

Supporting Information

Gilbreth et al. 10.1073/pnas.1102294108

SI Materials and Methods

NMR Epitope Mapping. Uniformly ^{15}N -labeled yeast homolog of small ubiquitin-related modifier (ySUMO) and human SUMO1 (hSUMO1) were produced by culturing BL21(DE3) cells harboring a derivative of the pHFT2 expression vector containing the ySUMO or hSUMO1 gene in M9 media with $^{15}\text{NH}_4\text{Cl}$ as the sole nitrogen source. pHFT2 is a pHFT1 (1) derivative containing a 10-His-tag instead of 6-His. A hSUMO1 mutant was used containing the C52A mutation. Protein expression was induced by the addition of 1 mM IPTG. Proteins were purified using a Ni-Sepharose column (GE Healthcare). After cleaving the N-terminal tag sequence with tobacco etch virus (TEV) protease, the proteins were concentrated and dissolved in 50 mM phosphate, 100 mM NaCl, pH 6.5. ^1H , ^{15}N heteronuclear single quantum coherence spectra were collected on a Varian INOVA 600 NMR spectrometer using pulse sequences provided by the manufacturer. All ySUMO spectra were recorded at 20 °C. All hSUMO1 spectra were recorded at 17 °C. ySUMO resonances were assigned using previously reported assignments by Sheng and Liao (2). hSUMO1 resonances were assigned using previously reported assignments by Macauley et al (3). Spectra were collected for the free [^{15}N]-ySUMO (380 μM), free [^{15}N]-hSUMO1 (228 μM), [^{15}N]-ySUMO (100 μM) in complex with unlabeled monobody (200 μM) and [^{15}N]-hSUMO1 (242 μM) in complex with unlabeled monobody (484 μM) in the above buffer. Residues affected by monobody binding were identified by comparing the free and monobody bound spectra. Amide cross-peaks were classified into five categories: strongly affected (a shift of greater than two peak widths), moderately affected (a shift of between one and two peak widths), weakly affected (a shift of approximately 1 peak width), unaffected (a shift of less than one peak width), and excluded (resonances that could not be unambiguously assigned) (4, 5).

Phage Enzyme-Linked Immunosorbent Assay (ELISA). For phage amplification, *Escherichia coli* XL1-blue cells transformed with the *Lacl* containing plasmid pMCSG21 (6) (termed “XL21” hereafter) were used. Monobody-displaying phage particles were prepared by growing XL21 cells transfected with the phagemids of individual clones in the presence of 0.2 mM IPTG and helper phage K07 (7, 8). Cultures were then centrifuged and the phage-containing supernatant was used for ELISA assays. All incubations were at room temperature. In all instances except for the phage titration experiment used to test hSUMO1-binding monobody specificity (Fig. 5), wells of a 96-well Microolon® (Greiner) ELISA plate were treated with a 2 $\mu\text{g}/\text{mL}$ solution of a GST-fusion of the appropriate target protein, or GST alone in 50 mM Tris Cl buffer containing 150 mM NaCl, pH 7.5 (TBS) and incubated for 1 h followed by blocking with 0.5% BSA in TBS for 1 h. In the hSUMO1 binder specificity experiment, 2 $\mu\text{g}/\text{mL}$ neutravidin in TBS was coated, followed by blocking with 0.5% BSA and an addition of a 50 nM solution of His-tagged ySUMO, hSUMO1, or hSUMO2 in complex with the biotin-Tris nitrilotriacetate (BT-Tris-NTA) compound, which noncovalently links a biotin moiety to a His-tag (9, 10), and incubated for 30 min. In epitope mapping competition experiments, wells coated with GST-target were then incubated with either 1 μM of the ySMB-1 monobody, or 1 μM of the hS1MB-4 monobody in TBS or TBS only for 1 h. In other experiments, this step was not performed. After washing the wells with TBS + 0.1% Tween 20 (TBST), 50 μL of a 30% solution of phage supernatant in TBS + 0.5% BSA was added to the wells and incubated for 30 min. In the phage titration experi-

ment, serial fivefold dilutions of this 30% solution were also tested. In competition experiments, 1 μM ySMB-1 or 1 μM hS1MB-4 was included in the binding mixture. Bound phages were then detected using an anti-M13 antibody conjugated to horseradish peroxidase (GE Healthcare) in conjunction with the Ultra 3,3',5,5'-tetramethylbenzidine (TMB) ELISA colorimetric substrate (Pierce). Reactions were quenched after 5 min by addition of H_2SO_4 and phage binding quantified by absorbance measured at 450 nm.

Protein Expression and Purification. GST-fusion proteins were produced by cloning genes into a previously described vector (11). All other proteins were expressed by cloning genes into the pHFT2 vector. Unless otherwise noted, all proteins were expressed by growing BL21(DE3) cells harboring the appropriate pHFT2 vector in ZYP-5052 autoinduction media according to the methods of Studier (12). Proteins were purified using Ni-Sepharose columns (GE Healthcare), or His-Mag magnetic particles (Novagen) in conjunction with a Kingfisher instrument (Thermo).

Surface Plasmon Resonance. Monobody proteins purified as described above were immobilized via His-tag to an NTA surface using a BIAcore 2000 instrument so that the theoretical maximum response (R_{max}) from target binding was 100–200 response units. Target protein at varying concentrations was then flowed over the surface at a flow rate of 30 $\mu\text{L}/\text{min}$ and the binding signal recorded. Kinetic fitting of traces was carried out using the BIAevaluation software. For equilibrium experiments, the equilibrium binding response was recorded for multiple target concentrations and fit to a simple 1:1 saturation binding curve.

X-Ray Crystallography. ySMB-1 and ySUMO proteins were expressed and purified as described above. After removal of the tag sequence with TEV protease, the two proteins were mixed in a 1:1 molar ratio, concentrated to a total protein concentration of 4.9 mg/mL, and dissolved in 10 mM Tris, 50 mM NaCl, pH 8.0. The formation and monodispersity of the complex was asserted by gel filtration. The ySMB-1/ySUMO complex was crystallized in 14% PEG 8000, 16% glycerol at 19 °C using the hanging drop vapor diffusion method. Crystals were frozen in a mixture of 80% mother liquor and 20% glycerol as a cryoprotectant. Diffraction data were collected at Advanced Photon Source beamline 21-ID-F (Argonne National Laboratory). Crystal and data collection information are reported in Table S2. X-ray diffraction data were processed and scaled with HKL2000 (13). The structure was determined by molecular replacement using sequential search with two different models with the program MOLREP in CCP4 (14). The ySUMO structure (residues 1013–1098 of chain C, PDB ID code 2EKE) was used as a search model, along with the fibronectin type III domain (FN3) structure with the variable loop regions deleted (PDB ID code 1FNA) (15, 16). Rigid body refinement was carried out with REFMAC5 (17). Model building and the search for water molecules was carried out using the Coot program (18). Simulated annealing was performed in CNS1.1 (19). The Translation/Libration/Screw and bulk solvent parameters, restrained temperature factor and final positional refinement were completed with REFMAC5 (17). Molecular graphics were generated using PyMOL (www.pymol.org).

Design of the SUMO-Targeted Monobody Library. Choice of positions and diversity in the SUMO-targeted library carried the following rationale. All residues of the FG loop were varied except one, S77

that did not contact γ SUMO in the γ SMB-1/ γ SUMO crystal structure and did not appear to be capable of direct participation in any similar interface. We varied Y76 to D, H, N, and Y, because, although it did not directly contact γ SUMO in the γ SMB-1 interface, we suspected that this position may be capable of interacting with the conserved R55 in all SUMOs (Fig. 44). Leucine 81 of γ SMB-1 is buried in a pocket in the γ SUMO surface that is conserved across all SUMO isoforms, and an equivalent "anchor" leucine or valine is conserved in all SUMO-interacting motif (SIM)/SUMO complexes for which there are structures. As a result, we restricted the amino acid diversity at this position to F, L, I, and V. E47 and S86 of the FN3 scaffold made very minimal contact in the γ SMB-1 interface and were not varied. Though P87 of the scaffold did make significant contact in the γ SMB-1 interface, it was held constant to avoid perturbation of the turn structure it introduces which would likely change the overall positioning of the FG loop.

Phage Display Selection. For use in selection, γ SUMO, hSUMO-1, and hSUMO-2 were expressed as a C-terminal fusion to an engineered GST variant devoid of cysteine residues (C to S mutations) except for a single cysteine near the N terminus (11). In the case of hSUMO1 a C52A mutant was used and in the case of hSUMO2 a C53S mutant was used. The GST-fusion targets were modified with a redox cleavable biotin moiety using EZ-link Biotin HPDP (Pierce). Monobody-displaying phage particles were prepared by growing XL21 cells transfected with the phagemid library in the presence of 0.2 mM IPTG and helper phage K07 (8, 20). In the first round of library selection, 50 nM biotinylated GST target was mixed with a sufficient amount of streptavidin-conjugated magnetic beads (Streptavidin MagneSphere Paramagnetic Particles; Promega, Z5481/2) in TBST. Beads were blocked with a 5- μ M solution of biotin in TBST. To this target solution, 10^{11-12} phage particles suspended in 0.5 mL TBST + 0.5% BSA were added, and the solution was mixed and incubated for 15 min at room temperature. After washing the beads twice with TBST, the bead suspension containing bound phages were added to a fresh XL21 culture. Phages were amplified as described before (8). In the second round, phages were preincubated in TBST + 0.5% BSA with 500 nM unbiotinylated GST to remove GST binders from the population. Target binding phages were then captured by streptavidin-conjugated magnetic beads

loaded with 10 nM GST target. Phages bound to the target protein were eluted from the beads by cleaving the linker within the biotinylation reagent with 100 mM DTT in 50 mM Tris pH 8.0. The phagemids were washed and recovered as described above. After amplification, the third and fourth rounds of selection were performed using 1 and 0.1 nM target, respectively.

Monobody Effects on SUMO/SIM Interactions. Wells of a Microton[®] (Greiner) ELISA plate were coated with 2 μ g/mL GST-RanBP2 for 1 hour at room temperature. This IR1-M-IR2 construct of Ran-GTP binding protein 2 (RanBP2) has been described previously (21). A complex was preformed between His-tagged SUMOylated Ran GTPase activating protein (RanGAP) (modified with SUMO1) and the BT-Tris-NTA reagent which noncovalently attaches a biotin moiety to a His-tag (9, 10). This complex was incubated with varying concentrations of hS1MB-4 for 1 h and then the mixture was added to the ELISA plate and incubated for 30 min. Bound SUMO-RanGAP was then detected using a streptavidin-horseradish peroxidase conjugate in conjunction with the Ultra TMB ELISA reagent (Pierce). The reaction was quenched with 2 M H₂SO₄, and the absorbance at 450 nm was measured.

Monobody Effects on SUMOylation. A mixture of hSUMO1 and His₆-SUMO3 (24 μ M each) was combined with either monobody hS1MB-4 or hS1MB-5 at varying concentrations and incubated for 1 h. A mixture of E1 [SUMO activating enzyme subunit 1/2 (SAE1/2), 1.7 μ M], E2 (Ubc9, 13.7 μ M), and ATP (5.5 mM) was then added and the SUMOylation reaction allowed to proceed for 10 min at 37 °C. The reaction was then quenched by an addition of SDS-PAGE loading dye and reaction mixture was analyzed by SDS-PAGE.

Monobody Effects on DeSUMOylation. YFP-hSUMO1-enhanced cyan fluorescent protein fusion (63 μ g/mL) was mixed with varying concentrations of monobody hS1MB-4, hS1MB-5, or γ SMB-1 as a control and incubated at room temperature for 30 min. sentrin(SUMO)-specific protease 1 was then added at a final concentration of 32 nM and the mixture incubated for 15 min at 37 °C. The reaction was stopped by putting the reaction containers on ice, adding SDS-PAGE sample buffer, and then boiling for 5 min. The reaction mixture was then analyzed by SDS-PAGE.

- Huang J, Koide A, Nettle KW, Greene GL, Koide S (2006) Conformation-specific affinity purification of proteins using engineered binding proteins: Application to the estrogen receptor. *Protein Expr Purif* 47:348–354.
- Sheng W, Liao X (2002) Solution structure of a yeast ubiquitin-like protein Smt3: The role of structurally less defined sequences in protein-protein recognitions. *Protein Sci* 11:1482–1491.
- Macauley MS, et al. (2004) Structural and dynamic independence of isopeptide-linked RanGAP1 and SUMO-1. *J Biol Chem* 279:49131–49137.
- Farmer BT, II, et al. (1996) Localizing the NADP+ binding site on the MurB enzyme by NMR. *Nat Struct Biol* 3:995–997.
- Huang X, Yang X, Luft BJ, Koide S (1998) NMR identification of epitopes of Lyme disease antigen OspA to monoclonal antibodies. *J Mol Biol* 281:61–67.
- Stols L, et al. (2007) New vectors for co-expression of proteins: Structure of Bacillus subtilis ScoAB obtained by high-throughput protocols. *Protein Expr Purif* 53:396–403.
- Koide A, Gilbreth RN, Esaki K, Tereshko V, Koide S (2007) High-affinity single-domain binding proteins with a binary-code interface. *Proc Natl Acad Sci USA* 104:6632–6637.
- Sidhu SS, Lowman HB, Cunningham BC, Wells JA (2000) Phage display for selection of novel binding peptides. *Methods Enzymol* 328:333–363.
- Koide A, et al. (2009) Accelerating phage-display library selection by reversible and site-specific biotinylation. *Protein Eng Des Sel* 22:685–690.
- Reichel A, et al. (2007) Noncovalent, site-specific biotinylation of histidine-tagged proteins. *Anal Chem* 79:8590–8600.
- Wojcik J, et al. (2010) A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain. *Nat Struct Mol Biol* 17:519–527.
- Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41:207–234.
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
- Anon (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr Sec D Biol Crystallogr* 50:760–763.
- Dickinson CD, et al. (1994) Crystal structure of the tenth type III cell adhesion module of human fibronectin. *J Mol Biol* 236:1079–1092.
- Duda DM, et al. (2007) Structure of a SUMO-binding-motif mimic bound to Smt3p-Ubc9p: Conservation of a non-covalent ubiquitin-like protein-E2 complex as a platform for selective interactions within a SUMO pathway. *J Mol Biol* 369:619–630.
- Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr Sec D Biol Crystallogr* 53:240–255.
- Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr Sec D Biol Crystallogr* 60:2126–2132.
- Brunger AT, et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr Sec D Biol Crystallogr* 54:905–921.
- Koide A, et al. (2007) Identification of regions within the F domain of the human estrogen receptor alpha that are important for modulating transactivation and protein-protein interactions. *Mol Endocrinol* 21:829–842.
- Tatham MH, et al. (2005) Unique binding interactions among Ubc9, SUMO and RanBP2 reveal a mechanism for SUMO paralog selection. *Nat Struct Mol Biol* 12:67–74.

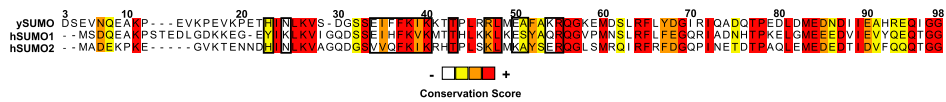


Fig. S1. Sequence alignment of ySUMO and hSUMOs. Residues are colored by conservation score among ySUMO and hSUMO isoforms. Conservation score reflects conservation of a residue's chemical properties and structure (1–4). Residues located in the SIM binding site are boxed in black.

- 1 Livingstone CD, Barton GJ (1993) Protein sequence alignments: A strategy for the hierarchical analysis of residue conservation. *Comput Appl Biosci* 9:745–756.
- 2 Clamp M, Cuff J, Searle SM, Barton GJ (2004) The Jalview Java alignment editor. *Bioinformatics* 20:426–427.
- 3 Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948.
- 4 Lopez R, Lloyd A (1997) ClustalW WWW Service at the European Bioinformatics Institute (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>).

Clone	BC Loop	DE Loop	FG Loop	Estimated K_d (nM)	Library
ySMB-1	S S S S V S	G S S S	Y Y S Y Y D L Y Y S Y	82	YSGW
ySMB-2	S S S S V S	G S S S	Y W T Y E W Y M Y D	45	YSGW
ySMB-3	Y S V W D V A	G Y S S	Y P Y Y L Y Y S Q		YSGW
ySMB-4	P W A Y S Q S V A	G S S S	Y Y G T P W E G W Y S W	210	YSGW
ySMB-5	P A N S V S	G S S S	Y E W Y W W T Y	100	YSGW
ySMB-6	S S S S V S	S Y S S	Y W F E Y G Y W S Y	166	YSGW
ySMB-7	S S S S V S	G S S S	Y Q E W S Y G W S S E	800	YSGW
ySMB-8	Y V Y G S S V A	G S S S	Y Y E G D L Y S S M	350	YSGW
ySMB-9	G Y W F I D	G Y S S	Y Y D N Y W	8000	YSGW
ySMB-10	S S S S V S	G S S S	Y Y D Y Y Y G M Y W S Q		YSGW
ySMB-11	Y Q N S S W S G V S	G S S S	L Y S E Y W E Y Y S		YSGW
ySMB-12	S S S S V S	G Y S S	Y T G F Y D Y G F W S Y W		YSGW
ySMB-13	S S S S V S	G S S S	Y Y M E W D Y Y S E		YSGW
ySMB-14	S S S S V S	G S S S	T W L Y D W Y M W Y		YSGW
ySMB-15	G Y Y Q P W Y V D	S S S Y	Y T F I S A S E Y Y Y M		YSGW
ySMB-16	Y G R N Y T M D V S	S S S Y	V V Y E D W S Y W S		YSGW
ySMB-17	S S S S V S	G S S S	Y I Q Y Y D L Y Y A	40	YSGW
ySMB-18	S S S S V S	G S S S	Y L P Y Y D I Y Y S Y		YSGW
ySMB-19	S S S S V S	G S S S	F Y D V W Y Y S I		YSGW
ySMB-20	G Y Y M N Y Y M V D	G S S S	Y W Y D E Y G Q Y W W S E	66	YSGW
ySMB-21	S S S S V S	G S S S	T M Y S Y W E Y Y Y		YSGW
ySMB-22	S S S S V S	G S S S	Y Y S E Y Y Y W W S E		YSGW
ySMB-23	A Y A S S Q S T V H	S S S Y	Y W E Y Q W H M Y		YSGW
ySMB-24	G Y A Y Y G E S V W	G S S S	N Y G D N Q I E Y Y F		YSGW
ySMB-25	S S S S V S	G S S S	Y Q E W L Y W Y S T		YSGW
ySMB-26	S W E Y Y G Y S V D	G S S S	Y Y S W Y G Y W S E		YSGW
ySMB-27	A R Q H Y E D V Q	S S S S	Y H W E F G Y W S Q		YSGW
ySMB-28	Y E S E V S	G S S S	Y W V Y W S Y W K		YSGW
ySMB-29	S S S S V S	G S S S	Y Y D E A Y E W Y S W		YSGW
ySMB-30	S S S S V S	G S S S	Y W Y M W D G Y V S S W Y		YSGW
ySMB-31	S S S S V S	G S S S	Y E F E E Y G W Y S Y Y		YSGW
ySMB-32	S S S S V S	G S S S	Y W E W V P W G Y Y S Y		YSGW
ySMB-33	S S S S V S	G S S S	Y Y S E W Y Y S Y		YSGW
ySMB-34	S S S S V S	G S S S	Y Y W A D W Y Y W S V		YSGW
ySMB-35	Y E S Y Q Y V W	G S S S	E H G Y W Y W		YSGW
ySMB-36	D R Y Y V S	G S S S	E Y E S N Y Y Y K W		YSGW
ySMB-37	S S S S V S	G S S S	L T D Y Y A Y Y S Y		YSGW
ySMB-38	W S R Y W E W V S	G S S S	S S W H P Q S P Y Y		YSGW
ySMB-39	S S S S V S	S S G Y	T W Y P Q W D G Y S V Y A		YSGW
ySMB-40	Y Q T R L Y P Y V S	S S S S	V E E F D W G Y S W S		YSGW
ySMB-41	S S S S V S	G S S S	W F Q Y H W G S Y Y Y		YSGW
ySMB-42	W K Q W T M G V D	G Y S S	W T V M D L Y Y Y Y		YSGW
ySMB-43	E S D Y Y S V Y	G S S S	W M I S E Y W E Y H W S N		YSGW
ySMB-44	S S S S V S	G S S S	W Y P F S E Y G W A S L		YSGW
ySMB-45	S S S S V S	G Y Y S	Y F A Y Y D Y Y S Y		YSGW
ySMB-46	S S S S V S	G S S S	Y F K E D W Y Y S Y		YSGW
ySMB-47	Y F G W S V S	G S S S	Y S S Y Y V P F Y		YSGW
ySMB-48	I M E S G V S	G S S S	Y H L F D P V Y S S D		YSGW
ySMB-49	S S S S V S	S S G Y	Y H S Y E Y S S L Y Y S W		YSGW
ySMB-50	S S S S V S	G S S S	Y H W S P F Y Y S Y		YSGW
ySMB-51	Y G G Y W Y V S	Y Y S S	Y P W Y D D T Y Y		YSGW
ySMB-52	Y Q G G S Y W V S	G S S S	Y Q H P Q M S T Y Y		YSGW
ySMB-53	S S S S V S	G S S S	Y Q S F Y N E Y W S S G	380	YSGW
ySMB-54	G Y D E S V M	G S S S	Y R E Y D L Y Y S Y	950	YSGW
ySMB-55	V W G E W A Y Y V H	S S Y S	Y S D W Y A Y Y E W		YSGW
ySMB-56	S S S S V S	G S S S	Y T D W S E M L S W Y		YSGW
ySMB-57	Q Y H Y D S G S V S	G Y S S	Y V Y H W G I E Y Y		YSGW
ySMB-58	S S S G Y V Q	G S S S	Y W A G Y Y Y Y S Y		YSGW
ySMB-59	S S S S V S	G S Y S	Y W Y G Y Y Y Y S H		YSGW
ySMB-60	Y E S S F Y H V S	G S S S	Y Y A F T Y G Y Y S Y		YSGW
ySMB-61	S S S S V S	G S S S	Y Y D S M Y W W A D		YSGW
ySMB-62	E G Y S M V D	G S Y S	Y Y E M F S D E G W W S W		YSGW
ySMB-63	S S S S V S	G S Y S	Y Y G W E W D G W W S W Q		YSGW
ySMB-64	S Q F S G V S	G S Y S	Y Y H E H W Y Y S Y	280	YSGW
ySMB-65	S S S S V S	G S S S	Y Y Q E Y W Y W S Y		YSGW
ySMB-66	S S S S V S	G S S S	Y Y Q Y A P Y Q W Y S W Q		YSGW
ySMB-67	G S F F H D P V Y	G S S S	Y Y Q Y M D Y Y Y S Y	1000	YSGW
ySMB-68	S S S S V S	G S Y S	Y Y S E Y W Y D F S W		YSGW
ySMB-69	S S S S V S	G S S S	Y Y S T Y W Y D W S G	138	YSGW
ySMB-70	D H E D G L Y D V H	G S S S	Y Y S Y W S Y Y		YSGW
ySMB-71	S S S S V S	G S S S	Y Y V D Y Y Y W S E	230	YSGW
ySMB-72	S S S S V S	G S S S	Y Y W D E Y Y W S I		YSGW
ySMB-73	S S S S V S	S S G Y	Y Y W H D S Y W W S V		YSGW
ySMB-74	S S S S V S	G S S S	Y Y Y F Y D Y W M A Y	544	YSGW
ySMB-75	S S S S V S	G S S S	Y Y Y I Y R D L S Y A E		YSGW
ySMB-76	S S S S V S	G S S S	Y Y Y S Y Y Y W A T		YSGW
ySMB-77	S S S S V S	Y Y G S	Y Y Y D Y Y Y D L Y Y S N		YSX
ySMB-78	S S S S V S	Y Y G T	Y Y Y D Y Y Y D L Y Y S N		YSX
ySMB-79	Y Y S Y Y S Y S V S	Y Y G S	Y Y Y D Y Y Y D L Y Y S N		YSX
ySMB-80	S S S S V S	Y Y G S	Y Y Y D Y Y Y D L Y Y S Y		YSX
ySMB-81	S S S S V S	S Y Y S	Y Y Y S Y Y Y Y Y S Y		YSX
ySMB-82	S S S S V S	G S K S	Y Y Y S Y Y Y Y Y S F		YS
ySMB-83	S S P S V S	G S K S	Y Y Y S Y Y Y Y Y S Y		YS
ySMB-84	S S S S V S	S Y Y S	Y Y Y S Y Y Y Y Y S Y		YS
ySMB-85	S S S S V S	G S K S	Y Y Y S Y Y Y Y Y S Y		YS
ySMB-86	S S S S V S	S S Y S	Y Y Y S Y Y Y Y Y S Y		YS
ySMB-87	S S S S V S	P Y Y S	Y Y Y S Y Y Y Y Y S Y		YS
ySMB-88	S S S S V Y	G S K S	Y Y Y S Y Y Y Y Y S Y		YS
ySMB-89	S S S Y S V S	S S Y S	Y Y Y S Y Y Y Y Y S Y		YS
ySMB-90	S S S S V S	G S S S	Y Y Y S Y Y Y Y Y S Y	320	YS
ySMB-91	S Y S S S V S	G S S S	Y Y Y S Y Y Y Y Y S Y	1000	YS
ySMB-92	Y Y S Y Y S Y S V S	Y Y Y S	Y S S Y S S S S Y		YS
ySMB-93	S S S Y V S	G S K S	Y S S S Y Y S Y S Y		YS
ySMB-94	S S S S V S	G S Y S	Y S S Y Y Y S Y S Y		YS
ySMB-95	S S S S S Y V S	S Y Y S	Y S Y Y Y Y Y S Y		YS

Fig. S2. Sequences and affinities of ySUMO-binding monobodies. Amino acid sequences of the variable loops of all ySUMO-binding monobodies recovered in our laboratory. If available, K_d values from SPR are given. Monobodies originated from one of three libraries: a binary Tyr/Ser library in which loop lengths and sequences were varied using a combination of 50% Y and 50% S (1), a "YSX" library which used a combination of 40% Y, 20% S, 10% G, and 5% each of R, L, H, D, N, A (2), or a "YSGW" library which used a combination of 30% Y, 15% S, 10% G, 5% each of W, F, and R, and 2.5% each of all other amino acids except cysteine in the BC and FG loops and 50% Gly, 25% Tyr, and 25% Ser at position 52, and a 50/50 mixture of Tyr and Ser at positions 53–55 in the DE loop (3). Residues are colored as follows: Tyr (yellow), Ser (red), Gly (green), Arg and Lys (dark blue), His (blue), polar amino acids (C, N, T, Q,) (pale blue), hydrophobic amino acids (A, L, V, F, I, W, M, P) (white), and Asp and Glu (pink), residues originating from the vector template (not mutagenesis) (gray).

- 1 Koide A, Gilbreth RN, Esaki K, Tereshko V, Koide S (2007) High-affinity single-domain binding proteins with a binary-code interface. *Proc Natl Acad Sci USA* 104:6632–6637.
- 2 Gilbreth RN, Esaki K, Koide A, Sidhu SS, Koide S (2008) A dominant conformational role for amino acid diversity in minimalist protein-protein interfaces. *J Mol Biol* 381:407–418.
- 3 Wojcik J, et al. (2010) A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain. *Nat Struct Mol Biol* 17:519–527.

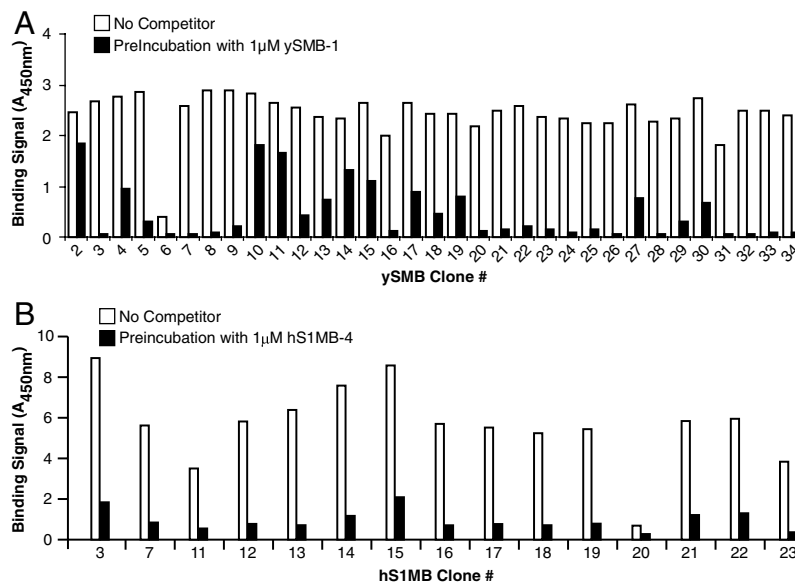


Fig. S3. Epitope mapping ELISA of ySUMO-binding and hSUMO-1 binding monobodies. (A) Binding of 34 phage-displayed ySUMO-binding monobodies measured by ELISA in the presence and absence of 1 μ M ySMB-1 competitor. Clone numbers correspond to those of the format ySMB-X in Table S1. (B) Binding of 16 phage-displayed hSUMO1-binding monobodies to hSUMO1 measured by ELISA in the presence or absence of 1 μ M hS1MB-4 competitor. Clone numbers correspond to those of the format hS1MB-X in Fig. 4B and Fig. S6. The absorbance values were scaled to 1-cm path length.

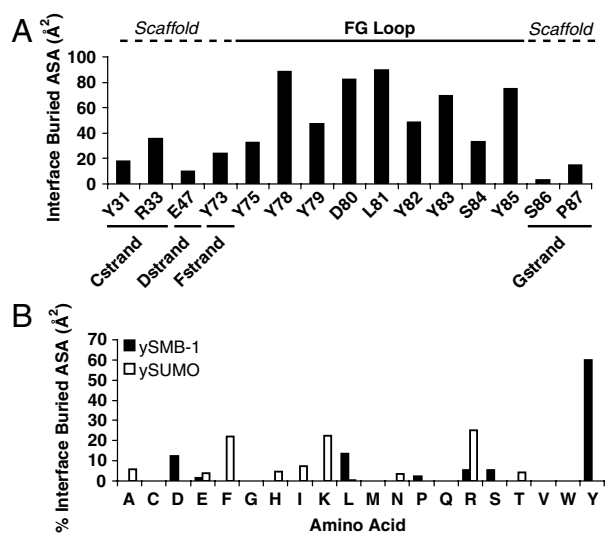


Fig. S4. ySMB-1/ySUMO interface analysis. (A) Buried surface area contributed by each residue in the ySMB-1 paratope. (B) Percent of total ySMB-1 and ySUMO buried surface area contributed by each amino acid type.

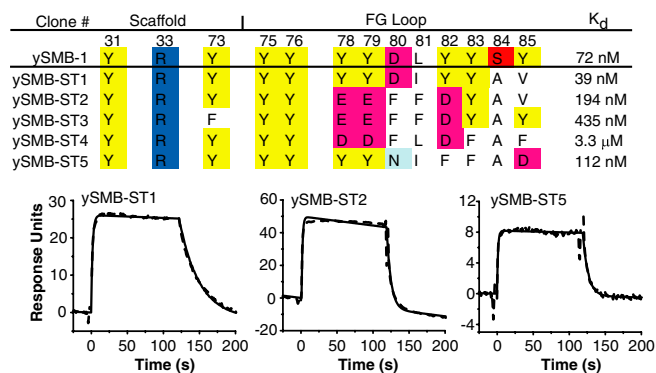


Fig. S5. ySUMO-binding monobodies isolated from the SUMO-targeted library. The amino acid sequences of monobodies recovered against ySUMO from the SUMO-targeted library with K_d values from SPR. Residues are colored as in Fig. S2. Representative SPR traces are shown below.

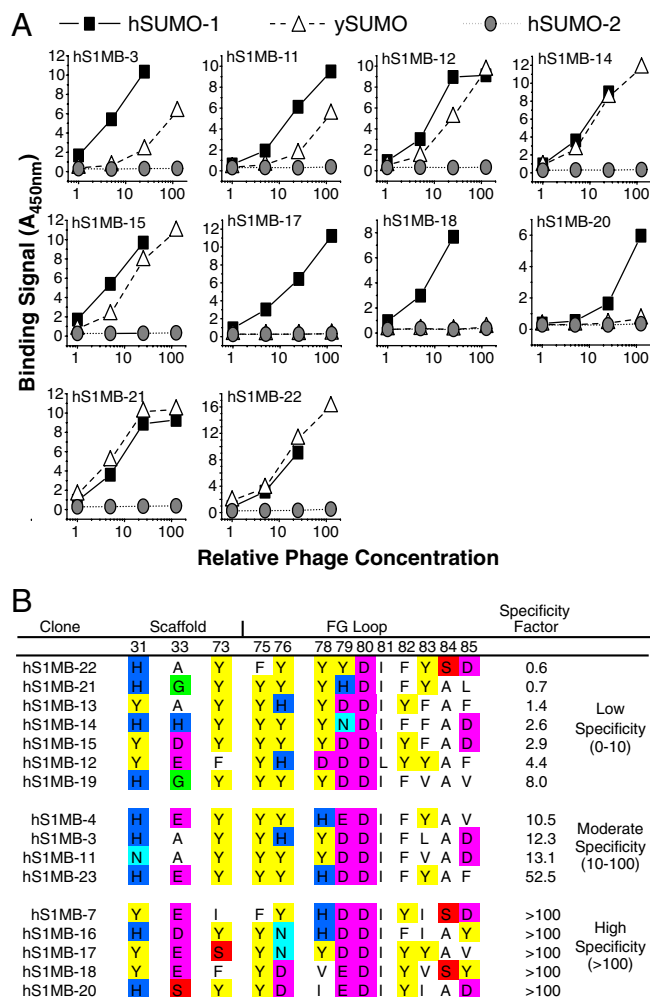


Fig. S6. Selectivity of hSUMO1-binding monobodies. (A) Binding curves derived from phage ELISA of 10 hSUMO-binding monobodies binding to ySUMO, hSUMO1, and hSUMO2. Data for six additional monobodies are shown in Fig. 5A. (B) The amino acid sequences of 16 hSUMO1-binding monobodies are shown and grouped according to their specificity factor for hSUMO1 over ySUMO. Residues are colored as in Fig. S2. The specificity factor is the ratio of apparent affinity measured for hSUMO1 to that for ySUMO in the titration phage ELISA experiment shown in Fig. 5A and Fig. S6A.

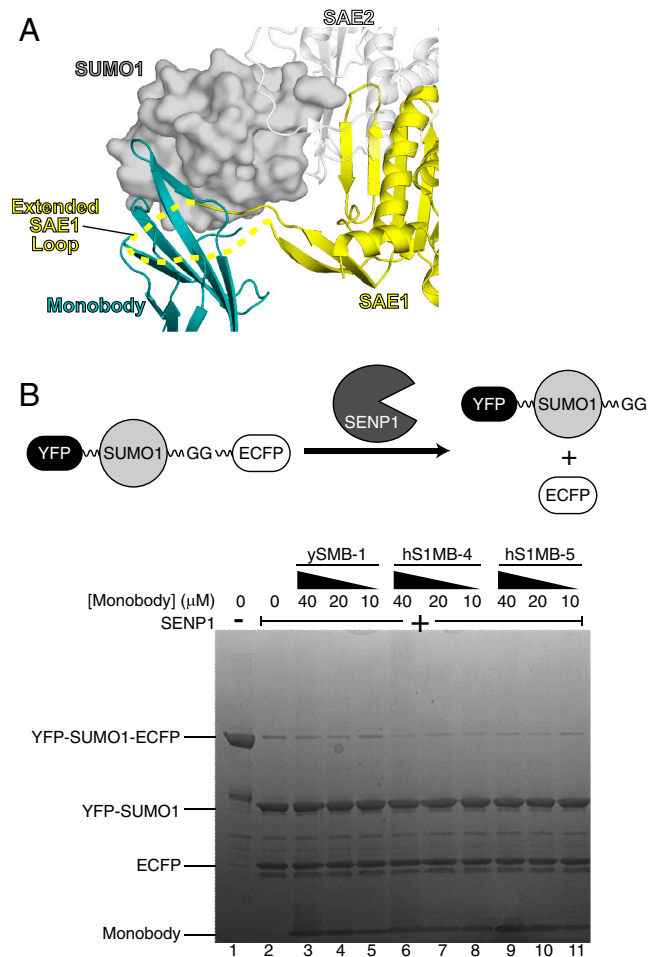


Fig. S7. Proposed mechanism for monobody inhibition of hSUMO1 conjugation and monobody effects on deSUMOylation. (A) A modeled structure of a ySMB-1-like monobody bound to an E1-hSUMO1 complex (PDB ID 3KYD; ref. 1). The trajectory of a long loop of SAE1 that is disordered in the crystal structure is illustrated by a dashed yellow line. (B, Upper) Schematic of the deSUMOylation assay in which a YFP-hSUMO1-enhanced cyan fluorescent protein (ECFP) fusion is cleaved by sentrin(SUMO)-specific protease 1 (SENP1) at the hSUMO1 C-terminal diglycine sequence. (Lower) SDS-PAGE analysis of deSUMOylation reactions carried out in the presence of hSMB-4 (lanes 6–8) or hS1MB-5 (lanes 9–11). Controls are also shown without SENP1 or a monobody (lane 1) or with SENP1 cleavage carried out in the presence of the ySUMO specific ySMB-1 (lanes 2–5). Bands corresponding to the YFP-hSUMO1-ECFP fusion and the YFP-hSUMO1 and ECFP cleavage products are indicated as well as the band corresponding to the monobodies.

1 Olsen SK, Capili AD, Lu X, Tan DS, Lima CD (2010) Active site remodelling accompanies thioester bond formation in the SUMO E1. *Nature* 463:906–912.

Table S1. Crystallographic information and refinement statistics for the structure of the ySMB-1/ySUMO complex (PDB ID code 3QHT)

Data collection*	
Beamline	APS 21-ID-F
Space group	$P2_12_12$
Cell parameters, Å	$a = 59.64, b = 175.46, c = 52.83, \alpha = \beta = \gamma = 90^\circ$
Wavelength, Å	0.97872
Resolution, Å	50.00–2.40 (2.49–2.40)
Unique reflections	22,586
R_{merge}^\dagger	0.085 (0.643)
Completeness	100.0% (99.5%)
Redundancy	7.1 (6.6)
$I/\sigma(I)$	18.9 (2.2)
Refinement statistics	
Resolution range, * Å	20.00–2.40 (2.46–2.40)
Unique reflections	
Working set	21,341
Free set	1,151
R^\ddagger	0.223
R_{free}^\S	0.272
Overall mean B values, Å ²	49.82
No. of amino acid residues	338
No. of water molecules	85
Matthews coefficient	3.20 (water content 61.6%)
rmsd from ideal values	
Bonds/angle	0.02 Å/1.9°
Estimated overall coordinate error based on maximum likelihood, Å	0.2
Estimated overall error for B values based on maximum likelihood	14.4
Ramachandran plot statistics	
Residues in most favored regions, %	87.8 (258)
Residues in additionally allowed regions, %	9.2 (27)
Residues in generously allowed regions, %	1.7 (5)
Residues in disallowed regions, %	1.4 (4)

APS, Advanced Photon Source.

*Values for highest resolution shell shown in parentheses.

$^\dagger R_{\text{merge}} = \frac{\sum_{\text{hkl}} \sum_i |I(\text{hkl})_i - \langle I(\text{hkl}) \rangle|}{\sum_{\text{hkl}} \sum_i \langle I(\text{hkl}) \rangle}$ over I observations of a reflection hkl.

$^\ddagger R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|}$.

$^\S R_{\text{free}}$ is R with 5% of reflections sequestered before refinement.

Table S2. Interface statistics for monobody ySMB-1 and SIM peptides

	Monobody ySMB-1	Avg SIM peptide*
Buried surface, Å ²	670	635 ± 80
SC value	0.72	0.77 ± 0.02
% neutral and nonpolar atoms in interface	64	64 ± 12

*Values reported are the average for five SUMO/SIM complexes (PDB ID codes 1WYW, 1Z5S, 2ASQ, 2KQS, and 2RPQ). The standard deviations in these values across all five complexes are given. Buried surface and % composition values calculated using the PROTORP server (1). Shape complementarity (SC) values calculated using the SC program in the CCP4 suite (2, 3).

1 Reynolds C, Damerell D, Jones S (2009) ProtorP: A protein–protein interaction analysis server. *Bioinformatics* 25:413–414.

2 Lawrence MC, Colman PM (1993) Shape complementarity at protein/protein interfaces. *J Mol Biol* 234:946–950.

3 Anon (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50:760–763.