcagctcagggtcgaatttgccatggcggccaccaggtaccaccggcgcctcaggcatttgagaagca
cacggtcacactgcttccggtagtcaataaaccggtaaaccagcaatagacataagcggctatttaacgaccctgccctgaaccgacgacc
gggtcatcgtggccggatcttgcggcccctcggcttgaacgaattgttagacattatttgccgactaccttggtgatctcgcctttcacgtagtg
gacaaattcttccaactgatctgcgcgcgaggccaagcgatcttcttcttgtccaagataagcctgtctagcttcaagtatgacgggctgatac
tgggccggcaggcgctccattgcccagtcggcagcgacatccttcggcgcgattttgccggttactgcgctgtaccaaatgcgggacaac
gtaagcactacatttcgctcatcgccagcccagtcgggcggcgagttccatagcgttaaggtttcatttagcgcctcaaatagatcctgttcag
gaaccggatcaaagagttcctccgccgctggacctaccaaggcaacgctatgttctcttgcttttgtcagcaagatagccagatcaatgtcga
tcgtggctggctcgaagatacctgcaagaatgtcattgcgctgccattctccaaattgcagttcgcgcttagctggataacgccacggaatga
tgtcgtcgtgcacaacaatggtgacttctacagcgcggagaatctcgctctctccaggggaagccgaagtttccaaaaggtcgttgatcaaa
gctcgccgcgttgtttcatcaagccttacggtcaccgtaaccagcaaatcaatatcactgtgtggcttcaggccgccatccactgcggagcc
gtacaaatgtacggccagcaacgtcggttcgagatggcgctcgatgacgccaactacctctgatagttgagtcgatacttcggcgatcaccg
cttccctcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaaca
aatagctagctcactcggtcgctacgctccgggcgtgagactgcggcgggcgctgcggacacatacaaagttacccacagattccgtgga
taagcaggggactaacatgtgaggcaaaacagcagggccgcgccggtggcgtttttccataggctccgccctcctgccagagttcacata
aacagacgcttttccggtgcatctgtgggagccgtgaggctcaaccatgaatctgacagtacgggcgaaacccgacaggacttaaagatcc
ccaccgtttccggcgggtcgctccctcttgcgctctcctgttccgaccctgccgtttaccggatacctgttccgcctttctcccttacgggaagt
gtggcgctttctcatagctcacacactggtatctcggctcggtgtaggtcgttcgctccaagctgggctgtaagcaagaactccccgttcagc
ccgactgctgcgccttatccggtaactgttcacttgagtccaacccggaaaagcacggtaaaacgccactggcagcagccattggtaactg
ggagttcgcagaggatttgtttagctaaacacgcggttgctcttgaagtgtgcgccaaagtccggctacactggaaggacagatttggttgct
gtgctctgcgaaagccagttaccacggttaagcagttccccaactgacttaaccttcgatcaaaccacctccccaggtggtttttttcgtttacag
ggcaaaagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctactgaaccgctctagatttcagtgcaatttatctcttcaa
atgtagcacctgaagtcagccccatacgatataagttgtaattctcatgttagtcatgccccgcgcccaccggaaggagctgactgggttgaa
ggctctcaagggcatcggtcgagatcccggtgcctaatgagtgagctaacttacattaattgcgttgcgctcactgcccgctttccagtcggg
aaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttgcgtattgggcgccagggtggtttttcttttcacca
gtgagacgggcaacagctgattgcccttcaccgcctggccctgagagagttgcagcaagcggtccacgctggtttgccccagcaggcga
aaatcctgtttgatggtggttaacggcgggatataacatgagctgtcttcggtatcgtcgtatcccactaccgagatgtccgcaccaacgcgc
agcccggactcggtaatggcgcgcattgcgcccagcgccatctgatcgttggcaaccagcatcgcagtgggaacgatgccctcattcagc
atttgcatggtttgttgaaaaccggacatggcactccagtcgccttcccgttccgctatcggctgaatttgattgcgagtgagatatttatgccag
ccagccagacgcagacgcgccgagacagaacttaatgggcccgctaacagcgcgatttgctggtgacccaatgcgaccagatgctccac
gcccagtcgcgtaccgtcttcatgggagaaaataatactgttgatgggtgtctggtcagagacatcaagaaataacgccggaacattagtgc
aggcagcttccacagcaatggcatcctggtcatccagcggatagttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgc
cgctttacaggcttcgacgccgcttcgttctaccatcgacaccaccacgctggcacccagttgatcggcgcgagatttaatcgccgcgacaa
tttgcgacggcgcgtgcagggccagactggaggtggcaacgccaatcagcaacgactgtttgcccgccagttgttgtgccacgcggttgg
gaatgtaattcagctccgccatcgccgcttccactttttcccgcgttttcgcagaaacgtggctggcctggttcaccacgcgggaaacggtct
gataagagacaccggcatactctgcgacatcgtataacgttactggtttcacattc

pXL-shRNA1:

gacgtcgctagctgtacaaaaaagcaggctttaaaggaaccaattcagtcgactggatccggtaccaaggtcgggcaggaagagggccta
tttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaa
tacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttc
gatttcttggctttatatatcttgtggaaaggacgaaacaccgggccatgctggacgcacatcttcgatagagatgtgcgtccagcatggccct
tttttaagcttgggccgctcgaggtacctctctacatatgacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgt
tgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactata
aagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcg
ggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccc

gttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactgg
taacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtatttg
gtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttt
tgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaa
ctcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatat
atgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccc
cgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagat
ttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccg
ggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatgg
cttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgtt
gtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtga
ctggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgcca
catagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgt
aacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaa
agggaataagggcgacacggaaatgttgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggata
catatttgaatgtatttagaaaaataaacaaataggggttccgcgcacatttccccgaaaagtgccacct

pXL-shRNA2:

gacgtcgctagctgtacaaaaaagcaggctttaaaggaaccaattcagtcgactggatccggtaccaaggtcgggcaggaagagggccta
tttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattagaattaatttgactgtaaacacaaagatattagtacaaaa
tacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgttttaaaatggactatcatatgcttaccgtaacttgaaagtatttc
gatttcttggctttatatatcttgtggaaaggacgaaacaccgggacagcaggtgctagagtgttcgatagacactctagcacctgctgtccct
tttttaagcttgggccgctcgaggtacctctctacatatgacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgt
tgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactata
aagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcg
ggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccc
gttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactgg
taacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaagaacagtatttg
gtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttt
tgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaa
ctcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatat
atgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccc
cgtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagat
ttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccg
ggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatgg
cttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgtt
gtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtga
ctggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgcca
catagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgt
aacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaa
agggaataagggcgacacggaaatgttgaatactcatactcttcctttttcaatattattgaagcatttatcagggttattgtctcatgagcggata
catatttgaatgtatttagaaaaataaacaaataggggttccgcgcacatttccccgaaaagtgccacct

Figure S6. Plasmid sequences.

Sequence	Name	EC_{50} (nM)	C.I. (nM)
YYYYY	**3**	1.854	1.352-2.543
YYYYY	**7**	0.883	0.694-1.124
YYYYY	**11**	2.218	1.956-2.516
YYYYY	**13**	1.735	1.403-2.144
YYYYY	**17**	0.947	0.834-1.075
YYYYY	**18**	5.980	5.362-6.669
YYYYY	**19**	1.375	1.145-1.651
YYYYY	**24**	2.000	1.379-2.902
YYYYY	**27**	2.231	1.783-2.789
YYYYY	**28**	0.899	0.666-1.212
YYYYY	**29**	2.209	1.699-2.426
YYYYY	**32**	2.584	1.898-3.517
	E51-TPSTs	0.934	0.815-1.071
	E51-shRNAs	3.657	3.275-4.082

Table S1. EC_{50}s of E51 sulfoforms, E51-TPSTs, and E51-shRNAs based on interpolation of ELISA binding curves in Figure 3B. C.I. = confidence interval. Red "Y" denotes sY.