Supporting Information for

Fast and Sequence-Specific Palladium-Mediated Cross-Coupling Reaction Identified from Phage Display

Reyna K. V. Lim,[†] Nan Li,[‡] Carlo P. Ramil, and Qing Lin*

Department of Chemistry, State University of New York at Buffalo, Buffalo, New York 14260-3000, United States, qinglin@buffalo.edu [†]Current address: California Institute for Biomedical Research, 11119 N Torrey Pines Rd #100, La Jolla, CA 92037, USA. [‡]Current address: School of Chemistry and Chemical Engineering, Shaanxi Normal University, Xi'an, 710062, P. R. China.

Table of Contents

Supplemental Figures and Tables

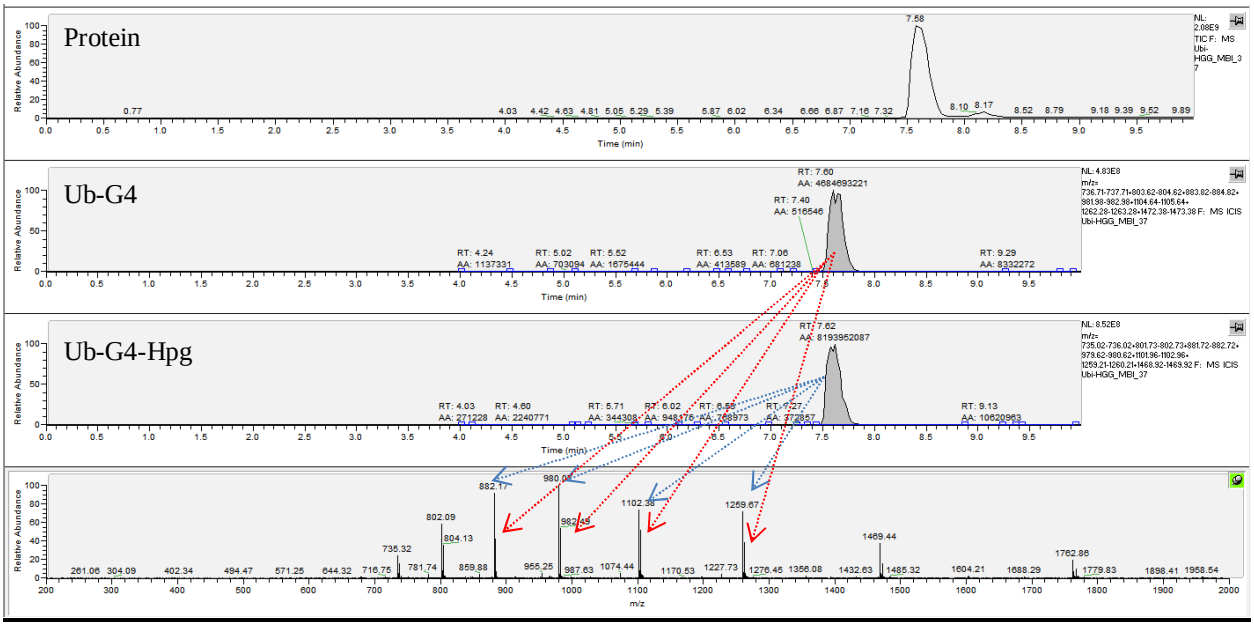


Figure S1. Determination of Hpg occupancy in Ub-G4-Hpg after expressing the proteins in XL1- Blue cells in SelenoMet Plus Nutrient Mix medium supplemented with Hpg and methionine biosynthesis inhibitors**.** The occupancy was calculated to be 64% based on the ion counts of Ub-G4-Hpg and Ub-G4-Met.

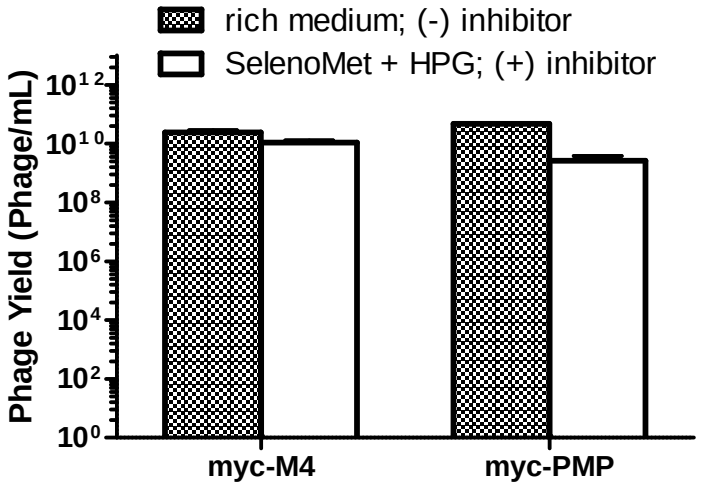


Figure S2. Comparison of phage yield after amplifying the phage in XL1-Blue cells grown in rich medium or SelenoMet plus Nutrient Mix medium supplemented with Hpg and methionine biosynthesis inhibitors. The phage UV absorbance at 268 nm was measured and the phage yield was calculated based on the following: A_{268} = 1.0 corresponds to a phage yield of 5×10^{12} phage/mL.[1]

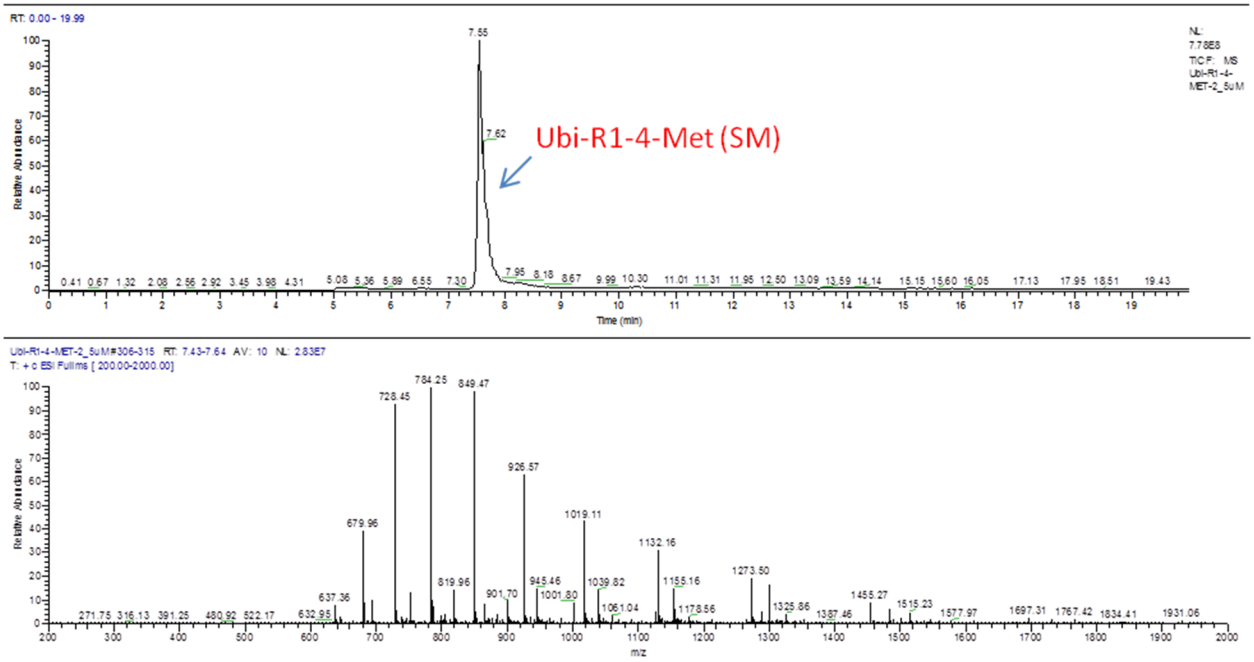


Figure S3. LC-MS analysis of the palladium-mediated cross-coupling reaction involving Ub-R1-4-Met in PBS buffer at 37 ºC for 30 min. Only charge ladders corresponding to Ub-R1-4-Met were detected, indicating a complete recovery of the starting materials (the minor peaks correspond to the impurity present in the starting materials).

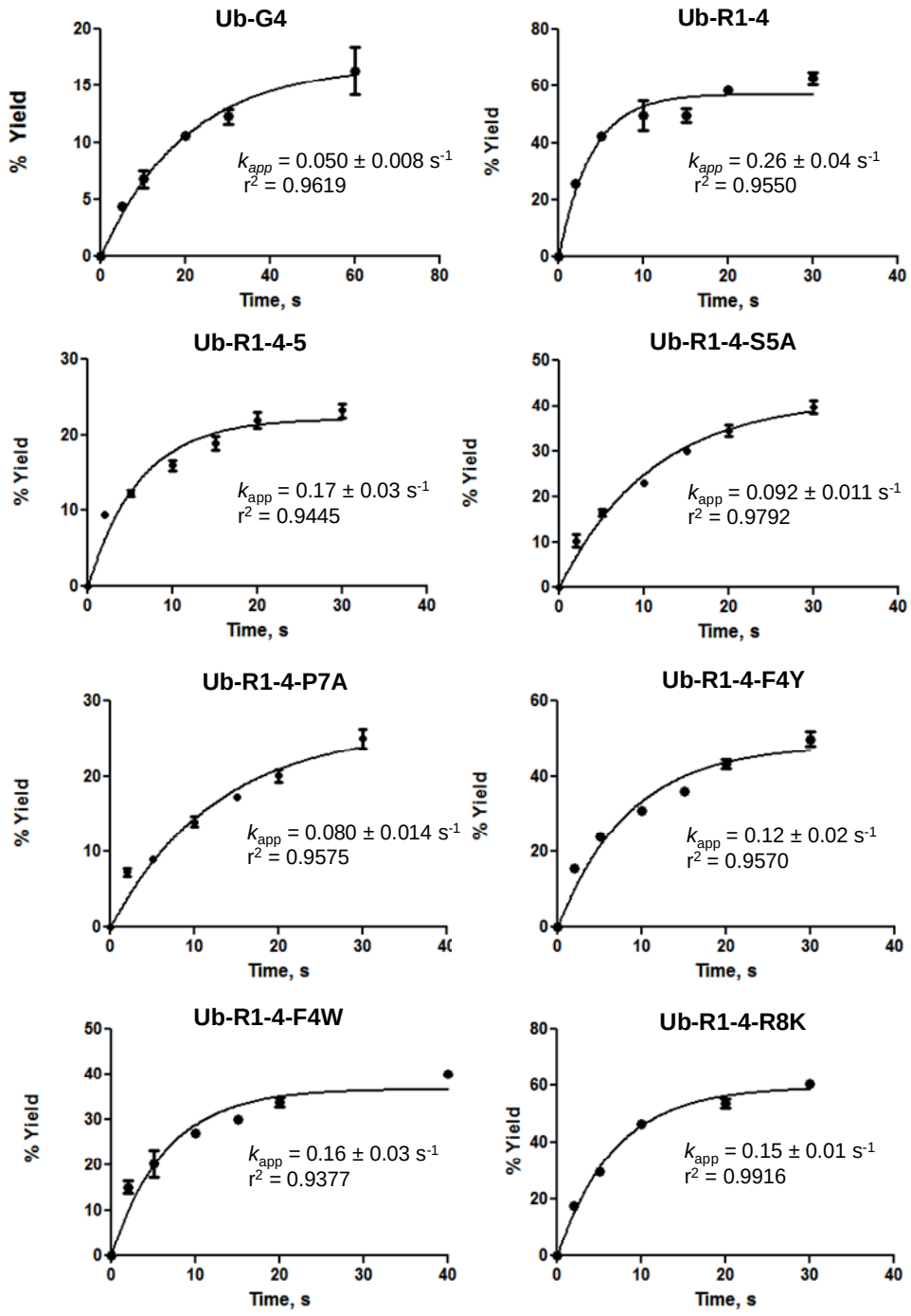
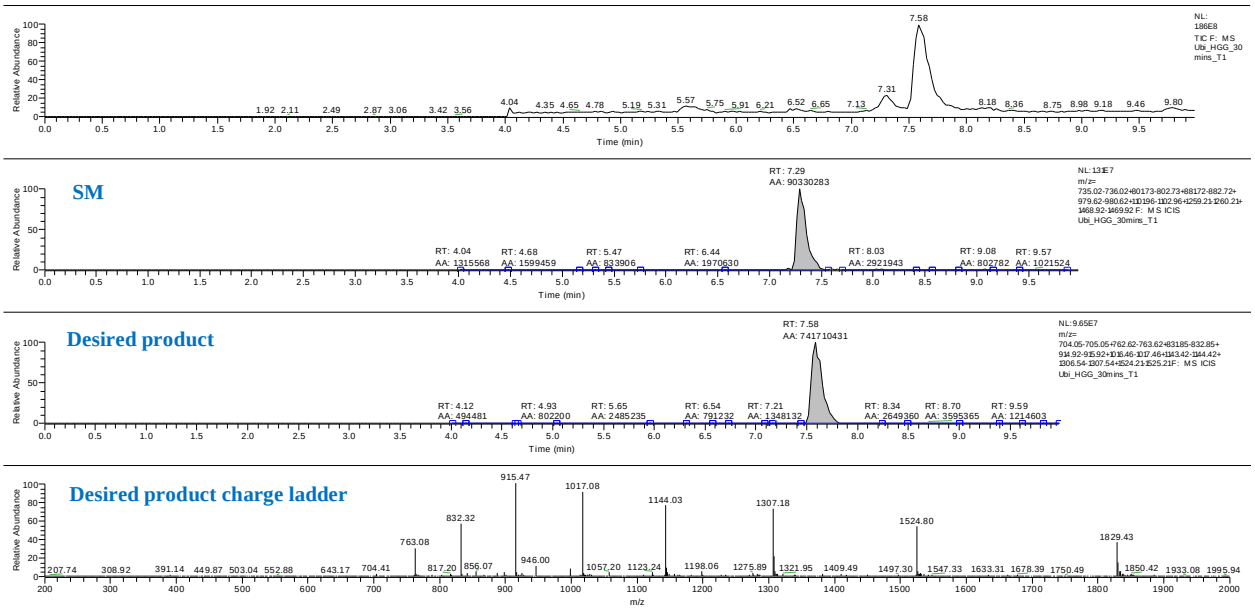
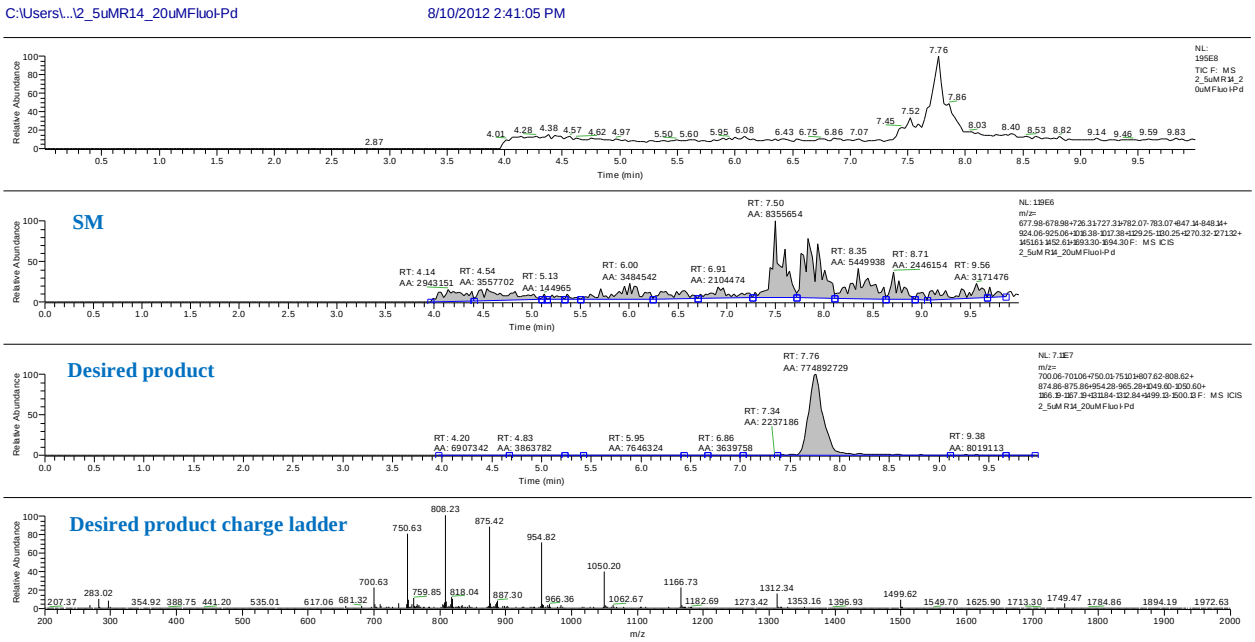


Figure S4. Determining the initial rate constants. The reaction yields were plotted vs. time for the palladium-mediated Sonogashira cross-coupling between Ub-G4, Ub-R1-4 and its mutants and the preactivated fluorescein iodide-Pd complex. Each data point represents the average of at least two independent measurements. The initial rate constant, k_{app}, was derived by fitting the curve to the equation: $Y=Y_{max}(1 - e^{-kapp*t})$, where Y is the % yield. The second-order rate constant k_2 was calculated using the equation: $k_2 = k_{app}/[\text{Ar-Pd}]$, where the concentration of the preactivated fluorescein iodide-Pd complex was 2×10^{-5} M.

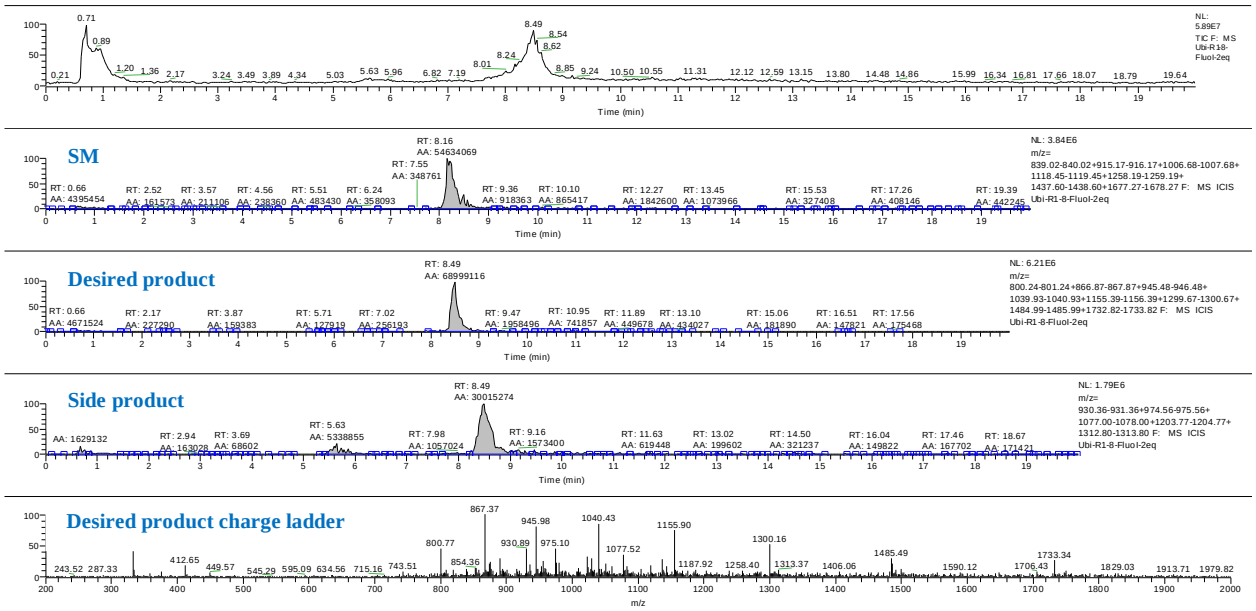
a) Ub-G4 (Table 1, entry 1)



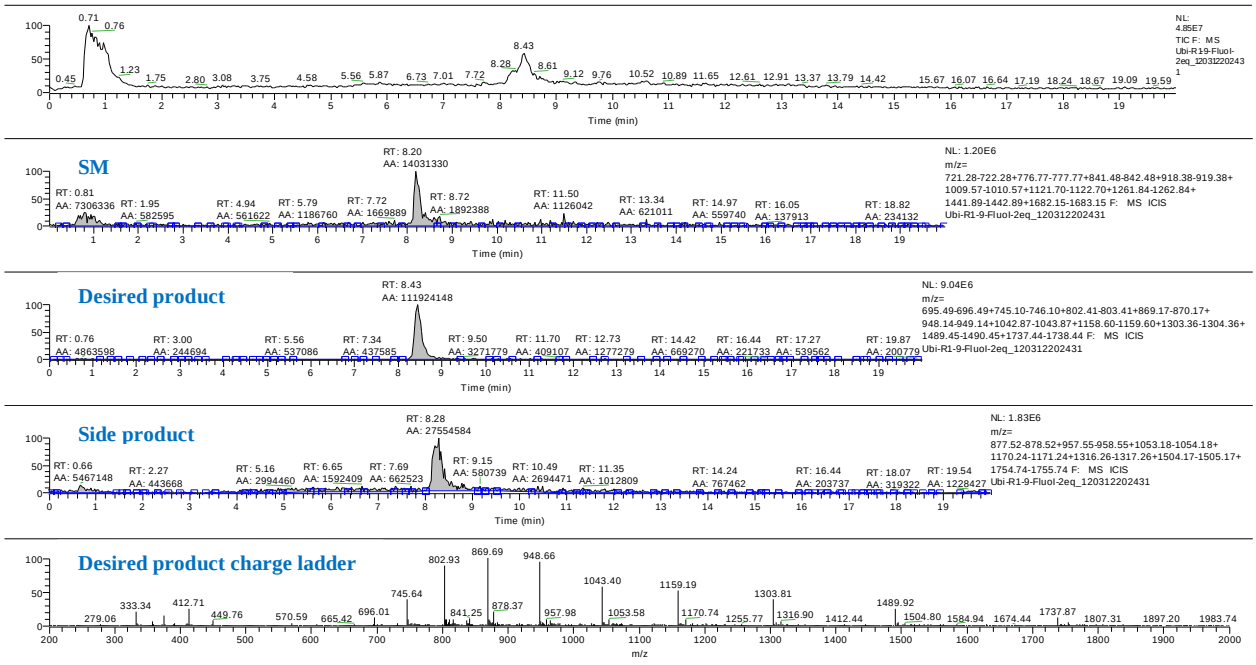
b) Ub-R1-4 (Table 1, entry 2)



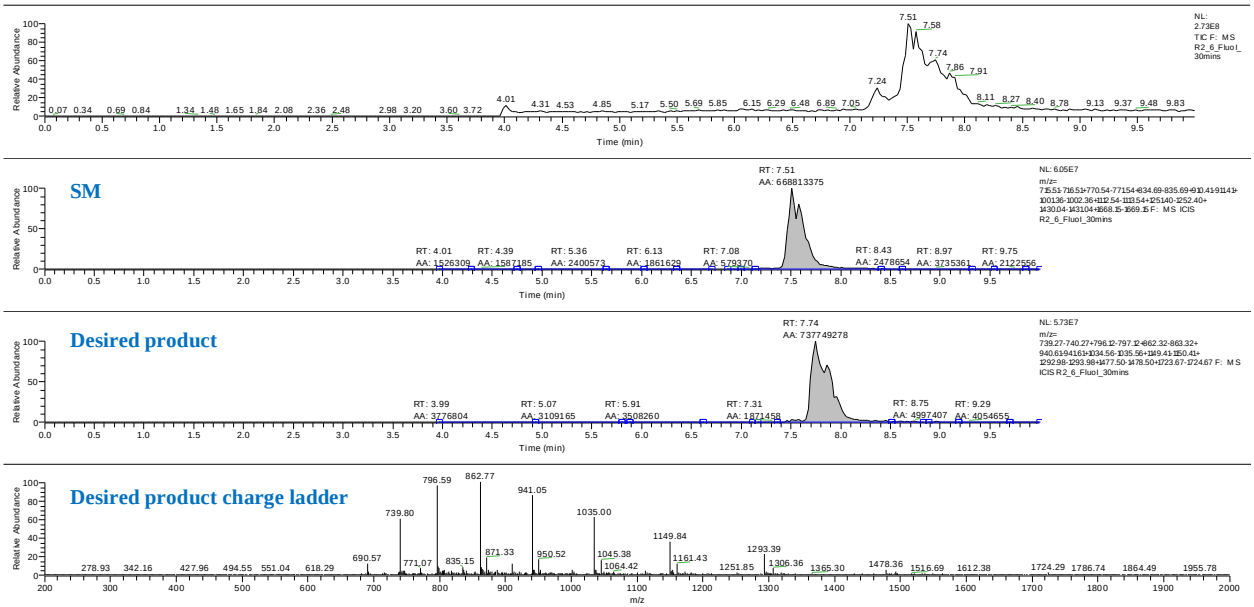
c) Ub-R1-8 (Table 1, entry 3)



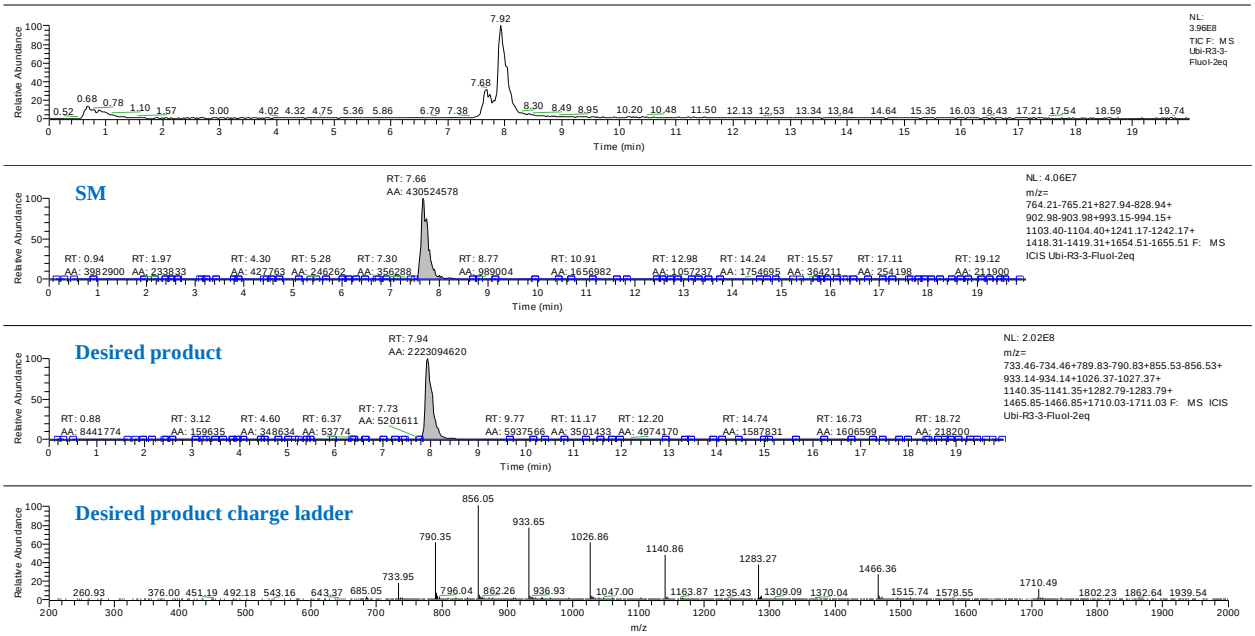
d) Ub-R1-9 (Table 1, entry 4)



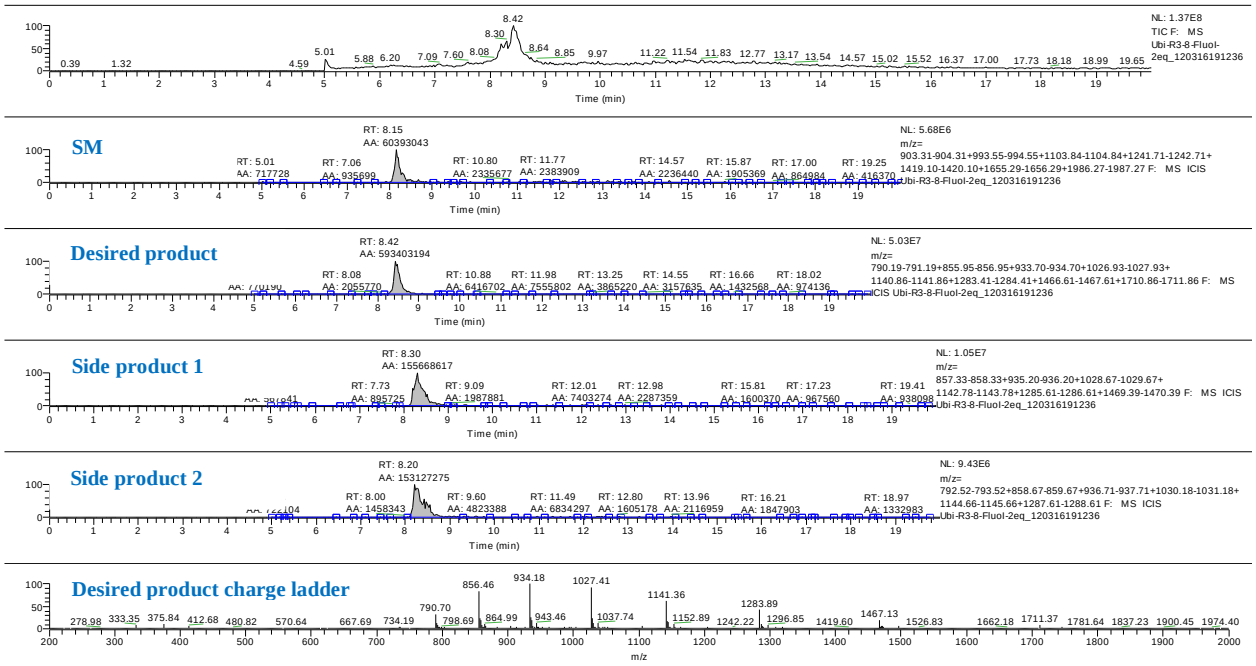
e) Ub-R2-6 (Table 1, entry 5)



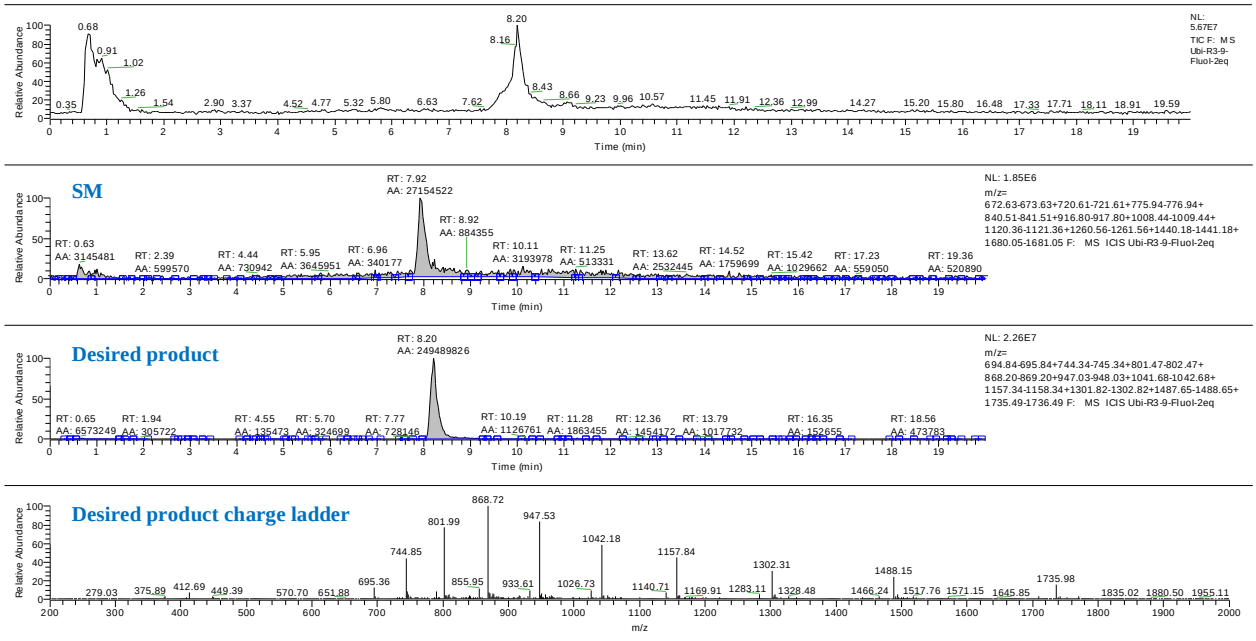
f) Ub-R3-3 (Table 1, entry 6)



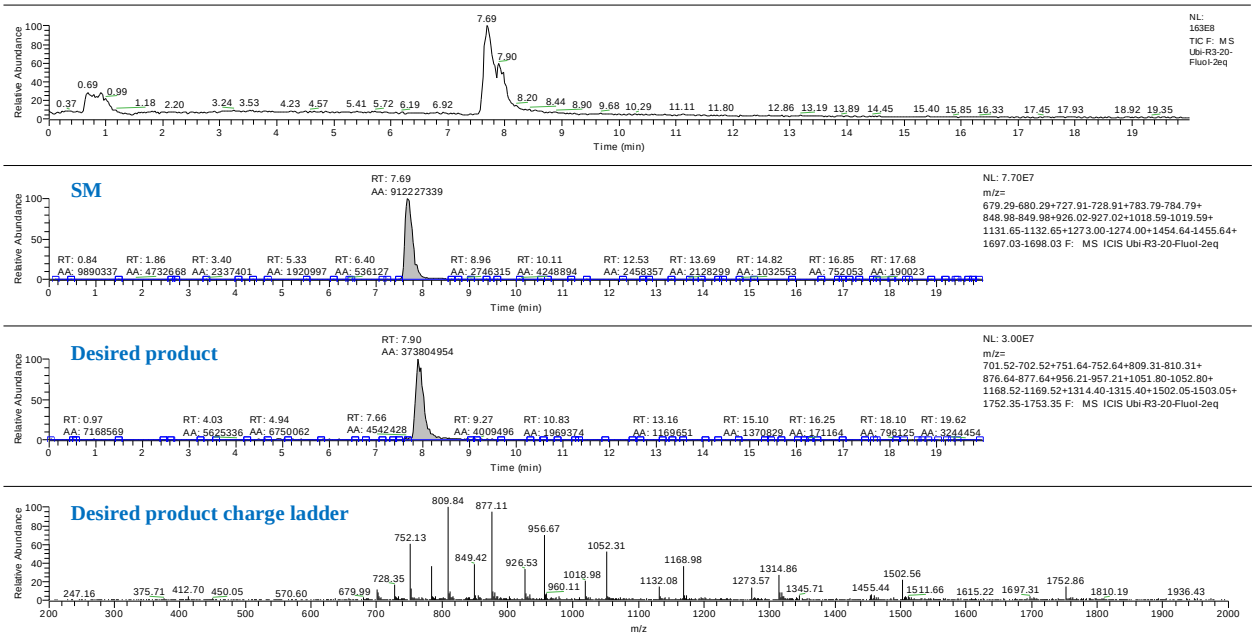
g) Ub-R3-8 (Table 1, entry 7)



h) Ub-R3-9 (Table 1, entry 8)



i) Ub-R3-20 (Table 1, entry 9)



j) R3-24 (Table 1, entry 10)

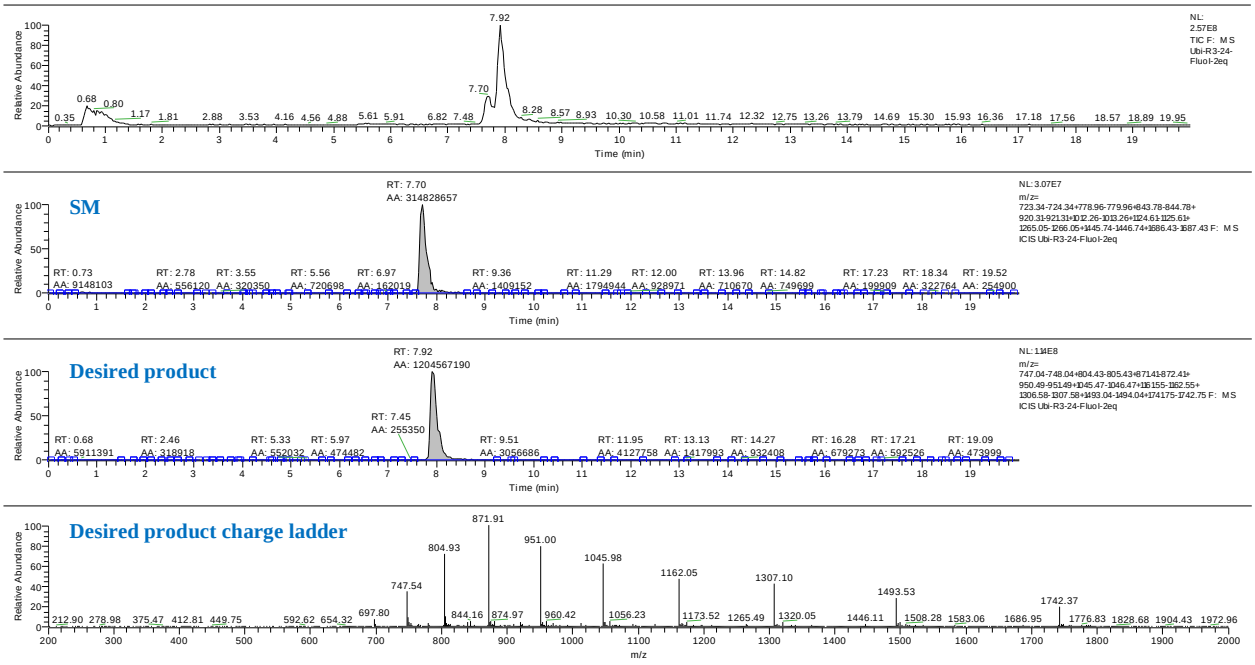
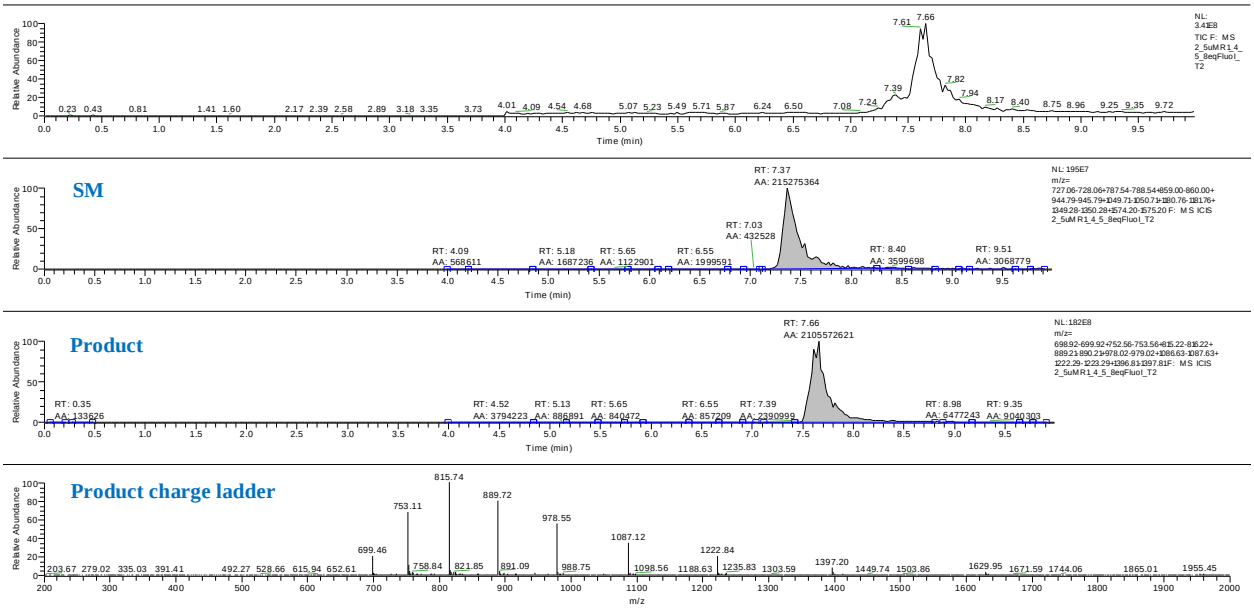
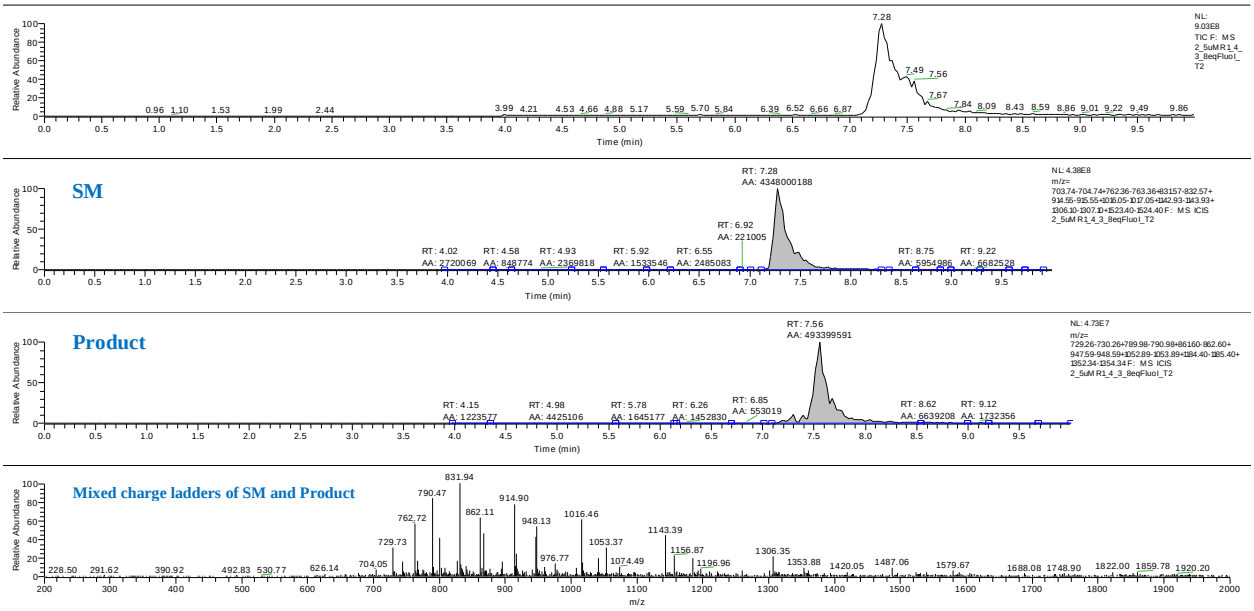


Figure S5. LC-MS analyses of the reaction mixtures for the various ubiquitin fusion proteins listed in Table 1. The yields of the products were determined by LC-MS based on ion counts and calculated using the following equation: yield % = $I_{product}/(I_{Ub\text{-}peptide} + I_{product} + I_{side\ product(s)})$, where $I_{Ub\text{-}peptide}$, $I_{product}$, and $I_{side\ product(s)}$ represent the ion counts of the remaining Ub-peptide, product, and side product(s) (if present), respectively.

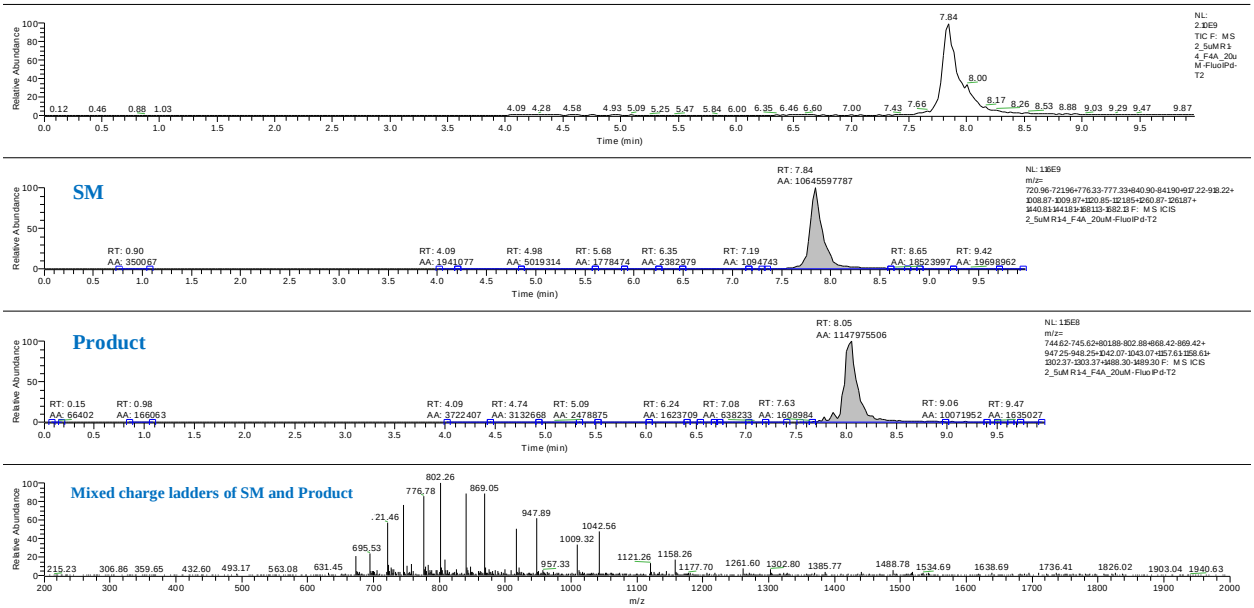
a) R1-4-5 (Table 2, entry 3)



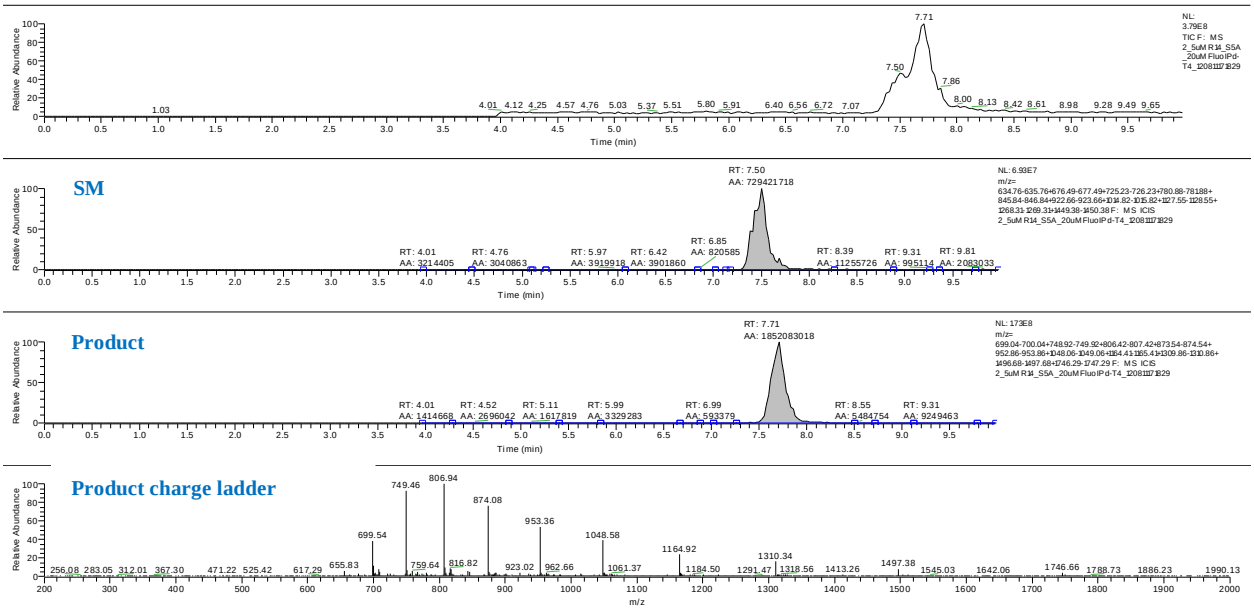
b) R1-4-3 (Table 2, entry 4)



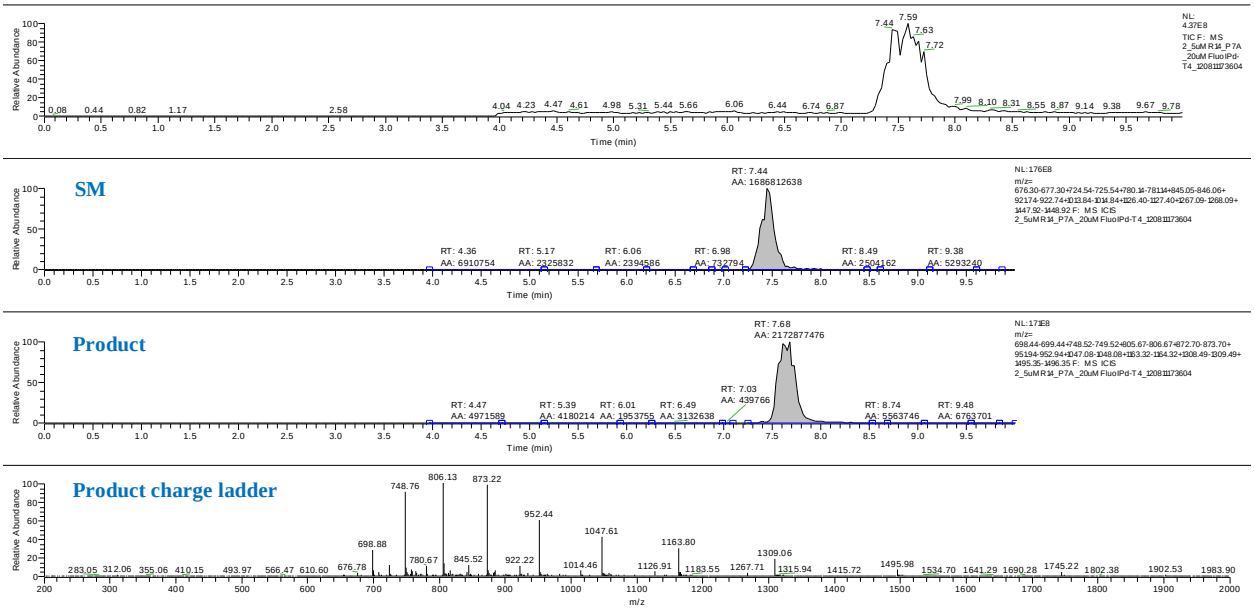
c) R1-4-F4A (Table 2, entry 5)



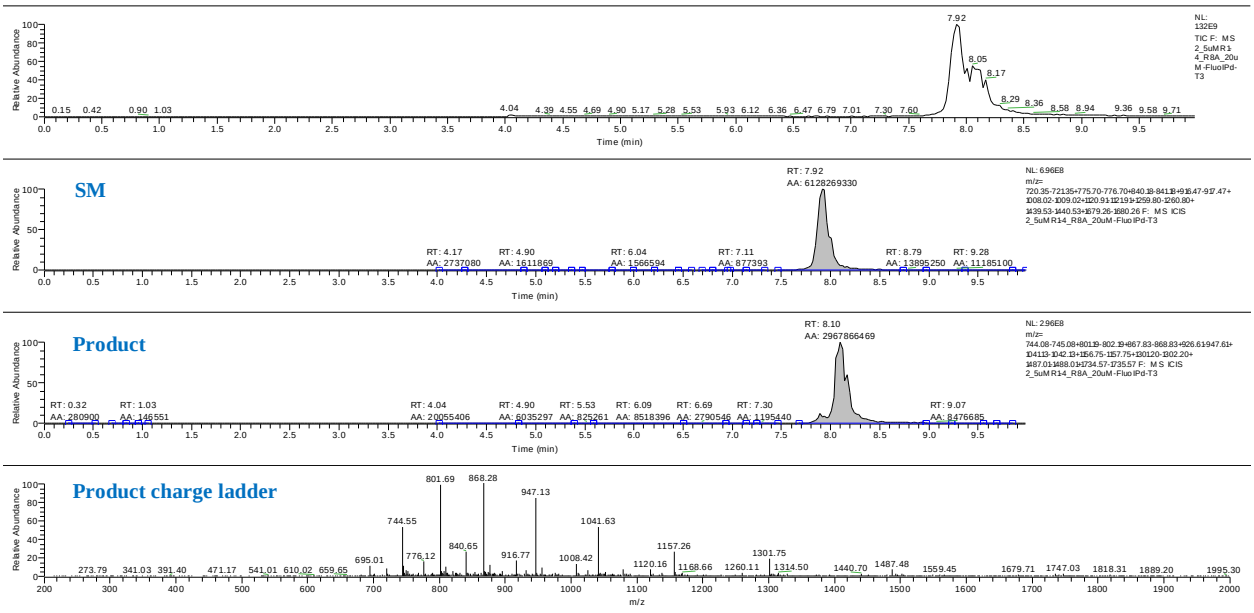
d) R1-4-S5A (Table 2, entry 6)



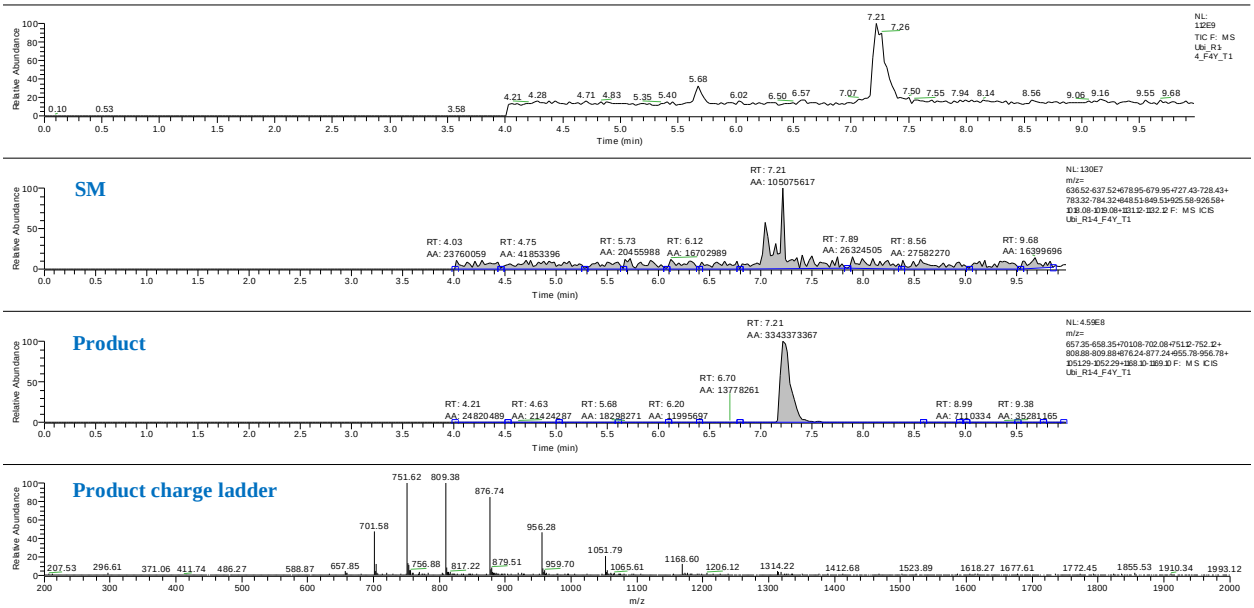
e) R1-4-P7A (Table 2, entry 7)



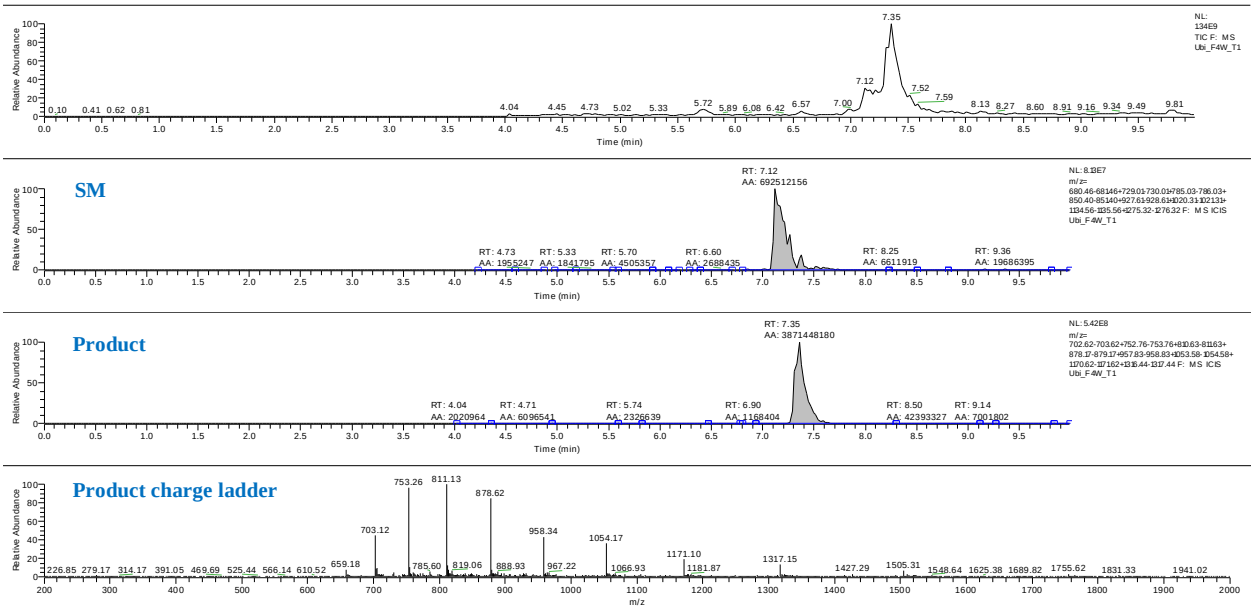
g) R1-4-R8A (Table 2, entry 8)



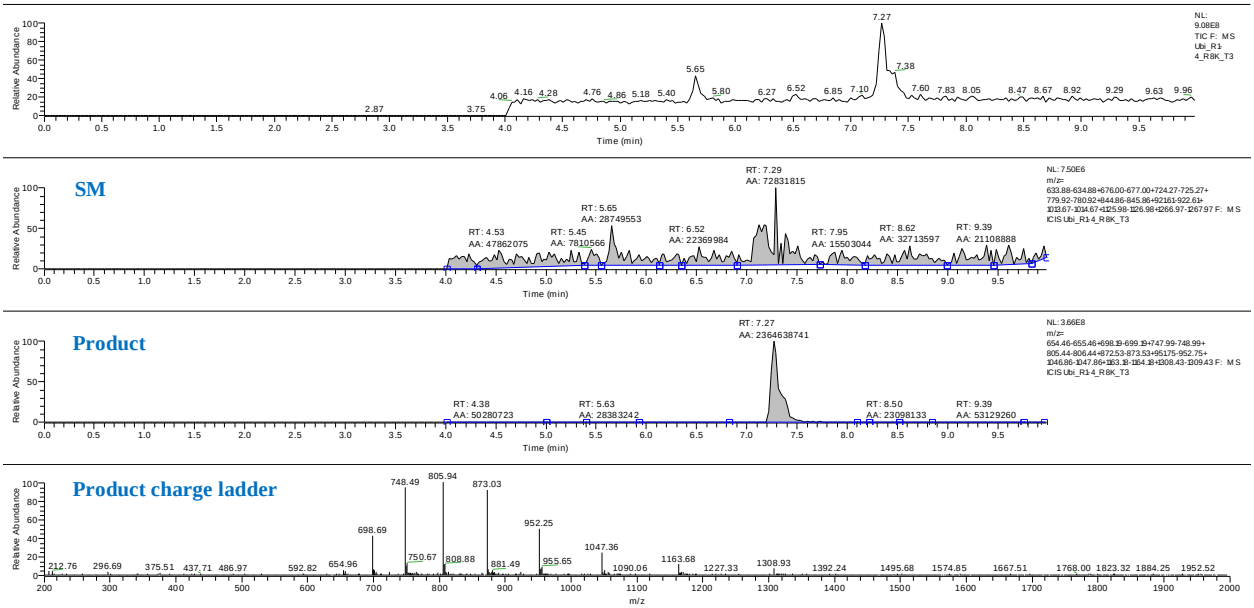
h) R1-4-F4Y (Table 2, entry 9)



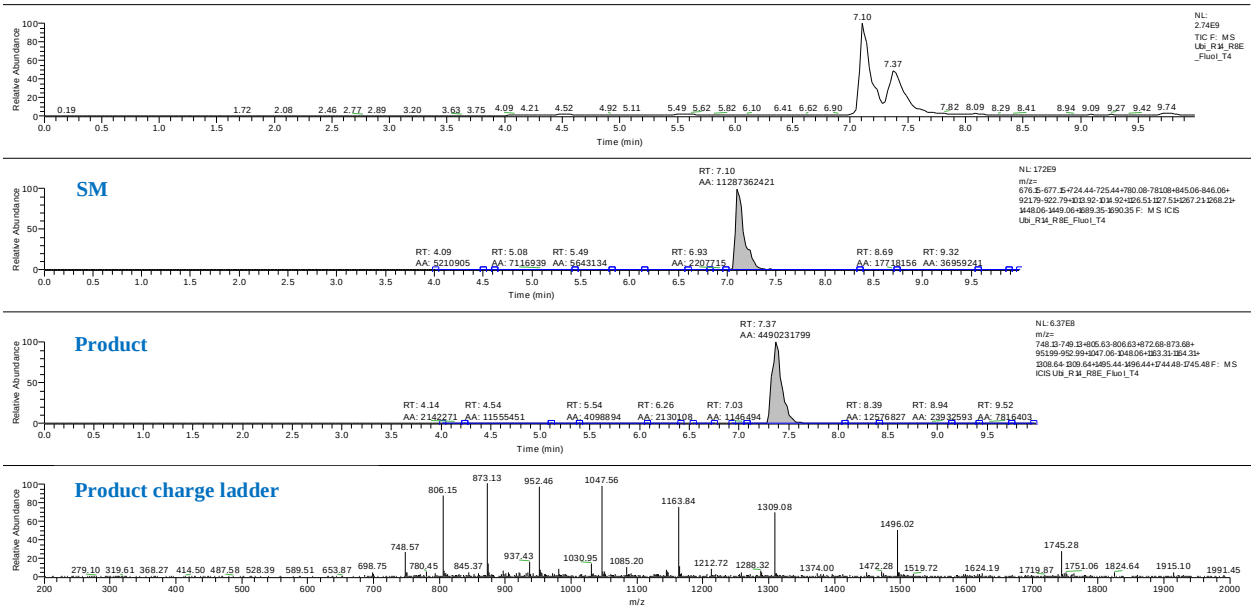
i) R1-4-F4W (Table 2, entry 10)



j) R1-4-R8K (Table 2, entry 11)



k) R1-4-R8E (Table 2, entry 12)



l) R1-4-R8H (Table 2, entry 13)

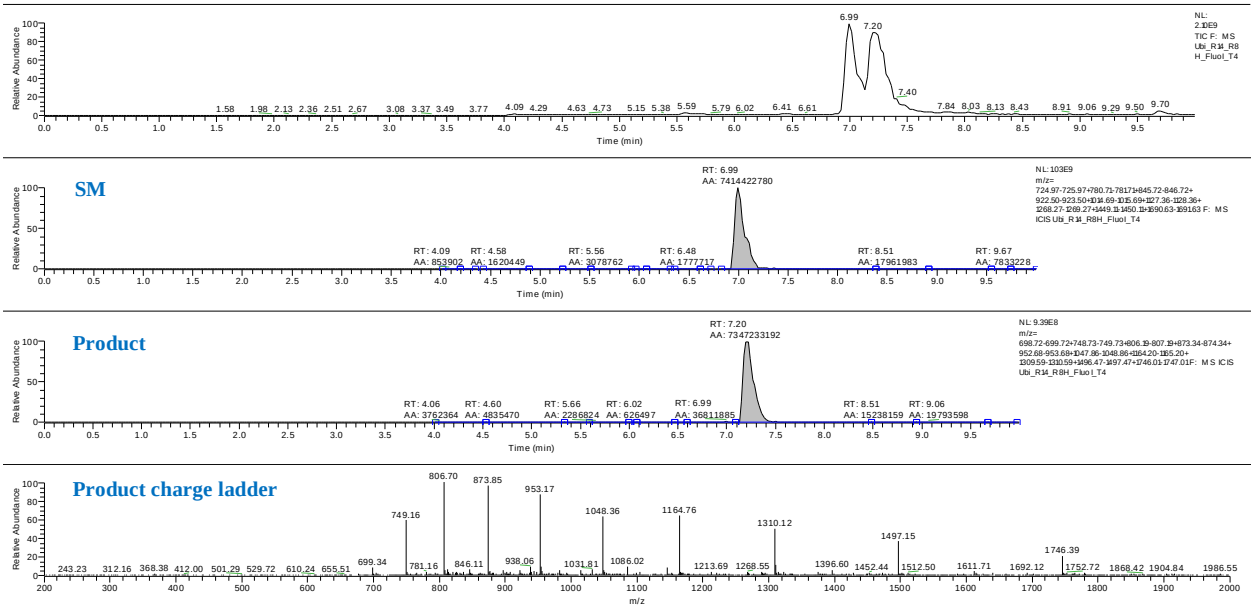
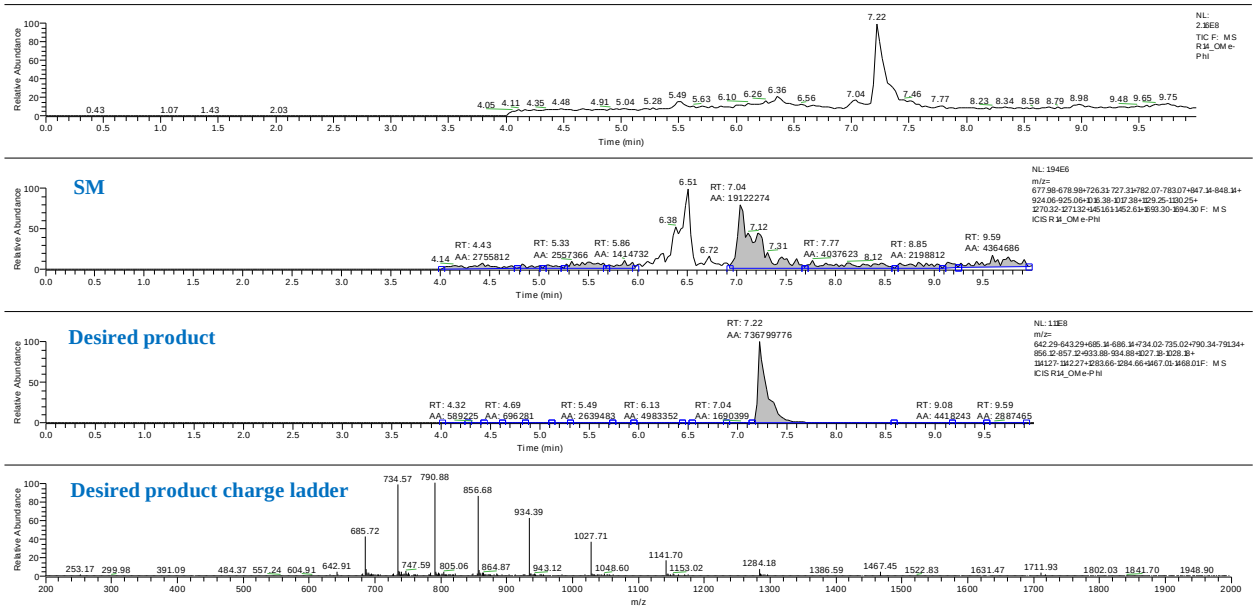
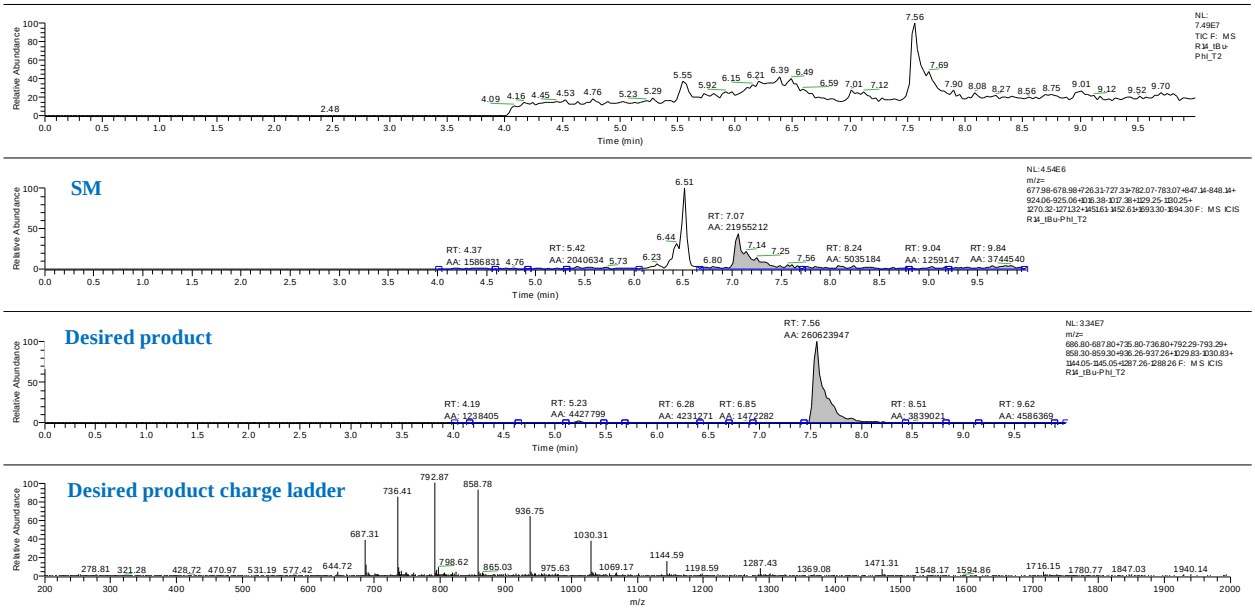


Figure S6. LC-MS analyses of the reaction mixtures for the various ubiquitin R1-4 mutant proteins listed in Table 2. The yields of the products (average of at least two trials) were determined by LC-MS based on ion counts and calculated using the following equation: yield % = $I_{product}/(I_{Ub\text{-}peptide} + I_{product} + I_{side\ product(s)})$, where $I_{Ub\text{-}peptide}$, $I_{product}$, and $I_{side\ product(s)}$ represent the ion counts of the remaining Ub-peptide, product, and side products if present, respectively.

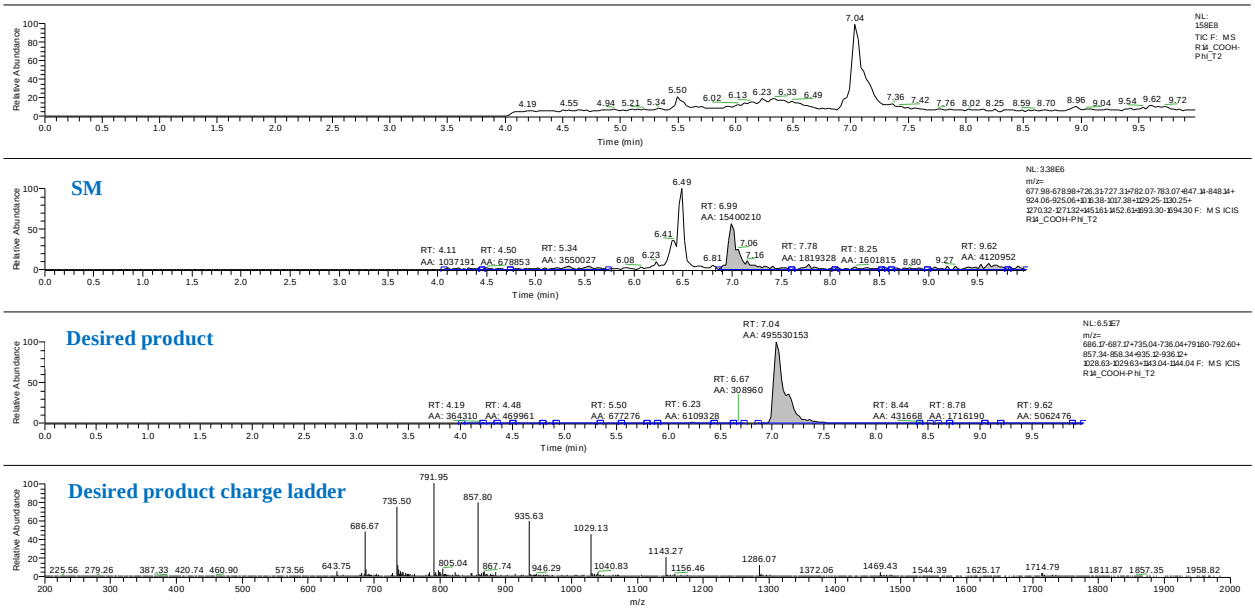
p-Methoxyphenyl iodide (97%)



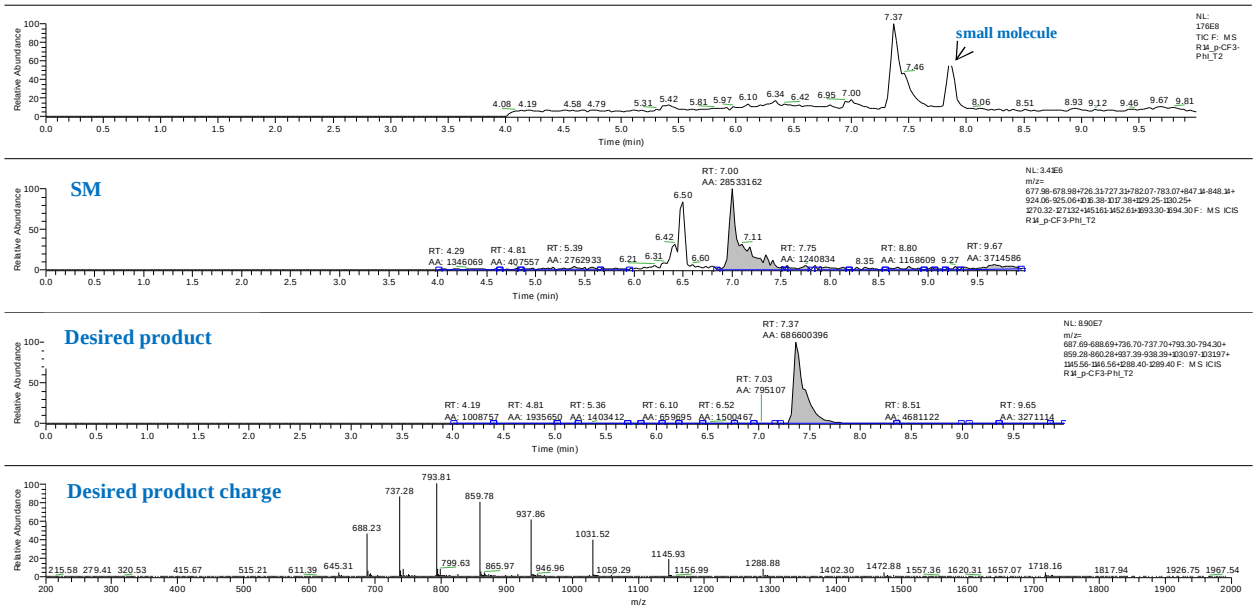
p-tert-Butylphenyl iodide (92%)



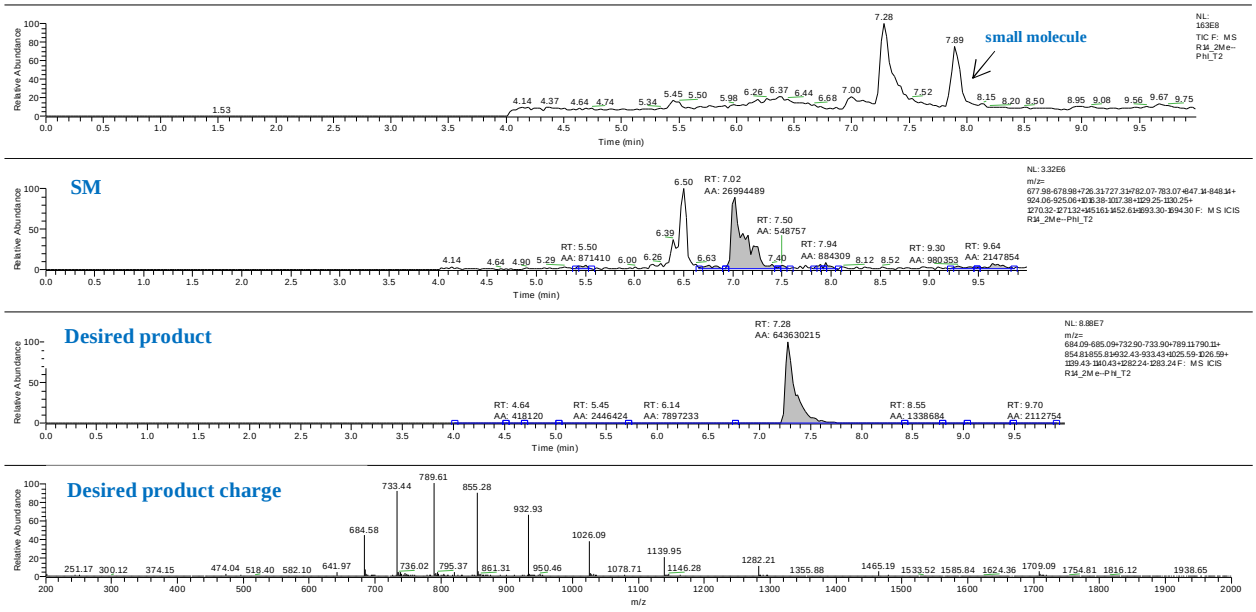
p-Iodobenzoic acid (97%)



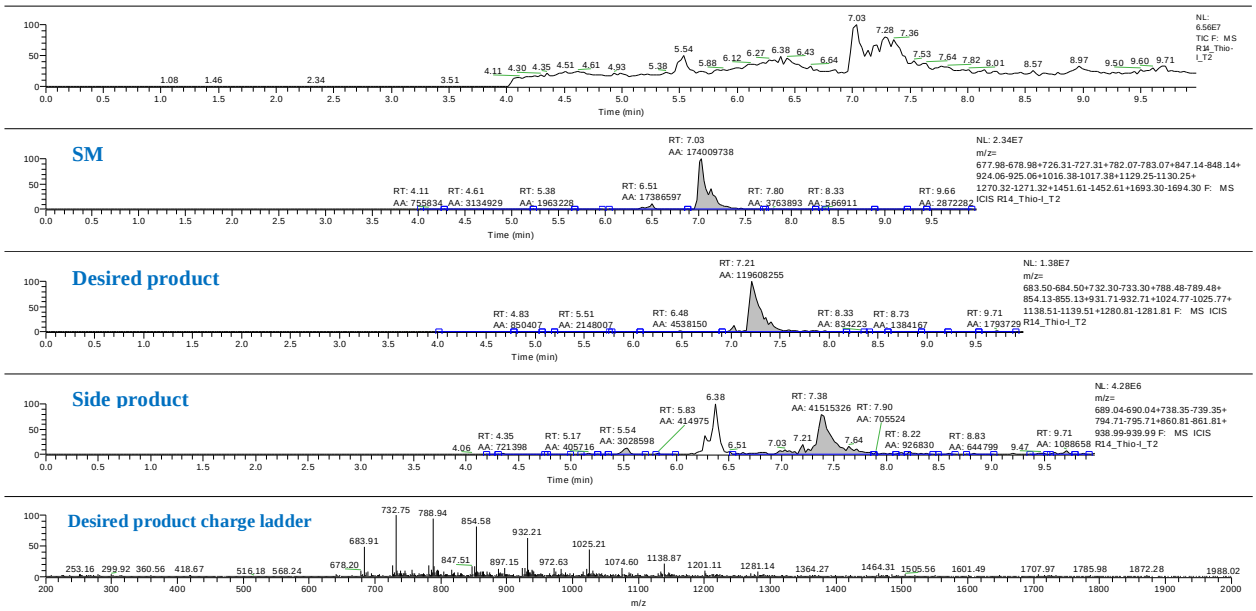
p-Trifluoromethylphenyl iodide (96%)



o-Methylphenyl iodide (96%)



Thiophenyl iodide (36%)



Coumarin iodide (54%)

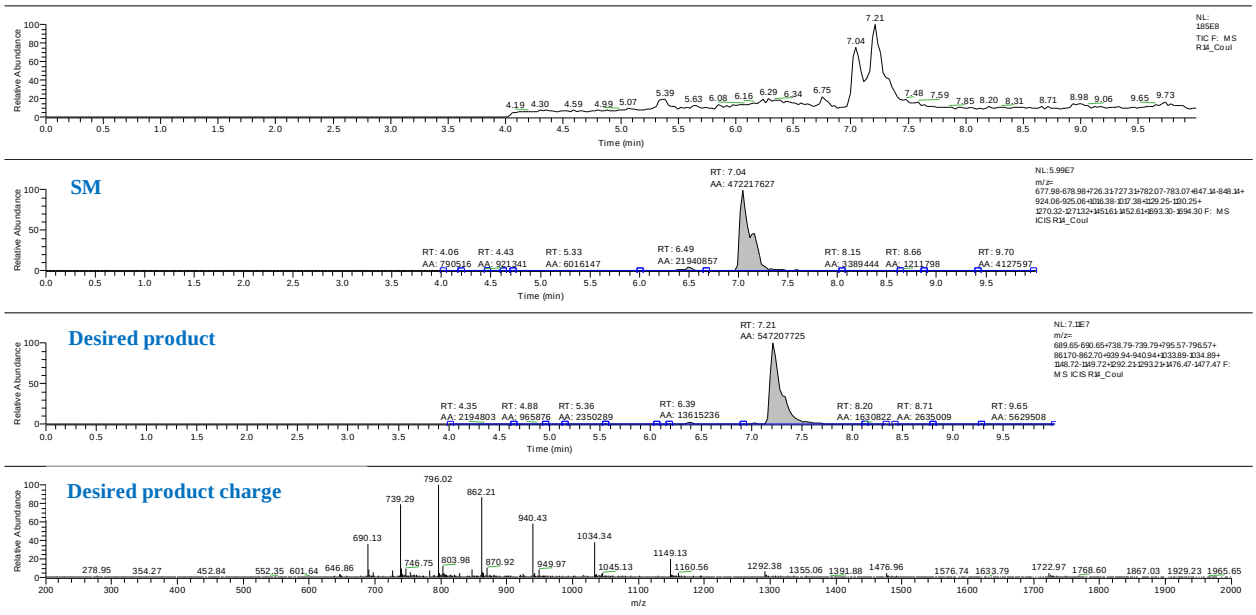
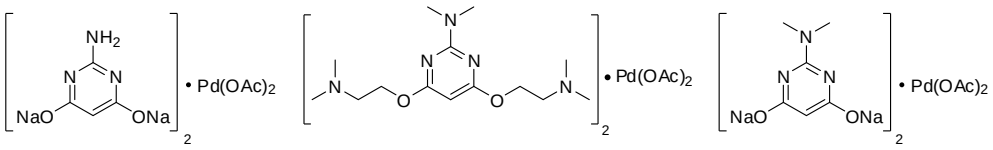


Figure S7. LC-MS analyses of the reaction mixtures of Ub-R1-4 toward various aryl/vinyl iodides listed in Figure 4. The yields of the products (average of at least two trials) were determined by LC-MS based on ion counts and calculated using the following equation: yield % = $I_{product}/(I_{Ub\text{-}peptide} + I_{product} + I_{side\ product(s)})$, where $I_{Ub\text{-}peptide}$, $I_{product}$, and $I_{side\ product(s)}$ represent the ion counts of the remaining Ub-peptide, product, and side products if present, respectively.

Table S1. Structure and effect of palladium catalysts on phage viability.



Pd-cat-1 **Pd-cat-2** **Pd-cat-3**

	Dilution	No. of Plaques
Control (ER 2738)	---	0
M13KE	10^7	80-90
M13KE + 37°C	10^7	100-150
M13KE + 37°C + 0.5 mM **Pd-cat-1**	10^6	800-1000
M13KE + 37°C + 1.0 mM **Pd-cat-1**	10^6	800-1000
M13KE + 37°C + 2.0 mM **Pd-cat-1**	10^6	800-1000
M13KE + 37°C + 4.0 mM **Pd-cat-1**	10^6	800-1000
M13KE + 37°C + 8.0 mM **Pd-cat-1**	10^6	240-500
M13KE + 37°C + 9.0 mM **Pd-cat-1**	10^6	290-500

Table S2. Fraction eluted after each round of Cu-free Sonogashira cross-coupling reaction-based solution-phase panning.

	Phage loaded (*cfu*)	**Phage eluted (*cfu*)**	**Fraction eluted** [a]
Round 1	5.0×10^{10}	1.92×10^{6}	3.84×10^{-5}
Round 2	4.0×10^{9}	5.50×10^{5}	1.38×10^{-4}
Round 3	1.4×10^{9}	8.31×10^{4}	5.95×10^{-5}

[a] Fraction eluted was calculated by dividing phage eluted over phage loaded.

Table S3. Representative sequences from the naive library and after each round of the reaction-based panning.

PRESELECTION											
1	S	N	G	Y	T	M	L	G	C	S	C
2	V	R	T	S	T	M	Y	I	Y	S	P
3	A	Y	P	V	S	M	N	L	R	F	R
4	S	T	H	G	P	M	M	T	N	F	I
5	V	C	D	N	S	M	S	I	S	P	S
6	D	S	I	V	T	M	F	P	A	S	A
7	S	D	L	T	A	M	V	H	Y	Y	P
8	D	S	V	C	S	M	P	T	V	S	N
9	R	T	D	N	C	M	A	I	I	T	I
10	H	V	A	Y	D	M	P	I	N	L	H

ROUND 1											
R1-1	G	T	P	D	P	M	L	V	T	I	F
R1-2	T	C	A	L	D	M	H	S	S	R	I
R1-3	C	V	T	N	C	M	Y	Y	Y	S	Y
R1-4	G	R	Y	F	S	M	P	R	P	S	R
R1-5	C	Y	S	H	F	M	N	C	L	C	G
R1-6	S	R	P	H	T	M	P	T	F	R	H
R1-8	P	S	L	P	C	M	Y	S	F	P	D
R1-9	A	L	F	L	P	M	S	R	V	H	D
R1-12	C	L	P	R	C	M	A	S	P	P	I

ROUND 2											
R2-1	T	N	S	G	N	M	A	V	S	R	N
R2-2	R	T	N	R	C	M	Y	D	C	P	F
R2-3	P	Y	C	T	I	M	N	G	C	V	N
R2-6	D	P	P	P	T	M	D	V	P	P	H
R2-7	N	F	R	F	I	M	A	G	I	C	H
R2-8	F	S	D	G	T	M	K	L	R	R	H
R2-9	Y	T	S	S	R	M	S	I	S	S	Y
R2-11	S	P	N	L	P	M	L	P	S	V	P
R2-12	S	A	D	A	S	M	P	I	H	I	A

ROUND 3											
R3-3	S	P	P	H	S	M	V	P	T	P	A
R3-6	S	T	I	R	L	M	H	S	N	F	P
R3-8	N	V	P	P	L	M	G	T	P	S	L
R3-9	F	P	R	P	P	M	T	P	P	P	H
R3-20	F	P	T	N	H	M	H	R	H	P	T
R3-24	D	G	R	D	A	M	S	P	H	Y	Y

Table S4. Primers used for the construction of the various ubiquitin fusion proteins.

Fusion protein	Oligonucleotides
Ub-R1-4	F: 5'-CGG CCG CTA CTT CAG CAT GCC CCG CCC CTC CCG CG-3' R: 5'-TCG ACG CGG GAG GGG CGG GGC ATG CTG AAG TAG CGG CCG GTA C-3'
Ub-R1-8	F: 5'-CCC CAG CCT CCC CTG CAT GTA CTC CTT CCC CGA CG-3' R: 5'-TCG ACG TCG GGG AAG GAG TAC ATG CAG GGG AGG CTG GGG GTA C-3'
Ub-R1-9	F: 5'-CGC CCT CTT CCT CCC CAT GTC CCG CGT CCA CGA CG-3' R: 5'-TCG ACG TCG TGG ACG CGG GAC ATG GGG AGG AAG AGG GCG GTA C-3'
Ub-R2-6	F: 5'-CGA CCC CCC CCC CAC CAT GGA CGT CCC CCC CCA CG-3' R: 5'-TCG ACG TGG GGG GGG ACG TCC ATG GTG GGG GGG GGG TCG GTA C-3'
Ub-R3-3	F: 5'-CTC CCC CCC CCA CTC CAT GGT CCC CAC CCC CGC CG-3' R: 5'-TCG ACG GCG GGG GTG GGG ACC ATG GAG TGG GGG GGG GAG GTA C-3'
Ub-R3-6	F: 5'-CTC CAC CAT CCG CCT CAT GCA CTC CAA CTT CCC CG-3' R: 5'-TCG ACG GGG AAG TTG GAG TGC ATG AGG CGG ATG GTG GAG GTA C-3'
Ub-R3-8	F: 5'-CAA CGT CCC CCC CCT CAT GGG CAC CCC CAG CCT CG-3' R: 5'-TCG ACG AGG CTG GGG GTG CCC ATG AGG GGG GGG ACG TTG GTA C-3'
Ub-R3-9	F: 5'-CTT CCC CCG CCC CCC CAT GAC CCC CCC CCC CCA CG-3' R: 5'-TCG ACG TGG GGG GGG GGG GTC ATG GGG GGG CGG GGG AAG GTA C-3'
Ub-R3-20	F: 5'-CTT CCC CAC CAA CCA CAT GCA CCG CCA CCC CAC CG-3' R: 5'-TCG ACG GTG GGG TGG CGG TGC ATG TGG TTG GTG GGG AAG GTA C-3'
Ub-R3-24	F: 5'-CA CGG CCG CGA CGC CAT GTC CCC CCA CTA CTA CG-3' R: 5'-TCG ACG TAG TAG TGG GGG GAC ATG GCG TCG CGG CCG TCG GTA C-3'
Ub-R4-3	F: 5'-CCC CCT CAG CTC CCA CAT GGG CTG CGG CAT CTG CG-3' R: 5'-TCG ACG CAG ATG CCG CAG CCC ATG TGG GAG CTG AGG GGG GTA C-3'

Fusion protein	Oligonucleotides
Ub-R1-4-5	F: 5'-CTT CAG CAT GCC CCG CG-3' R: 5'-TCG ACG CGG GGC ATG CTG AAG GTA C-3'
Ub-R1-4-3	F: 5'-CAG CAT GCC CG-3' R: 5'-TCG ACG GGC ATG CTG GTA C-3'
Ub-R1-4-F4A	F: 5'-TAC CGG CCG CTA CGC CAG CAT GCC CCG-3' R: 5'-CGG GGC ATG CTG GCG TAG CGG CCG GTA-3'
Ub-R1-4-S5A	F: 5'-GCC GCT ACT TCG CCA TGC CCC GCC C-3' R: 5'-GGG CGG GGC ATG GCG AAG TAG CGG C-3'
Ub-R1-4-P7A	F: 5'-TAC TTC AGC ATG GCC CGC CCC TCC C-3' R: 5'-GGG AGG GGC GGG CCA TGC TGA AGT A-3'
Ub-R1-4-R8A	F: 5'-TTC AGC ATG CCC GCC CCC TCC CGC G-3' R: 5'-CGC GGG AGG GGG CGG GCA TGC TGA A-3'
Ub-R1-4-F4Y	F: 5'-CGG CCG CTA CTA CAG CAT GCC CCG CCC CTC CCG CG-3' R: 5'-TCG ACG CGG GAG GGG CGG GGC ATG CTG TAG TAG CGG GTA C-3'
Ub-R1-4-F4W	F: 5'-CGG CCG CTA CTG GAG CAT GCC CCG CCC CTC CCG CG -3' R: 5'-TCG ACG CGG GAG GGG CGG GGC ATG CTC CAG TAG CGG CCG GTA C-3'
Ub-R1-4-R8K	F: 5'-CGG CCG CTA CTT CAG CAT GCC CAA ACC CTC CCG CG-3' R: 5'-TCG ACG CGG GAG GGT TTG GGC ATG CTG AAG TAG CGG CCG GTA C-3'

Table S5. Determination of the intact mass of the Hpg-encoded ubiquitin fusion proteins.

Name	Sequence	Calculated mass (Da)	Found mass (Da)
G4	GGΘGG	8,811.1	8,812.2 ± 1.3
R1-4	GRYFSΘPRPSR	10,159.7	10,159.1 ± 1.7
R1-8	PSLPCΘYSFPD	10,062.6	10,062.1 ± 1.5
R1-9	ALFLPΘSRVHD	10,091.7	10,091.4 ± 1.6
R2-6	DPPPTΘDVPPH	10,008.5	10,008.3 ± 1.7
R3-3	SPPHSΘVPTPA	9,926.4	9,926.1 ± 1.8
R3-8	NVPPLΘGTPSL	9,931.5	9,929.9 ± 0.7
R3-9	FPRPPΘTPPPH	10,079.7	10,079.6 ± 1.9
R3-20	FPTNHΘHRHPT	10,180.7	10,180.8 ± 1.9
R3-24	DGRDAΘSPHYY	10,117.5	10,117.0 ± 1.8
R1-4-5	FSΘPR	9,442.9	9,442.4 ± 1.4
R1-4-3	SΘP	9,139.6	9,139.3 ± 1.3
R1-4-F4A	GRY**A**SΘPRPSR	10,083.6	10,084.1 ± 1.8
R1-4-S5A	GRYF**A**ΘPRPSR	10,143.7	10,143.7 ± 1.5
R1-4-P7A	GRYFSΘ**A**RPSR	10,133.6	10,133.9 ± 1.6
R1-4-R8A	GRYFSΘP**A**PSR	10,074.6	10,075.4 ± 2.0
R1-4-F4Y	GRY**Y**SΘPRPSR	10,175.7	10,175.8 ± 0.8
R1-4-F4W	GRY**W**SΘPRPSR	10,198.7	10,198.6 ± 0.6
R1-4-R8K	GRYFSΘP**K**PSR	10,131.7	10,131.8 ± 1.0
R1-4-R8E	GRYFSΘP**E**PSR	10,132.5	10,134.1 ± 0.7
R1-4-R8H	GRYFSΘP**H**PSR	10,140.5	10,141.9 ± 0.8

Θ = Hpg.

Table S6. Determination of the intact mass of the product derived from the palladium-mediated cross-coupling reaction with fluorescein iodide.

C-terminal peptide sequence	Calculated mass (Da)	Found mass (Da)
G4	9,142.2	9144.1 ± 1.3
R1-4	10,489.4	10,491.8 ± 1.1
R1-8	10,391.3	10,394.4 ± 1.7
R1-9	10,420.9	10,423.6 ± 1.0
R2-6	10,338.1	10,341.0 ± 1.5
R3-3	10,256.5	10,259.2 ± 1.6
R3-8	10,259.9	10,264.0 ± 1.4
R3-9	10,409.6	10,412.0 ± 1.6
R3-20	10,511.1	10513.1 ± 1.4
R3-24	10,447.9	10449.6 ± 1.3
R1-4-5	9,772.5	9,774.7 ± 1.6
R1-4-3	9,469.4	9,471.9 ± 1.5
R1-4-F4A	10,414.0	10,415.5 ± 1.5
R1-4-S5A	10,473.8	10,475.7 ± 1.2
R1-4-P7A	10,464.0	10,465.8 ± 1.7
R1-4-R8A	10,405.5	10,406.9 ± 1.8
R1-4-F4Y	10,505.9	10,508.1 ± 1.0
R1-4-F4W	10,528.8	10,530.9 ± 1.5
R1-4-R8K	10,462.2	10,464.5 ± 0.6
R1-4-R8E	10,464.2	10,465.6 ± 1.0
R1-4-R8H	10,472.0	10,473.5 ± 0.9

Table S7. LC-MS determination of the intact mass of the product derived from the palladium-mediated cross-coupling reaction between Ub-R1-4 and the various aryl/vinyl iodides.

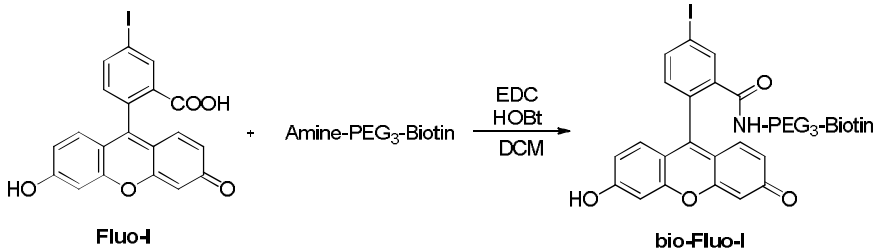
Aryl/vinyl iodide	Calculated mass (Da)	Found mass (Da)
Fluorescein iodide	10,489.4	10,491.8 ± 1.1
p-Methoxyphenyl iodide	10,265.8	10,267.5 ± 1.4
p-*tert*-Butylphenyl iodide	10,291.9	10,293.5 ± 0.8
p-Iodobenzoic acid	10,279.7	10,282.9 ± 1.6
p-Trifluoromethylphenyl iodide	10,303.7	10,305.7 ± 1.4
o-Methylphenyl iodide	10,249.8	10,251.7 ± 1.4
Thiophenyl iodide	10,241.7	10,243.5 ± 1.1
Coumarin iodide	10,333.7	10,334.5 ± 1.6

General methods

Solvents and chemicals were purchased from commercial sources and used directly without further purification. Flash chromatography was performed with SiliCycle P60 silica gel (40-63 μm, 60 Å). ^{1}H NMR spectra were recorded with Inova-300 or -500 MHz spectrometers and chemical shifts were reported in ppm using either TMS or deuterated solvents as internal standards (TMS, 0.00; $CDCl_3$, 7.26; DMSO-d_6, 2.50). Multiplicity was reported as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, brs = broad. ^{13}C NMR spectra were recorded at 75 MHz and chemical shifts were reported in ppm using the deuterated solvents as internal standards ($CDCl_3$, 77.0; DMSO-d_6, 39.5). Electrospray LC-MS analysis was performed using a Finnigan LCQ Advantage IonTrap mass spectrometry coupled with a Surveyor HPLC system. Protein liquid chromatography was performed using a Phenomenex Aeris C4 column (3.6 μm, 200 Å, 2.10 × 100 mm) with a flow rate of 250 μL/min and a linear gradient of 5-50% ACN/H_2O containing 0.1% HCOOH for 10 min. All fluorescence image acquisitions were carried out using a Zeiss LSM-710 confocal microscope equipped with a continuous laser and fluorescence lifetime (FLIM) detectors.

Experimental procedures and characterization data

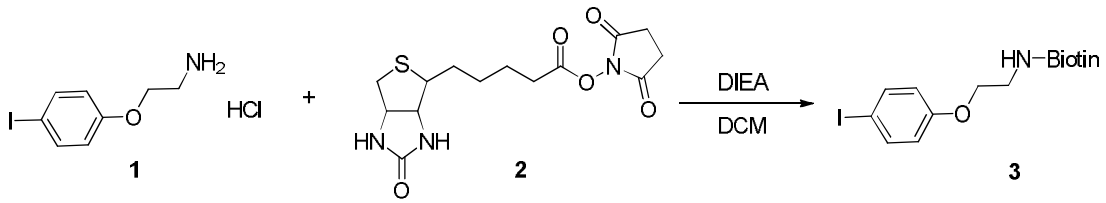
Synthesis of biotinylated fluorescein iodide



Fluo-I bio-Fluo-I

To a solution of **Fluo-I** (4.6 mg) in 5 mL DCM was added 1 equiv of hydroxybenzotriazole (HOBt), 1 equiv of NH_2-PEG_3-biotin, and 1 equiv of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC). The reaction was left at room temperature overnight. After evaporation, crude reaction mixture was purified by column chromatography using DCM/methanol as solvents (40:1, 20:1, 15:1). The pure fraction was dried to afford a yellow solid (3.9 mg, 45%). ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.13 (broad, 2H), 8.09 (d, J =1.5 Hz, 1H), 7.87 (dd, J = 7.8, 2.0 Hz, 1H), 7.81 (t, J =5.6 Hz, 1H), 6.86 (d, J = 7.5 Hz, 1H), 6.61 (d, J = 2.0 Hz, 2H), 6.49-6.47 (m, 2H), 6.43-6.41 (m, 3H), 6.36 (s, 1H), 4.32-4.29 (m, 1H), 4.14-4.11 (m, 1H), 3.45-3.41 (m, 5H), 3.18-3.07 (m, 8H), 2.97 (t, J =5.8 Hz, 2H), 2.81 (dd, J = 12.5, 4.5 Hz, 1H), 2.58 (d, J = 12.5 Hz, 1H), 2.06 (t, J = 7.2 Hz, 2H), 1.61-1.27 (m, 6H); ^{13}C NMR (75 MHz, DMSO-d_6) δ 172.56, 172.54, 165.88, 163.16, 159.27, 152.93, 152.40, 132.73, 131.40, 129.28, 126.35, 113.01, 108.96, 105.99, 102.74, 70.12, 69.94, 69.86, 69.60, 67.27, 64.21, 61.48, 59.63, 55.86, 38.88, 35.52, 28.62, 28.47, 25.69; MS (ESI): [M+H^+] 859.2 (calc.), 859.3 (found).

Biotinylated phenyl-iodide



Compound **1** was synthesized following the literature report.[2,3] To a solution of D-biotin (240 mg, 1.0 mmol) in DMF was added *N*-hydroxysuccinimide (1.2 equiv) and EDC (1.2 equiv). The reaction was allowed to proceed at room temperature overnight. The resulting mixture was filtered and the residue was washed with water and methanol to afford the NHS-activated biotin **2** as a white solid (163.8 mg, 48% yield)): ^{1}H NMR (300 MHz, DMSO-d_6) δ 6.41 (s, 1H), 6.35 (s, 1H), 4.31-4.27 (q, 1H), 4.15-4.10 (s, 1H), 3.12-3.06 (m, 1H), 2.87-2.79 (m, 5H), 2.71-2.58 (m, 3H), 1.68-1.54 (m, 3H), 1.52-1.36 (m, 3H); ^{13}C NMR (75 MHz, DMSO-d_6) δ 175.46, 174.13, 167.88, 110.00, 66.18, 64.37, 60.43, 35.20, 33.03, 32.77, 30.64, 29.51; MS (ESI) calcd for $C_{14}H_{19}N_3O_5SNa$ 364.1 [M+Na$^+$], found 364.1.

Subsequently, to a solution of **2** (102 mg, 0.30 mmol) and **1** (108 mg, 0.36 mmol) in 10 mL DCM was added DIEA (312 μL, 1.80 mmol). The reaction was allowed to proceed overnight. The resulting mixture was diluted using DCM and then washed sequentially with 2 N HCl, H_2O, and brine. The organic layer was separated, dried over anhydrous Na_2SO_4, and concentrated to afford **3** as a white solid (65.7 mg, 45% yield): ^{1}H NMR (300 MHz, DMSO-d_6) δ 8.02 (t, J = 5.7 Hz, 2H), 7.59-7.54 (m, 2H), 6.80-6.75 (m, 2H), 6.37 (d, J = 17.1 Hz, 2H) 4.29-4.25 (m, 1H), 4.09-4.04 (m, 1H), 3.93 (t, J = 5.4 Hz, 2H), 3.40-3.34 (q, 2H), 2.81-2.75 (q, 1H), 2.55 (d, J = 12.3 Hz, 1H), 2.06 (t, J = 7.5 Hz, 2H), 1.63-1.36 (m, 4H), 1.33-1.21 (m, 2H); ^{13}C NMR (75 MHz, DMSO-d_6) δ 172.84, 163.13, 158.79, 138.42, 117.75, 83.66, 66.93, 61.44, 59.62, 55.86, 38.45, 35.49 28.56, 28.45, 25.69; MS (ESI) calcd for $C_{18}H_{24}N_3IO_3S$ 489.1 [M+H$^+$], found 489.2.

Phage viability assay

Equal amounts of M13KE phages (10^7 *Pfu*) in tris-buffered saline (TBS) were loaded to the wells of a 96-well microtiter plate and were either treated with palladium catalyst, **Pd-cat-1**, or not (100 μL total volume per well) at 37 °C for 30 min. Afterwards, all phage in each well were titered by following New England Biolab's protocol[4] using mid-log phase, freshly grown ER 2738 cells.

Construction of phage model systems

Complementary synthetic oligonucleotides encoding either MMPGMM (myc-M4) or THDYPMPGANP (myc-PMP) bearing the restriction sites overhangs (*Sac*I and *Spe*I) were first annealed and then ligated using T4 DNA ligase (Invitrogen) into pComb3HSS phagemid vector predigested with *Sac*I/*Spe*I (Invitrogen) at 16 °C overnight. After heat inactivation at 65 °C for 10 min, 1-2 μL of the ligation mixture was transformed into XL1-Blue cells through heat-shock method (42 °C for 45 s). Cells were recovered using SOC medium with incubation at 37 °C for

1 h. An aliquot was plated on LB-ampicillin plate (100 µg/mL) for overnight incubation at 37 °C. Plasmids extracted from colonies were sent for DNA sequencing to confirm the presence of the desired phagemid.

Hpg-encoded phage propagation of single phagemid *via* methionine biosynthesis inhibition

A single colony of XL1-Blue cells harboring the desired phagemid was grown in LB medium (2 mL) supplemented with ampicillin (100 µg/mL) at 37 °C, 280 rpm, for 8 h. Next, cells were spun down and resuspended in 1 mL of SelenoMet plus nutrient mix (methionine-deficient) medium supplemented with 1g/L dextrose and ampicillin. A 100-µL aliquot was taken and added to a 100 mL of the same medium (1000x dilution) followed by incubation at 37 °C, 250 rpm, for 15-17 h (or until OD_{600} reaches 1.0). The morning after, a corresponding dilution to 0.5 OD_{600} was made using SelenoMet plus nutrient mix medium supplemented with ampicillin (total volume of culture was 100 mL). After this, the inhibitor cocktail consisting of 100 mg/L of Lys, Phe, Thr and 50 mg/L of Iso, Leu, Val, was added. After 15-20 min incubation at 37 °C, 1 mM of HPG and 1 mM IPTG were also added to the culture. Phage packaging and propagation was then carried out by addition of M13KO7 helper phage (1×10^{10} phage/mL, final concentration) and the temperature was lowered to 30 °C. After 90 min of infection, kanamycin (10 µg/mL) was added and left the culture for incubation at 200 rpm for additional 16.5 h. Phage purification was carried out following NEB's protocol.

Quantification of phage by UV absorbance or by phage titering

The amount of phage is quantified using two methods: by UV absorbance at 268 nm and by phage titering. For application where phage infectivity is not of main concern such as western blot analysis, we measure phage amount by UV (phage/mL); otherwise, we employ phage titering (during selection). However, we observed that for similar amount of Met-encoded phage and HPG-encoded phage, the latter tend to give higher UV absorbance reading presumably due to possible absorbance contribution of HPG that sometimes results to unequal loading.

For phage titering, 10 µL each of the diluted phages was mixed with 90 µL of freshly grown XL1-blue cells ($OD_{600} \sim 0.5$) and incubated at 37 °C, 200 rpm, for 20 min. Each dilution was plated on LB-ampicillin plate (100 µg/mL) and incubated overnight at 37 °C. Colonies are counted the morning after to estimate the infectivity titer of the phage measured in *cfu*/mL.

Cu-free Sonogashira cross-coupling reaction on phage

Part 1a. Preactivation of the biotinylated fluorescein iodide, ***bio-Fluo-I*** *(Method A).* In a 0.6-mL microcentrifuge tube was added 142 µL of water, 2 µL of 10 mM biotinylated fluorescein iodide, **bio-Fluo-I**, (in DMSO), 2 µL of 80 mM sodium ascorbate, and 2 µL of 10 mM **Pd-cat-2**. The mixture was heated to 37 °C and reacted for 1 h with vigorous stirring and was later used for the solution-phase panning.

Part 1b. Preactivation of the biotinylated fluorescein iodide, ***bio-Fluo-I*** *(Method B).* To a 1.50-mL glass vial containing 6.36 mg of sodium carbonate (0.06 mmol) was added 100 µL of DMSO, 100 µL of 200 mM *N,N*-dimethyl-2-amino-4,6-dihydroxypyrimidine (**Pd-cat-3**, 0.02 mmol in DMSO), and 50 µL of 200 mM $Pd(OAc)_2$ (0.01 mmol in DMSO). The mixture was stirred at 65 °C for 30 min. Then, a second batch of reagents was added into the same tube consisting of 100 µL of ***bio-Fluo-I*** (0.01 mmol in DMSO), 50 µL of 200 mM ascorbic acid (0.01 mmol in DMSO), and 100 µL DMSO, and continued stirring for another 30 min to obtain the pre-activated reagent cocktail. Note that the result of western blot analyses were the same regardless of the pre-activation method used (data not shown).

Part 2. Several pairwise set-ups of Met-encoded phage and Hpg-encoded phage (for anti-myc, anti-pIII, anti-biotin, and silver staining) containing approximately 10^{11} phage/mL (as measured by UV-Vis) in 200 µL PBS were reacted with 20 µM of the biotinylated fluorescein iodide-palladium complex (8 µL of 500 µM) via the Cu-free Sonogashira cross-coupling reaction at 37 °C, 300 rpm, for 30 min. One set-up for Hpg-encoded phage (for anti-biotin) was also reacted with the 'preactivated' reagent that was prepared in the absence of $Pd(OAc)_2$. After the reaction, one portion of the samples was analyzed by western blot analyses while the other portion was subjected to phage ELISA (continued below).

Western blot analysis and silver staining

After the reaction, Met/Hpg-encoded phages were precipitated with 20% PEG/NaCl (1/5 of the volume) at room temperature for 15 min, and then centrifuged at 13,500 rpm at 4 °C for 10 min. Next, phage pellets were resuspended in 25 µL PBS with addition of 6 µL 6x SDS loading buffer. After boiling the phages at 95 °C for 5 min, all phage lysates were loaded in SDS-PAGE gel, along with the pre-stained protein molecular weight marker (Fermentas). After running the SDS-PAGE (200V for 30-35 min), the gel was incubated in 1x transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol) for 30-40 min. Then, the gel was sandwiched to a PVDF membrane (Millipore) for transferring using a semi-dry blotting apparatus at 15 V for 40 min. Next, the membrane was blocked with 1x casein in PBS for 1-2 h at room temperature then rinsed 3× with PBST (0.1% Tween) within 15-min period. Afterwards, the membrane was cut with its individual protein ladder and incubated separately with the corresponding antibody. One of the membranes was incubated with the mouse-anti-pIII antibody (NEB, 1:3333 dilution) for 1.5 h at rt with rocking then rinsed again for 3x with PBST (0.1% Tween) before incubation with the alkaline-phosphatase (AP)-conjugated anti-mouse secondary antibody (1:5000, Santa Cruz Biotechnology) for at least 2 h at rt. After the final washing, the membrane was soaked in 0.1 M Tris (pH 9.5) for 5 min without agitation and subsequently, excess liquid from the membrane was removed before addition of the AP substrate (BCIP/NBT-Blue, SIGMA) to visualize the blue/purple band. For the anti-biotin blot, membrane was incubated with Vectastain (Vector Lab) right after the blocking following the manufacturer's protocol and directly stained with AP substrate. For anti-myc blot, 1:2500 dilution was used for overnight incubation at 4 °C followed

by incubation with 1:5000 AP-conjugated anti-mouse secondary antibody for 30 min at room temperature and similarly stained with the AP substrate until a blue/purple band was observed.

For the silver staining, the gel was stained right after the SDS-PAGE analysis using the silver staining kit (Invitrogen), following the manufacturer's suggested protocol.

Model phage selection experiment (Phage ELISA)

After the reaction, Met/Hpg-encoded phages were precipitated with 20% PEG/NaCl (1/5 of the volume) at room temperature for 15 min, and then centrifuged at 13,500 rpm at 4 °C for 10 min. Next, Met- and Hpg-encoded phage pellets were each resuspended in 200 µL PBST (0.1% tween 20) and separately added to each well of clear streptavidin-coated microtiter plate that was blocked with 5 mg/mL BSA in 0.1 M $NaHCO_3$ (pH 8.6) for overnight at 4 °C and washed 6x with PBST (0.05%). Affinity capture was done for 40-50 min at room temperature with rocking followed by another 10x PBST (0.1%) washing. Next, we added 200 µL of HRP-conjugated anti-p8 antibody (1:3000) in 1X casein PBS to each well and incubated further for 1 h at room temperature with rocking. Following the washing, ABTS (2,2-azino-bis(3-ethylbenzothiazoline)-6-sulfonic acid-diammonium salt) substrate was added and the mixture was incubated for 20-60 min before reading the absorbance at 416 nm.

Construction of ubiquitin-peptide fusion expression vector

Using our previous construct pQE80L-PUbF[2] as the template (expresses ubiquitin protein with additional MGG sequence at its C-terminus), we replaced the GGMGG sequence at the C-terminus with *Kpn*I and *Sal*I restriction sites by PCR using the pair of primers: FWD - 5'- TGG TGC TGC GTC TGC GCG GTA CCA TGG TCG ACT AAT AAA AGC TTA ATT AGC TGA GCT TGG -3' and REV - 5'- CCA AGC TCA GCT AAT TAA GCT TTT ATT AGT CGA CCA TGG TAC CGC GCA GAC GCA GCA CCA -3' (*Kpn*I and *Sal*I sites are underlined) to obtain the ubiquitin expression vector, pQE80L-Ub-KS, used in subsequent studies (see below).

Efficiency of methionine replacement by Hpg via methionine biosynthesis inhibition

A single colony of XL1-Blue cells harboring the plasmid expressing Ub- G4 was grown in LB medium (2 mL) supplemented with ampicillin (100 µg/mL) at 37 °C, 280 rpm, for 8 h. Next, cells were spun down and resuspended in 1 mL of SelenoMet plus nutrient mix (methionine-deficient) medium supplemented with 1g/L dextrose and ampicillin. A 100-µL aliquot was taken and added to a 100 mL of the same medium (1000x dilution) followed by incubation at 37 °C, 250 rpm, for 15-17 h (or until OD_{600} reaches 1.0). The morning after, a corresponding dilution to 0.5 OD_{600} was made using SelenoMet plus nutrient mix medium supplemented with ampicillin (total volume of culture was 100 mL). After this, the inhibitor cocktail consisting of 100 mg/L of Lys, Phe, Thr and 50 mg/L of Iso, Leu, Val, was added. After 15 min incubation at 37 °C, 1 mM of Hpg and 1 mM IPTG were also added. Cells were allowed to express the Hpg-encoded protein at 250 rpm for 6-8 h.

Construction of the various ubiquitin mutant expression plasmids

All mutants for alanine scanning mutagenesis study of R1-4 were constructed via site-directed mutagenesis using the primers listed in Supplementary Table 4 with Ub-R1-4 as the template. The rest of the mutants were constructed by annealing the complementary oligonucleotides (encoding the desired peptide, Supplementary Table 4) bearing the restriction sites as overhangs (*Kpn*I and *Sal*I). The annealed oligonucleotides were then ligated to the purified pQE80L-Ub-KS cut at *Kpn*I and *Sal*I sites. All newly constructed plasmids were transformed into M15 methionine-auxotroph cells via heat-shock method (42 °C, 45 s). Cells were recovered with SOC medium, incubated at 37 °C for 1 h, and plated on LB-ampicillin plate, followed by overnight incubation at 37 °C. All inserts were confirmed by DNA sequencing.

Expression, purification, His_6-tag cleavage, and characterization of Met and Hpg-encoded ubiquitin fusion proteins

A single colony of bacteria expressing the desired ubiquitin fusion protein was grown in LB medium at 37 °C overnight. This starter culture was used to inoculate the larger culture (100× dilution) in SelenoMet plus nutrient mix medium (methionine-deficient) supplemented with 1 mM methionine and 100 µg/mL ampicillin. After reaching OD_{600} 0.8-1.0, cells were harvested by centrifugation and washed 3x with 0.9% NaCl by repeated re-suspension and centrifugation. Afterwards, the cell pellet was resuspended in SelenoMet plus nutrient mix medium supplemented with ampicillin, and incubated at 37 °C, 250 rpm, for 40-60 min. Then, into the flask was added 1 mM Hpg and 1 mM IPTG. Protein expression was carried out at 25 °C, 250 rpm, for 12 h. Cells were pelleted by centrifugation, resuspended in binding buffer (300 mM NaCl, 50 mM Na_2HPO_4, 10 mM imidazole, pH 8.0), and lysed by sonication on ice. Cell lysate was clarified by centrifugation at 4 °C and loaded to a chromatography column (Bio-Rad) containing HisPur Ni-NTA resin (Thermo Scientific) followed by 2 h incubation at room temperature with rocking. Resin was then washed with washing buffer 3-4x (300 mM NaCl, 50 mM Na_2HPO_4, 50 mM imidazole, pH 8.0). Finally, His_6-tagged fusion protein was eluted from the column by passing 4-mL of elution buffer (300 mM NaCl, 50 mM Na_2HPO_4, 250 mM imidazole, pH 8.0).

For the expression of the wild-type ubiquitin, a single colony of bacteria expressing the desired ubiquitin fusion was grown in LB medium at 37 °C, 280 rpm, overnight. This starter culture was used to inoculate the larger culture (100× dilution) in LB medium supplemented with 100 µg/mL ampicillin. Upon reaching OD_{600} 0.6-0.8, 1 mM IPTG was added to induce protein expression and was carried out at 37 °C, 280 rpm, for 4-5 h. From here on, cells were treated in a similar manner as the Hpg-encoded proteins.

After purification, buffer exchange was carried out with PreScission protease cleavage buffer (50 mM Tris-HCl, pH 7.0, 150 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, 4 °C) using 3-kDa MWCO centrifugal filter (Millipore). Concentrated protein was then transferred to a 1.7-mL

microcentrifuge tube and an appropriate amount of the PreScission protease (GE Healthcare) was added. Cleavage was carried out for 2-3 days with gentle agitation at 4 °C. Once cleavage is completed (as monitored by LC-ESI/MS), protein sample was loaded to a 50-kDa MWCO centrifugal filter (Millipore) with addition of 50 mM Na_2HPO_4 (pH 8.0) or PBS up to 4 mL. The flow-through was collected this time and re-subjected to 3-kDa MWCO centrifugal filter (Millipore) for concentration. Most proteins was pure (>95%) at this point; otherwise, the protein was subjected to fast protein liquid chromatography purification. Molecular weights (see below) as well as purity of the protein sample were assessed by LC-ESI/MS analysis whereas protein concentration was determined by Bradford Assay (BioRad) using bovine serum albumin (BSA) as the standard. Final yields of cleaved fusion proteins range from 0.5-1 mg/L.

Validation of reactivities of the selected peptides in Sonogashira cross-coupling reaction

Part 1. Preparation of the activated reagent cocktail. Preparation of stock solutions: 200 mM $Pd(OAc)_2$ was prepared by dissolving 22.4 mg of $Pd(OAc)_2$ in 500 µL DMSO; 200 mM *N*,*N*-dimethyl-2-amino-4,6-dihydroxypyrimidine (MADHP) was prepared by dissolving 18.6 mg MADHP in 600 µL DMSO; 100 mM fluorescein iodide was prepared by dissolving 4.6 mg fluorescein iodide in 100 µL DMSO; and 200 mM ascorbic acid was prepared by dissolving 35.2 mg of ascorbic acid in 1 mL DMSO. To a 1.50-mL glass vial containing 6.36 mg of sodium carbonate (0.06 mmol) was added 100 µL of DMSO, 100 µL of MADHP (0.02 mmol), and 50 µL of $Pd(OAc)_2$ (0.01 mmol). The mixture was stirred at 65 °C for 30 min. Then, a second batch of reagents was added into the same tube consisting of 100 µL of fluorescein iodide (0.01 mmol), 50 µL of ascorbic acid (0.01 mmol), and 100 µL DMSO, and continued stirring for another 30 min to obtain the pre-activated reagent cocktail. Prior to the cross-coupling reaction, a 40x dilution in DMSO was made to afford a 500 µM solution.

Part 2. Sonogashira cross-coupling reaction with ubiquitin fusion proteins. 2.5 µM of the protein (in 48 µL PBS) in a 0.6-mL microcentrifuge tube was made to react with 8 equiv of the pre-activated fluorescein-iodide-palladium complex (2 µL of 500 µM) at 37 °C for 30 min with stirring. Afterwards, the reaction was quenched with 10 µL of 3-mercaptopropanoic acid (4% v/v in water) and directly injected to LC-MS for analysis.

Reaction specificity of the palladium-mediated cross-coupling with Hpg

To confirm that the alkylation product with fluorescein iodide was a result of the reaction with Hpg and not to the nearby serine residue, 2.5 µM of Ub-R1-4-Met [mass: 10,181.62 (calcd); 10,181.29 ± 1.8 (found)] was reacted with 8 equiv of the preactivated fluorescein-iodide-palladium complex in PBS at 37 °C for 30 min. Afterwards, the reaction was quenched with 10 µL of 3-mercaptopropanoic acid (4% v/v in water) and directly injected to LC-MS for analysis. No peaks corresponding to the expected alkylation product were observed, only the starting material and its slight impurity.

Construction of R1-4EGFR-EGFP plasmid

R1-4 peptide was introduced between A21 and S22 of EGFR in the pcDNA3-EGFR-EGFP$^{Ampr/Neor}$ plasmid using a pair of primers: FWD 5'-CTG GCT GCG CTC TGC CCG GCG GGC CGC TAC TTC AGC ATG CCA CGC CCA TCC CGC AGT CGG GCT CTG GAG GAA AAA-3 and REV 5'-TTT TTC CTC CAG AGC CCG ACT GCG GGA TGG GCG TGG CAT GCT GAA GTA GCG GCC CGC CGG GCA GAG CGC AGC CAG-3'. After PCR using standard protocol (with enhancer), the reaction mixture was analyzed by agarose gel electrophoresis. After desired band was observed, the rest of the PCR reaction mixture was digested with 1 µL of *Dpn*I (Promega) at 37 °C for 3 h. Subsequently, 3 µL of the digested mixture was transformed using XL1-Blue cells by heat-shock method. After recovery, cells were plated on LB-ampicillin (100 µg/mL) plate and incubated overnight at 37 °C. Colonies harboring the desired plasmid were confirmed by DNA sequencing and was subsequently used for the mammalian cell labeling studies.

References

1 Carlos F. Barbas, I., Dennis R. Burton, Jamie K. Scott, Gregg J. Silverman. *Phage Display: A Laboratory Manual* (2004).

2 Li, N., Lim, R. K. V., Edwardraja, S. & Lin, Q. *J. Am. Chem. Soc.* **133**, 15316-15319 (2011).

3 Nguyen, D. P. *et al. J. Am. Chem. Soc.* **131**, 8720-8721 (2009).

4 *Ph.D.™ Phage Display Libraries Instruction Manual*, New England Biolabs (2009).