<Supporting Information>

A Fragment-Based Method to Discover Irreversible Covalent Inhibitors of Catalytic Cysteines

Stefan G. Kathman, Ziyang Xu, and Alexander V. Statsyuk*

Department of Chemistry, Center for Molecular Innovation and Drug Discovery, Chemistry of Life Processes Institute, Northwestern University, Silverman Hall, 2145 Sheridan Road, Evanston, Illinois 60208

Table of Contents

I. Supplemental Chemical Synthesis S1

II. Supplemental Methods S14

III. Supplemental Figures S16

IV. Supplemental Tables S61
**Chemical Synthesis**

**General Information**

Methanol (ACS grade), ethyl acetate (ACS grade), chloroform (ACS grade), toluene (ACS grade), and diethyl ether (ACS grade), acetonitrile (HPLC grade), and hexanes (ACS grade) were purchased from Fisher Scientific and used without further purification. Dichloromethane, tetrahydrofuran and dimethylformamide were purified by passing over activated alumina. Commercially available reagents were used without further purification. Unless otherwise specified, all reagents were purchased from Sigma-Aldrich. Reactions were monitored by thin-layer chromatography (TLC) on pre-coated glass backed plates (60 Å silica gel, 0.25mm, Whatman), and components were visualized by UV light (254 and 365 nm) or by treating the plates with anisaldehyde, KMnO₄, and ninhydrin stains followed by heating. Flash column chromatography was performed over ultra pure silica gel (230–400 mesh) from Silicycle. ¹H and ¹³C NMR spectra were obtained on a Bruker AVANCE III 500 MHz spectrometer or an Agilent DDR2 400 MHz spectrometer (Funded by NSF CHE-1048773, 2010). Chemical shifts were reported in ppm relative to the residual solvent peak (CDCl₃ or DMSO-d₆). Multiplicity was indicated as follows: s (singlet); d (doublet); t (triplet); q (quartet); m (multiplet); dd (doublet of doublets); ddd (doublet of doublet of doublets); dt (doublets of triplets); td (triplet of doublets); brs (broad singlet). Coupling constants were reported in Hz. Small molecule ESI-MS was performed on an Agilent 1100 MSD quadropole instrument. For compounds tested in enzymatic assays, purity was confirmed by analytical HPLC on a Shimadzu LC-6AD instrument with a Restek Pinnacle C18 column with UV detection at 220nm with a 5→95% acetonitrile/water gradient, 0.1% trifluoracetic acid.

**Synthesis of 1a-c**

![Chemical Structure](image_url)

Aniline, p-methoxyaniline, or p-nitroaniline (1.07 mmol) was dissolved in THF (0.1 M, 10.7 mL) and cooled to 0°C with stirring. Diisopropylethylamine (1.4 mL, 8.58 mmol) was then added, followed by acryloyl chloride (175 µL, 2.14 mmol). After 5 min., the reaction was warmed to 23°C and stirred for 1 hour. TLC showed a full conversion to product. THF was evaporated under reduced pressure, and the residue was dissolved in 20mL dichloromethane and washed with a saturated aqueous solution of NaHCO₃ (2× 20 mL). The organic layer was dried with magnesium sulfate, filtered, and evaporated under reduced pressure. The residue was purified by flash chromatography with an ethyl acetate/hexanes gradient 25% EtOAc → 100% EtOAc.
**1a** (108 mg, 68% yield) $^1$H NMR (500 MHz, CDCl$_3$) δ 7.52 (d, J = 7.9 Hz, 2H), 7.35 – 7.23 (m, 3H), 7.06 (t, J = 7.4 Hz, 1H), 6.38 (dd, J = 16.9, 1.3 Hz, 1H), 6.19 (dd, J = 16.8, 10.2 Hz, 1H), 5.71 (dd, J = 10.3, 1.2 Hz, 1H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 163.53 , 137.71 , 131.16 , 129.07, 127.90 , 124.56. 

**1b** (166 mg, 87% yield) $^1$H NMR (500 MHz, CDCl$_3$) δ 7.52 (d, J = 8.9 Hz, 2H), 7.21 (s, 1H), 6.90 (d, J = 9.0 Hz, 2H), 6.45 (dd, J = 16.8, 1.3 Hz, 1H), 6.26 (dd, J = 16.8, 10.2 Hz, 1H), 5.78 (dd, J = 10.3, 1.3 Hz, 1H), 3.83 (s, 3H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 163.38 , 156.55 , 131.14 , 130.79 , 127.50 , 121.76 , 114.18 , 55.49. 

**1c** (44.4 mg, 22% yield) $^1$H NMR (500 MHz, CDCl$_3$) δ 8.17 (d, J = 9.1 Hz, 2H), 7.72 (d, J = 9.1 Hz, 2H), 7.43 (s, 1H), 6.45 (dd, J = 16.8, 1.0 Hz, 1H), 6.21 (dd, J = 16.8, 10.3 Hz, 1H), 5.83 (dd, J = 10.4, 1.0 Hz, 1H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 164.59, 144.55 , 143.20 , 130.47 , 128.98 , 124.92 , 119.23 .

**Synthesis of 2a-c**

Aniline, p-methoxyaniline, or p-nitroaniline (1.07 mmol) was dissolved in CH$_2$Cl$_2$ (0.1M, 10.7 mL), and cooled to 0°C with stirring. 2-chloroethane sulfonyl chloride (112 µL, 1.07 mmol) was then added, followed by triethylamine (150 µL, 1.07 mmol). After 1h of reaction time, a second equivalent of triethylamine (150 µL, 1.07 mmol) was added and the reaction was warmed to 23°C. After one hour TLC showed full conversion of the starting material to product, and the reaction was quenched with 20 mL water and extracted 2×20mL dichloromethane. The combined organic layers were washed with 20 mL 1M HCl and 20 mL saturated aqueous sodium chloride. The organic phase was then dried over magnesium sulfate, filtered, and evaporated under reduced pressure. Purified by flash column chromatography with a CH$_3$OH/CH$_2$Cl$_2$, CH$_3$OH gradient 0→5%.

**2a** (107.6 mg, 55% yield) $^1$H NMR (400 MHz, CDCl$_3$) δ 7.37 – 7.27 (m, 2H), 7.20 – 7.06 (m, 3H), 6.56 (s,1H), 6.55 (dd, J = 16.5, 9.9 Hz, 1H), 6.27 (d, J = 16.5 Hz, 1H), 5.94 (d, J = 9.9 Hz, 1H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 136.15 , 135.02 , 129.53 , 128.57 , 125.42 , 120.97.

**2b** (160 mg, 70% yield) $^1$H NMR (500 MHz, CDCl$_3$) δ 7.17 (d, J = 8.9 Hz, 2H), 6.88 (d, J = 8.9 Hz, 2H), 6.56 (dd, J = 16.5, 9.9 Hz, 1H), 6.28 (s, 1H), 6.22 (d, J = 16.6 Hz, 1H), 5.96 (d, J = 9.9 Hz, 1H), 3.82 (s, 3H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 158.03 , 134.96 , 128.44 , 128.27 , 125.12 , 114.61 , 55.50.
2c (64.8 mg, 26% yield) $^1$H NMR (500 MHz, CDCl$_3$) δ 8.24 (d, $J = 9.1$ Hz, 2H), 7.27 (d, $J = 9.1$ Hz, 2H), 6.64 (dd, $J = 16.4$, 9.8 Hz, 1H), 6.49 (d, $J = 16.5$ Hz, 1H), 6.15 (d, $J = 9.8$ Hz, 1H). $^{13}$C NMR (126 MHz, Chloroform-d) δ 143.55 , 143.33 , 135.05 , 128.84 , 125.35 , 117.82.

Synthesis of 3a-c

General Synthesis Scheme

Synthesis of 109

(±)-3-amino-1,2-propanediol (11.29g, 124 mmol) was dissolved in CH$_2$Cl$_2$:CH$_3$OH (1:5) (1M) and triethylamine (2mL, 14.7 mmol) was added. Di-tert-butyl dicarbonate (32.5g, 149 mmol) was dissolved in dichloromethane (0.8M, 186 mL) and added slowly to the reaction mixture. The resulting reaction was stirred at 23°C for 2h, followed by TLC analysis that showed a full consumption of the starting material. The reaction mixture was evaporated under reduced pressure, and the residue was purified by column chromatography with EtOAc:Hexanes 1:4, then dried on high vacuum to yield 109 as a white solid (23.7g, 94% yield). $^1$H NMR (500 MHz, CDCl$_3$) δ 5.28 – 4.96 (m, 1H), 3.83 – 3.73 (m, 1H), 3.60 (qd, $J = 11.7$, 4.9 Hz, 2H), 3.44 (s, 1H), 3.27 (dt, $J = 12.9$, 6.0 Hz, 2H), 1.46 (s, 9H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 157.45 , 80.13 , 71.37 , 63.58 , 28.35 , 27.42.

Synthesis of 110

109 (10 g, 52mmol) was suspended in H$_2$O (0.6M, 87.2mL) and the flask was covered in foil (to protect NaIO$_4$ from light). NaIO$_4$ (13.4g, 62.8 mmol) was then added and the reaction was stirred for 1h. A white precipitate had formed after 1h, and TLC analysis showed full consumption of the starting material. The precipitate was filtered off, and the aqueous layer was extracted with CHCl$_3$ (8×50 mL). The organic layer was dried with MgSO$_4$, filtered, and evaporated to yield 110 as a yellow oil, which was used immediately without further purification (7.7g, 93% yield). $^1$H NMR (500 MHz, CDCl$_3$) δ 9.68 (s, 1H), 5.23 (s, 1H), 4.10 (d, $J = 5.2$ Hz, 2H), 1.47 (s, 9H). $^{13}$C NMR (126 MHz, Chloroform-d) δ 157.45 , 80.13 , 71.37 , 63.58 , 28.35 , 27.42.
Synthesis of 111

Sodium hydride (60% dispersion in mineral oil) (1.9 g, 46.6 mmol) in tetrahydrofuran (0.17 M, 274 mL) was cooled to 0°C, then triethylphosphonoacetate (8.5 mL, 46.6 mmol) in THF was added dropwise. The reaction was stirred at 0°C for 20 min, then 110 (7.4 g, 46.6 mmol) in THF was added. The reaction was allowed to warm to 23°C and was stirred for 1h. TLC showed a full consumption of the starting materials and conversion to product. THF was removed under reduced pressure, and the residue was then diluted with ethyl acetate (200mL) and water (200 mL). The layers were separated, followed by the extraction of the aqueous layer with EtOAc (2×100 mL). The organic layer was then dried over MgSO₄, filtered, and evaporated. The residue was purified by flash column chromatography with an ethyl acetate/hexanes gradient 25% EtOAc → 50% EtOAc to yield 111 (6.6g, 66% yield). ¹H NMR (500 MHz, CDCl₃) δ 6.94 (dt, J = 15.7, 4.8 Hz, 1H), 5.97 (dt, J = 15.8, 1.9 Hz, 1H), 4.73 (s, 1H), 3.95 (t, J = 5.6 Hz, 2H), 3.76 (s, 3H), 1.48 (s, 9H). ¹³C NMR (126 MHz, CDCl₃) δ 166.55, 145.26, 120.71, 79.73, 60.37, 51.58, 41.28, 28.30.

Synthesis of 5

111 (6.6 g, 30.8 mmol) was dissolved in trifluoroacetic acid (47 mL, 617 mmol) and stirred at 23°C for 30 min. TLC at 30 min showed conversion to product. TFA was evaporated and azeotroped with toluene (2×100mL). The residue was then dried on high vacuum for 2 hours, dissolved in 2 mL methanol and dropped into ice cold diethyl ether (200 mL). The ether was then filtered to collect 5 as the TFA salt (6.2 g, 88% yield). ¹H NMR (500 MHz, DMSO-d₆) δ 8.08 (s, 3H), 6.86 (dt, J = 15.9, 5.6 Hz, 1H), 6.15 (dt, J = 16.0, 1.7 Hz, 1H), 3.70 (s, 3H), 3.70 (d, J = 1.8 Hz, 2H). ¹³C NMR (126 MHz, DMSO-d₆) δ 165.33, 140.61, 123.22, 51.72.

Synthesis of 3a-c

Benzoic acid, p-methoxybenzoic acid, or p-nitrobenzoic acid (0.35 mmol) was dissolved in dimethylformamide (0.2M, 1.75 mL), then 5 (42.6mg, 0.35 mmol), HBTU (128mg, 0.34 mmol), and HOBT (51.8 mg, 0.38 mmol) were added, followed by diisopropylethylamine (175 µL, 1.047 mmol). The reaction was stirred at 23°C for 16h. TLC at 16h showed conversion to product. The reaction was quenched with H₂O (5mL) and extracted with DCM (3×5mL). The combined organic layers were washed with 1M HCl (10mL), saturated aqueous NaHCO₃ (10mL), and saturated aqueous NaCl (10mL). The organic layer was dried over MgSO₄, filtered, and evaporated. Purification with flash column chromatography with CH₃OH/CH₂Cl₂ (CH₃OH gradient 0→5%) yielded 3a (65.6 mg, 86% yield) ¹H NMR (500 MHz, CDCl₃) δ 7.93 – 7.71 (m, 2H), 7.62 – 7.52 (m, 1H), 7.53 – 7.37 (m, 2H), 7.03 (dt, J = 15.7, 5.1 Hz, 1H), 6.39 (s, 1H), 6.02 (dt, J = 15.7, 1.9 Hz, 1H), 4.43 – 4.18 (m, 2H), 3.76 (s, 3H). ¹³C NMR (126 MHz, CDCl₃) δ 167.43, 166.42, 144.13, 133.85, 131.87, 128.72, 126.96, 121.61, 51.75, 40.61.
**3b** (75.7 mg, 87% yield) $^1$H NMR (500 MHz, CDCl$_3$) δ 7.79 (d, J = 8.8 Hz, 2H), 7.04 (dt, J = 15.7, 5.1 Hz, 1H), 6.97 (d, J = 8.8 Hz, 2H), 6.22 (s, 1H), 6.02 (dt, J = 15.7, 1.9 Hz, 1H), 4.28 (ddd, J = 6.1, 5.1, 1.9 Hz, 2H), 3.89 (s, 3H), 3.76 (s, 3H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 166.88, 166.46, 162.43, 144.39, 128.79, 126.08, 121.50, 113.87, 55.47, 51.73, 40.57.

**3c** (59.9 mg, 65% yield) $^1$H NMR (400 MHz, CDCl$_3$) δ 8.30 (d, J = 8.7 Hz, 2H), 7.94 (d, J = 8.7 Hz, 2H), 6.98 (dt, J = 15.7, 5.3 Hz, 1H), 6.31 (s, 1H), 5.99 (d, J = 15.7 Hz, 1H), 4.28 (td, J = 5.7, 1.9 Hz, 2H), 3.74 (s, 3H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 166.25, 165.42, 149.78, 143.19, 139.37, 128.24, 123.97, 122.17, 51.86, 40.87.

**Synthesis of 4**

**General Scheme**

---

**Synthesis of 112**

Sodium hydride (60% dispersion in mineral oil) (233.2 mg, 5.83 mmol) in tetrahydrofuran (0.17M, 34.3 mL) was cooled to 0°C with stirring, followed by the dropwise addition of diethyl(methylsulfonylmethyl)phosphonate (Oakwood) (1342.2 mg, 5.83 mmol) in 5 mL THF. The reaction was stirred at 0°C for 20 min, then **110** (928 mg, 5.83 mmol) in 5 mL THF was added. The reaction was allowed to warm to 23°C and was stirred for 1h. TLC showed conversion to the product. THF was removed under reduced pressure, and the residue was then diluted with ethyl acetate (30 mL) and water (30 mL). The layers were separated, and the aqueous layer was extracted with EtOAc (2×30 mL). The organic layer was then dried over MgSO$_4$, filtered, and evaporated. The residue was purified by flash column chromatography with an ethyl acetate/hexanes gradient 25% EtOAc → 50% EtOAc to yield **112** (530.4 mg, 56% yield). $^1$H NMR (400 MHz, CDCl$_3$) δ 6.38 (dd, J = 11.7, 5.8 Hz, 1H), 6.28 (dt, J = 11.4, 1.7 Hz, 1H), 5.01 (s, 1H), 4.23 (td, J = 6.3, 1.7 Hz, 2H), 3.00 (s, 3H), 1.41 (s, 9H).

**Synthesis of 113**

**112** (530.4 mg, 2.26 mmol) was dissolved in trifluoroacetic acid (3.45 mL, 45.1 mmol) and stirred at 23°C for 30 min. TLC at 30 min showed conversion to product. Trifluoroacetic acid was evaporated off and azeotroped with toluene (2×30 mL). The residue was then dried on high
vacuum for 2 hours, dissolved in 1 mL methanol and dropped into ice cold diethyl ether (100 mL). The resulting mixture was filtered to collect 113 as the TFA salt (435 mg, 77% yield). $^1$H NMR (500 MHz, DMSO-d$_6$) δ 8.18 (s, 1H), 6.99 (dt, J = 15.4, 1.7 Hz, 1H), 6.75 (dt, J = 15.4, 5.5 Hz, 1H), 3.76 (dd, J = 5.3, 1.7 Hz, 2H), 3.06 (s, 3H). $^{13}$C NMR (126 MHz, DMSO-d$_6$) δ 137.95 , 133.13 , 48.56 , 42.02 .

**Synthesis of 4a-c**

Benzoic acid, p-methoxybenzoic acid, or p-nitrobenzoic acid (0.2 mmol) was dissolved in dimethylformamide (0.2M, 1mL), then 113 (50mg, 0.2 mmol), HBTU (73.8 mg, 0.16 mmol), and HOBT (29.8 mg, 0.22 mmol) were added, followed by DIPEA (100.7 µL, 0.6 mmol). The reaction was stirred at 23°C for 16h. TLC at 16h showed conversion to product. The reaction was quenched with H$_2$O (5mL) and extracted with CH$_2$Cl$_2$ (3×5mL). The combined organic layers were washed with 1M HCl (10mL), saturated aqueous NaHCO$_3$ (10mL), and saturated aqueous NaCl (10mL). The organic layer was dried over MgSO$_4$, filtered, and evaporated. Purified by flash column chromatography with a CH$_3$OH/CH$_2$Cl$_2$, CH$_3$OH gradient 0→5% to yield 4a (30.5 mg, 64% yield). $^1$H NMR (500 MHz, CDCl$_3$) δ 7.83 (dt, J = 7.1, 1.4 Hz, 2H), 7.74 – 7.56 (m, 1H), 7.56 – 7.44 (m, 2H), 7.04 (dt, J = 15.2, 4.5 Hz, 1H), 6.56 (d, J = 15.1 Hz, 1H), 6.48 (t, J = 5.8 Hz, 1H), 4.37 (ddd, J = 6.2, 4.5, 1.9 Hz, 2H), 2.98 (s, 3H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 167.46 , 144.22 , 133.31 , 132.21 , 130.08 , 128.84 , 126.98 , 42.85 , 39.88.

4b (31.5 mg, 58% yield) $^1$H NMR (500 MHz, CDCl$_3$) δ 7.80 (d, J = 8.8 Hz, 2H), 7.04 (dt, J = 15.2, 4.4 Hz, 1H), 6.99 (d, J = 8.8 Hz, 2H), 6.55 (dt, J = 15.2, 2.0 Hz, 1H), 6.33 (t, J = 6.1 Hz, 1H), 4.35 (ddd, J = 6.2, 4.5, 1.9 Hz, 2H), 3.90 (s, 3H), 2.98 (s, 3H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 166.95 , 162.67 , 144.58 , 129.91 , 128.87 , 125.51 , 113.98 , 55.51 , 42.86 , 39.82.

4c (9.2 mg, 16% yield) $^1$H NMR (500 MHz, CDCl$_3$) δ 8.37 (d, J = 8.7 Hz, 2H), 8.01 (d, J = 8.7 Hz, 2H), 7.03 (dt, J = 15.2, 4.8 Hz, 1H), 6.58 (dt, J = 15.3, 1.9 Hz, 1H), 6.52 (s, 1H), 4.40 (ddd, J = 6.4, 4.8, 1.9 Hz, 2H), 3.00 (s, 3H). $^{13}$C NMR (126 MHz, CDCl$_3$) δ 165.92 , 149.72 , 144.05 , 139.07 , 129.84 , 128.53 , 123.74 , 42.63 , 39.90.
Structures of 6-108:
Characterization of compounds tested in enzymatic assays

\[ \text{M+Na} \]: 357.1 Da. HPLC purity: 95%.

\[ \text{M+H} \]: 301.1 Da. HPLC purity: 97%.
(s, 3H), 0.94 (d, J = 6.6 Hz, 6H). $^{13}$C NMR (126 MHz, CDCl$_3$) $\delta$ 173.93, 171.03, 166.31, 144.04, 121.52, 51.72, 43.11, 42.12, 40.03, 25.84, 22.76. [M+H]: 311.2 Da. HPLC purity: 98%.

$^{1}$H NMR (500 MHz, CDCl$_3$) $\delta$ 7.24 (d, J = 3.8 Hz, 1H), 6.80 (dt, J = 15.7, 5.1 Hz, 1H), 6.74 (dd, J = 3.8, 0.9 Hz, 1H), 6.08 – 5.96 (m, 1H), 5.81 (d, J = 15.7 Hz, 1H), 4.97 (t, J = 6.6 Hz, 1H), 4.18 – 3.99 (m, 2H), 3.89 (dd, J = 7.6, 5.9 Hz, 1H), 3.74 (dd, J = 7.5, 1.0 Hz, 1H), 3.56 (s, 3H), 2.34 – 2.04 (m, 1H), 1.93 – 1.69 (m, 4H). $^{13}$C NMR (126 MHz, CDCl$_3$) $\delta$ 166.38, 161.88, 153.36, 144.00, 136.10, 128.68, 123.75, 121.66, 68.64, 51.70, 40.47, 34.77, 25.84. [M+Na]: 318.1 Da. HPLC purity: 95%.

$^{1}$H NMR (500 MHz, CDCl$_3$) $\delta$ 7.55 – 7.00 (m, 5H), 6.72 (dt, J = 15.7, 5.2 Hz, 1H), 6.07 (d, J = 7.8 Hz, 1H), 5.94 (s, 1H), 5.69 (dt, J = 15.9, 1.7 Hz, 1H), 4.60 (td, J = 8.0, 6.2 Hz, 1H), 4.04 – 3.80 (m, 2H), 3.71 (s, 3H), 3.21 – 2.78 (m, 2H), 1.98 (s, 3H). $^{13}$C NMR (126 MHz, CDCl$_3$) $\delta$ 170.85, 170.20, 166.27, 143.28, 136.39, 129.18, 128.87, 127.25, 121.61, 54.83, 51.66, 40.03, 38.24, 23.21. [M+Na]: 327.1 Da. HPLC purity: 95%.

$^{1}$H NMR (500 MHz, CDCl$_3$) $\delta$ 7.33 – 7.19 (m, 1H), 7.11 (dt, J = 15.8, 5.0 Hz, 1H), 6.52 (d, J = 8.4 Hz, 1H), 6.13 (dt, J = 15.9, 1.9 Hz, 1H), 4.74 (td, J = 8.5, 5.9 Hz, 1H), 4.24 (dt, J = 5.8, 3.0 Hz, 2H), 3.95 (s, 3H), 2.23 (s, 3H), 1.89 (tt, J = 13.2, 6.2 Hz, 2H), 1.84 – 1.71 (m, 1H), 1.17 (dd, J = 11.5, 6.1 Hz, 6H). $^{13}$C NMR (126 MHz, CDCl$_3$) $\delta$ 172.34, 170.54, 166.42, 143.91, 121.38, 51.69 (d, J = 7.5 Hz), 40.86, 40.05, 24.80, 23.12, 22.82, 22.25. [M+Na]: 293.1 Da. HPLC purity: 95%.

$^{1}$H NMR (500 MHz, CDCl$_3$) $\delta$ 7.51 – 7.12 (m, 5H), 6.80 (dt, J = 15.7, 5.0 Hz, 2H), 6.61 (d, J = 8.0 Hz, 1H), 5.76 (dd, J = 15.8, 2.0 Hz, 1H), 4.78 (q, J = 7.6 Hz, 1H), 4.09 – 3.86 (m, 2H), 3.76 (s, 3H), 3.10 (dd, J = 7.5, 1.9 Hz, 2H), 2.00 (s, 3H). $^{13}$C NMR
(126 MHz, CDCl₃) δ 171.24, 170.32, 166.35, 143.55, 136.46, 129.22, 128.70, 127.11, 121.40, 54.75, 51.63, 39.99, 38.56, 23.10. [M+Na]: 327.1 Da. HPLC purity: 99%.

Supplemental methods

Dialysis experiments (Figure S5)

200 µL of papain and compound 6-8 adducts were prepared separately as described above in “irreversible tethering screening assay” by incubating papain (10 µM) with 100 µM of compounds 6-8 in 1% DMSO 50mM HEPES 150mM NaCl 0.1 mM EDTA pH 7.5. The adducts were loaded into Slide-A-Lyzer dialysis cassettes (Thermo). The cassettes were dialyzed against 500mL of 50mM HEPES 150mM NaCl 0.1mM EDTA pH 7.5 for 16h, then 20 µL were extracted and analyzed by ESI-MS. The buffer was then replaced with fresh buffer and samples continued and switched again after another 24 h, with 20 µL samples extracted for MS analysis prior to each buffer switch.

Recombinant expression of GST-HRV3C protease in E. coli

GST-HRV3C protease in a PGEX4T vector plasmid (GST-UbcH7) was transformed into Rosetta (DE3)pLysS cells (Millipore). 1L LB media containing 100µg/ml ampicillin was inoculated with 50 mL overnight cell culture and incubated at 37°C until OD reached ~0.5. Then, IPTG (0.5 mM final concentration) was added to the cell culture media at 28°C, followed by 5 hour incubation at the same temperature. Cells were then harvested and lysed by sonication in phosphate buffered saline (PBS) with 1mM DTT and 1mM PMSF. The supernatant was incubated with glutathione agarose beads (Pierce Biotechnology) for 1 hour at 4°C. The beads were washed three times with PBS + 1mM DTT + 1mM PMSF. The protease was then eluted with 100 mM Tris pH 8.0, 100 mM NaCl, 10 mM GSH (reduced), 1 mM DTT. The pooled fractions were then dialyzed three times (300, 400, 300 mL) versus 50 mM Tris Ph 8.5, 150 mM NaCl, 5 mM DTT, 20% glycerol.

Recombinant expression of UbcH7 in E. coli

UbcH7 in a PGEX6P1 vector plasmid (GST-UbcH7) was transformed into BL21 cells (Novagen). 1L LB media containing 100µg/ml ampicillin was inoculated with 50 mL overnight cell culture and incubated at 37°C until OD reached ~1.2. Then, IPTG (1.0 mM final concentration) was added to the cell culture media at 30°C, followed by 4 hour incubation at the same temperature. Cells were then harvested and lysed by sonication in phosphate buffered saline (PBS) with protease inhibitors (Complete Mini Protease Inhibitor Cocktail, Roche). The supernatant was incubated with glutathione agarose beads (Pierce Biotechnology) for 1 hour at 4°C. The beads were washed three times with PBS and incubated with PreScission Protease overnight at 4°C to elute UbcH7 (50mM HEPES, 150 mM NaCl, 0.1 mM EDTA).
Recombinant expression of USP08 catalytic domain in *E. coli*

USP08 catalytic domain in a PET21a-LIC vector plasmid (6×His-USP08, Addgene) was transformed into BL21 (DE3) cells (Invitrogen). 1L TB media containing 100µM kanamycin and 600µl antifoam 204 (Sigma A-8311) was inoculated with 50ml overnight culture and incubated at 37°C until OD reached ~3. Then, IPTG (100 µM final concentration) was added to the cell culture media at 15°C. The culture was incubated overnight at the same temperature. Cells were then harvested and lysed by sonication in 10 mM Tris-HCl pH 7.0, 0.5 M NaCl 5% glycerol 2 mM imidazole 1 mM β-mercaptoethanol 0.1 µM PMSF. The cleared lysate was then loaded onto TALON metal-affinity beads at 4°C. Beads were washed three times with 10 mM Tris-HCl, pH 7.0 0.5 M NaCl 5% glycerol 10 mM imidazole 1 mM β-mercaptoethanol 0.05% Tween 20. The protein was then eluted with 10 mM Tris-HCl pH 7.0, 0.5 M NaCl 5% glycerol 200 mM imidazole 1 mM β-mercaptoethanol before being exchanged into 50mM HEPES 150mM NaCl 0.1mM EDTA pH 7.5 with PD10 columns (GE Healthcare). MS analysis of USP08 showed that the resulting protein had a cleaved N-terminal methionine residue, and ~50% of the protein had been further modified by gluconic acid at the N-terminus.

**Supplemental figures**

**Figure S1** ESI-MS of 10 reaction mixtures containing 10 electrophilic fragments each screened against papain as described. See table S2 for list of fragments in each reaction mixture. The measured molecular weight of papain is 23422.56 Da.

Papain sequence (the starred cysteines are internal disulfides), Cys\(^{25}\) is a catalytic cysteine:

IPEYVDWRQKGAVTYPVKNQQGSC\(^{25}\)WAFSAVVTIEGIIKIRTGNLNEYSEQELLDC\(^{DR}\)RSYGCTGGLYVPWSALQLVAQYGIHYRNTYPYEGVQRYC\(^{RSREKGYPYAAKTDGVRQVQ}\)PYNEGALLYSIANQPVSVVLEAAGKDFQLYRGGIFVGPC\(^{GNKVDHAVAAVGYGPNYIL}\)IKNSWGTGWGENGYIRIKRGTGNSYGVC\(^{GLYTSSFYPVKN}\)

Calculated papain MW: 23,422.29 Da
Mix 2
Unmodified papain

Mix 3
Unmodified papain
Mix 4

Unmodified papain

Mix 5

Unmodified papain
Mix 6
Unmodified papain

Mix 7
Unmodified papain
Papain + 7 covalent complex
Mix 8
Unmodified papain

Mix 9
Papain + 6 covalent complex
Unmodified papain
**Figure S2** ESI-MS of the labeling of papain (10 µM) by 6, 7, or 8 (100 µM each, 1h) in the presence of 10mM glutathione (GSH).
6 + 10mM GSH

Papain + 6 covalent complex

Unmodified papain

7 + 10mM GSH

Unmodified papain

Papain + 7 covalent complex
8 + 10mM GSH

Unmodified papain

Papain + 8 covalent complex

Counts vs. Deconvoluted Mass (amu)
Figure S3. A) ESI-MS of papain treated with 100 µM of 6, 7, or 8 for 1h followed by addition of 106 (100µM) and incubation for 1h. Treatment of papain with 100µM of 106 alone is shown for comparison. The 7 + 106 and 8 + 106 7 + 106 spectra do not show separation between the peaks because inhibitors 7 and 8 are too close in MW to 106, but the peak is instead a weighted average of the two peaks. However, in no case did treatment with 6-8 followed by 106 result in dilabeling of papain. B) ESI-MS of papain treated with 100 µM of 106 for 1h, followed by addition of 100µM of 6-8. In no case did treatment with 106 followed by 6-8 result in dilabeling of papain.

A.
Papain + 7 + 106

Papain + 7/106 covalent complex

Unmodified papain

Papain + 8 + 106

Papain + 8/106 covalent complex

Unmodified papain
B.

Papain + $106$

Unmodified papain

Papain + $106$ covalent complex

Papain + $106 + 6$

Papain + $106$ covalent complex
Papain + 106 + 7

Papain + 106 covalent complex

Papain + 106 + 8

Papain + 106 covalent complex
Figure S4 ESI-MS of the labeling of papain (10 µM) by 6, 7, or 8. A) 100 µM each, 1h or B) 1mM each, 1h, zoomed out to show no dilabeling

A.

Papain + 6 (100 µM)

Unmodified papain

Papain + 6 covalent complex

Papain + 7 (100 µM)

Unmodified papain

Papain + 7 covalent complex
Papain + 7 (1 mM)

Papain + 7 covalent complex

Papain + 8 (1 mM)

Papain + 8 covalent complex
Figure S5. Papain+compound 6, Papain+compound 7, Papain+compound 8 covalent adducts after 16h, 40h, and 64h of dialysis as described in the experimental section.
6 + 64h dialysis

7+16h dialysis
7 + 40h dialysis

7 + 64h dialysis
$8 + 16\text{h dialysis}$

$8 + 40\text{h dialysis}$
Figure S6 Pseudo-first order papain inhibition plots at different concentrations of 6, 7, 8, 19, 106, 107, and 108.
Inhibition of papain with **8**

![Graph showing inhibition with 8](image)

Inhibition of papain with **19**

![Graph showing inhibition with 19](image)
Figure S7 ESI-MS spectra after incubating ten reaction mixtures containing 10 electrophilic fragments each with GST tagged human rhinovirus 3C protease for 1h.

GST-HRV3C protease sequence:

MSPIHGYWKIKGLVQPTRLLEELYEEHLYERDEGDKWRNKKFELGLEFPNLPPYYIDGVKLTQSMAIYIADKHKNMGCPKERAISLEGAVLDIRYGVSRIRAYSKDFETLKVDNFLSKEMLKMFEDRLCHKTYLNGHVTTHDFMLYDADTVVLYMDPMCLDAFPKLVCFKKRIEAIQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLVPRGSFEGPNTEFLASLRLRKINMIIITTSKGEFTGLGIDRVCVPIPTAQPGDDVVLVNGQKIRVKDKYKLVDEPNINLELTVLTLDNKEKRDIRGISEDLEGVDATLVHSNNFTNITILEVGPVTMAGLNLSTPTNRMIRYDYATKTQCGVLCATKGIFICHVGGNQRQGFSAQLKKQYFVEKQLERPHRD

Calculated MW: 47566.72 Da
Figure S8 ESI-MS spectra after incubating ten reaction mixtures containing 10 electrophilic fragments each with UbcH7 for 1h.

UbcH7 sequence:

GPLGSMAASRRLMKEEIRKCGMKNFRNIQVDEANLTTWQGLIVPDNPPYDKGAFRIE
INFPAEYPFKPPKTFKTIYHPNIDEKGVCLPVISAENWKPATKTDQVIQSLIALVNDPQ
PHEPLRADLAEEYSKDRKKFCFNAAEFTKKYGEKRFPVD

Calculated MW: 18,272.99 Da
**Figure S9** ESI-MS spectra after incubating ten reaction mixtures containing 10 electrophilic fragments with the catalytic domain of USP08 for 1h. USP08 was partially modified by gluconic acid at the N-terminus during bacterial expression.

USP08 catalytic domain sequence: (50% modification with gluconic acid)-GSSHHHHHSSGL VPRGSPTVTPTVNRKPTCYPKAIESRLSASQRNLNPVFGGSGPALTGLRNGLGNTCYM NSILQCLCNAPHLADYFNRNCYQDDINRSNLGHKGEVAEEFGIIMKALWTGQYRYISP KDFKITIGKINDQFAGYSQDQSQELLFLMDGLHEDLNKADNRKRYKEENNNDLDDFK AAEHAWQKHQLNEISIVALFQGQFKSTVQCLTCHKKRSRTFEAFMYLSLPLASTSKCTL QDCLRLFSEKKLTDNNRFYCSTHRDCARRDSLKIEIWKLPVLLVHLKRFSDGRWKQK LQTSVFPLENLDLSQYVIGPKNNLKYNLFSVSNHYGGGLGGHYTAYCKNAARQRWF KFDHEVSDISVSSVKSSAAYILFYTLG

Calculated MW: 45,218.13 Da, +gluconic acid: 45,396.29 Da

**Untreated USP08**

**Untreated USP08 + gluconic acid**
Unlabeled USP08

Unlabeled USP08 + gluconic acid

Mix 1

Unlabeled USP08

Unlabeled USP08 + gluconic acid

Mix 2
Figure S10. ESI-MS spectra of reaction mixtures containing the catalytic domain of USP08 treated with the mixture of 10 electrophilic fragments for 4h. USP08 was partially modified by gluconic acid at the N-terminus during bacterial expression.
Unlabeled USP08

Unlabeled USP08 + gluconic acid

Mix 5

Unlabeled USP08

Unlabeled USP08 + gluconic acid

Mix 6
Supplemental tables

Table S1: Pseudo-first order rate constants of fragments 6-55.

| Compound | $k$ pseudo-first order | Compound | $k$ pseudo-first order | Compound | $k$ pseudo-first order |
|----------|------------------------|----------|------------------------|----------|------------------------|
| 6        | 0.0007951 s$^{-1}$     | 22       | 0.0007605 s$^{-1}$     | 38       | 0.0007728 s$^{-1}$     |
| 7        | 0.0006978 s$^{-1}$     | 23       | 0.0004979 s$^{-1}$     | 39       | 0.000651 s$^{-1}$      |
| 8        | 0.0004232 s$^{-1}$     | 24       | 0.0005202 s$^{-1}$     | 40       | 0.0004793 s$^{-1}$     |
| 9        | 0.0006824 s$^{-1}$     | 25       | 0.0005202 s$^{-1}$     | 41       | 0.0005635 s$^{-1}$     |
| 10       | 0.0004656 s$^{-1}$     | 26       | 0.0006107 s$^{-1}$     | 42       | 0.0005281 s$^{-1}$     |
| 11       | 0.0007414 s$^{-1}$     | 27       | 0.0006665 s$^{-1}$     | 43       | 0.0007616 s$^{-1}$     |
| 12       | 0.000654 s$^{-1}$      | 28       | 0.0004200 s$^{-1}$     | 44       | 0.0006746 s$^{-1}$     |
| 13       | 0.0003582 s$^{-1}$     | 29       | 0.0004038 s$^{-1}$     | 45       | 0.0003961 s$^{-1}$     |
| 14       | 0.0005016 s$^{-1}$     | 30       | 0.0006579 s$^{-1}$     | 46       | 0.0007806 s$^{-1}$     |
| 15       | 0.0007733 s$^{-1}$     | 31       | 0.0005193 s$^{-1}$     | 47       | 0.0006539 s$^{-1}$     |
| 16       | 0.0006414 s$^{-1}$     | 32       | 0.0006296 s$^{-1}$     | 48       | 0.0004680 s$^{-1}$     |
| 17       | 0.0006156 s$^{-1}$     | 33       | 0.0006306 s$^{-1}$     | 49       | 0.0004975 s$^{-1}$     |
| 18       | 0.0006093 s$^{-1}$     | 34       | 0.0007717 s$^{-1}$     | 50       | 0.0006514 s$^{-1}$     |
| 19       | 0.0004396 s$^{-1}$     | 35       | 0.0005755 s$^{-1}$     | 51       | 0.0005048 s$^{-1}$     |
| 20       | 0.0005603 s$^{-1}$     | 36       | 0.0003400 s$^{-1}$     | 52       | 0.0005984 s$^{-1}$     |
| 21       | 0.0003327 s$^{-1}$     | 37       | 0.0004493 s$^{-1}$     | 53       | 0.0005521 s$^{-1}$     |
| **Average:** | 0.000575841 s$^{-1}$ | **Average:** | 0.000575841 s$^{-1}$ | **Average:** | 0.000575841 s$^{-1}$ |
| **Std Dev:** | 0.000127064 s$^{-1}$ | **Std Dev:** | 0.000127064 s$^{-1}$ | **Std Dev:** | 0.000127064 s$^{-1}$ |
Table S2: Composition of reaction mixes used for irreversible tethering.

| Mix 1 | MW  | Mix 2 | MW  | Mix 3 | MW  | Mix 4 | MW  |
|-------|-----|-------|-----|-------|-----|-------|-----|
| 25    | 234.1004 | 29    | 234.1004 | 76   | 239.1521 | 103  | 223.0957 |
| 53    | 254.0458 | 20    | 254.163  | 73   | 257.0567 | 28   | 249.1001 |
| 56    | 272.1161 | 68    | 273.1113 | 21   | 273.1113 | 59   | 259.0768 |
| 17    | 282.1038 | 60    | 283.0611 | 92   | 285.1113 | 67   | 276.1474 |
| 16    | 288.1474 | 91    | 288.1474 | 41   | 290.1267 | 50   | 286.9906 |
| 95    | 298.1317 | 27    | 298.1317 | 89   | 299.0969 | 44   | 290.1267 |
| 47    | 303.1019 | 86    | 303.1107 | 72   | 304.0882 | 23   | 299.1158 |
| 8     | 310.1893 | 35    | 312.0111 | 55   | 314.1267 | 11   | 315.0219 |
| 102   | 320.1736 | 98    | 323.0561 | 79   | 326.1267 | 88   | 326.163  |
| 51    | 342.1216 | 90    | 342.9878 | 69   | 347.1037 | 54   | 354.0215 |

| Mix 5 | MW  | Mix 6 | MW  | Mix 7 | MW  | Mix 8 | MW  |
|-------|-----|-------|-----|-------|-----|-------|-----|
| 66    | 249.1113 | 57    | 250.0954 | 37   | 251.127 | 39   | 251.127 |
| 36    | 263.1158 | 40    | 266.1267 | 104  | 268.0882 | 64   | 268.1423 |
| 100   | 277.0773 | 61    | 277.1426 | 62   | 277.1426 | 46   | 279.1583 |
| 84    | 287.1158 | 12    | 287.127  | 45   | 287.127  | 48   | 287.127  |
| 101   | 290.1267 | 99    | 291.0831 | 10   | 293.1086 | 19   | 295.0878 |
| 85    | 299.1158 | 49    | 299.127  | 7    | 300.1111 | 71   | 302.0878 |
| 26    | 305.1263 | 14    | 306.1016 | 22   | 308.0289 | 63   | 308.1736 |
| 24    | 315.1219 | 18    | 316.1423 | 43   | 316.1423 | 58   | 316.1423 |
| 38    | 327.0106 | 42    | 327.1583 | 87   | 330.158  | 65   | 334.072  |
| 80    | 357.0688 | 74    | 364.119  | 30   | 364.1787 | 83   | 370.1893 |

| Mix 9 | MW  | Mix 10 | MW  |
|-------|-----|--------|-----|
| 96    | 251.127 | 104   | 240.1474 |
| 32    | 270.1004 | 33    | 254.0458 |
| 70    | 280.1059 | 31    | 280.1423 |
| 94    | 287.127  | 97    | 288.1222 |
| 77    | 296.1736 | 34    | 297.0001 |
| 13    | 302.1267 | 15    | 302.1267 |
| 78    | 310.072  | 52    | 310.1317 |
| 81    | 316.1423 | 75    | 320.1736 |
| 6     | 334.0787 | 93    | 340.0423 |
| 82    | 380.0372 | 9     | 387.0986 |