SUPPLEMENTAL MATERIAL

General Methods

All Fmoc amino acid derivatives and HBTU (2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) were obtained from Novabiochem (San Diego, CA). All other reagents were from Sigma-Aldrich (St. Louis, MO). Analytical RP-HPLC was performed on a Hewlett-Packard 1100 series instrument equipped with a C18 Vydac column (5 μm, 4.6 x 150 mm) at a flow rate of 1 mL/min. Semi-preparative RP-HPLC was performed on a Waters DeltaPrep 4000 system with a Waters 486 tunable detector. All runs used linear gradients of 0.1 % TFA in water (solvent A) and 0.1% TFA 90 % acetonitrile in water (solvent B). Mass spectrometric analysis for all peptides and proteins was performed by ESI-MS on a Sciex-API-100 single quadrupole spectrometer using positive ionization. Protein purifications were carried out on an ÄKTA FPLC system (Amersham Biosciences, Uppsala, Sweden). NMR spectra were taken on a Bruker DPX 400 MHz instrument.

Synthesis of photo-methionine (5) and Fmoc-photo-methionine (6)

*Weinreb Amide (**2**).* To a solution of **1** (10.3 g, 33.9 mmol) in anhydrous CH_2Cl_2 (160 mL) at 0 °C was added *N*-methylmorpholine (9.3 mL, 84.8 mmol) followed by isobutyl chloroformate (5.5 mL, 42.4 mmol). The mixture was stirred for 15 min., and then *N,O*-dimethylhydroxylamine hydrochloride (3.9 g, 40 mmol) was added portion wise. The reaction was then allowed to warm up to r.t. and stirred under inert atmosphere for 2 h. At this time the reaction was diluted with HCl (1 M) and extracted 3 times with CH_2Cl_2. The organic layers were pooled, dried over Na_2SO_4, filtered, and concentrated. Silica gel chromatography (30% EtOAc in hexanes) yielded 11.75 g (99%) of **2** as an oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.16 (s, 1H), 4.12 (s, 1H), 3.67 (s, 3H), 3.18 (s, 3H), 2.64-2.39 (m, 2H), 2.23-2.07 (m, 1H), 2.01-1.82 (m, 1H), 1.46 (s, 9H), 1.44 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.87, 155.82,

82.10, 79.78, 61.51, 54.12, 32.56, 28.62, 28.52, 28.28, 27.98; FAB-HRMS calcd. for $C_{16}H_{30}N_2O_6$ ($M+H^+$) 347.2104, found 347.2178.

*Ketone (**3**).* To a solution of **2** (12.1 g, 44.5 mmol) in toluene (150 mL) at -78 °C under inert atmosphere was added methylmagnesium bromide (34 mL, 3 M solution in hexanes) over 30 min. The reaction was allowed to warm up to -5 °C over 3 h, quenched with HCl (1 M), and extracted 3 times with EtOAc. The organic layers were pooled, dried over Na_2SO_4, filtered, and concentrated. Silica gel chromatography (15% EtOAc in hexanes) yielded 6.9 g (67%) of **3** as a clear oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.05 (s, 1H), 4.13 (s, 1H), 2.68-2.39 (m, 2H), 2.12 (s, 3H), 1.93-1.73 (m, 2H), 1.46 (s, 9H), 1.44 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 207.76, 171.78, 155.81, 82.21, 79.87, 53.78, 39.70, 30.21, 28.59, 28.25, 27.02; FAB-HRMS calc'd. for $C_{15}H_{27}NO_5$ ($M+H^+$) 302.1889, found 302.1961.

*Diazirine (**4**).* A solution of **3** (6.9 g, 22.8 mmol) in liquid ammonia (50 mL) was stirred at reflux (dry ice condenser) for 5 h. At this time, the reaction flask was cooled to -50 °C and hydroxylamine-*O*-sulfonic acid (3.00 g, 26.20 mmol), dissolved in methanol (15 mL,), was added over 30 min. The reaction was then allowed to stir at reflux for 10 h and the ammonia was evaporated. The resulting slurry was filtered and the filter cake was washed with several portions of methanol. The combined washings were concentrated down until no smell of ammonia could be detected. The solution was then diluted with methanol (75 mL) and cooled to 0 °C. TEA (3.2 mL, 22.8 mmol) was added followed by portion wise addition of iodine until a brown color persisted. The reaction was stirred for 1 h at r.t., then concentrated to 30 mL, diluted with brine, and extracted 3 times with EtOAc. Organic layers were pooled, filtered, and concentrated. Silica gel chromatography (10% EtOAc in hexanes) yielded 3.55 g (49%) of **4** as a white powder. ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.98 (s, 1H), 4.13 (s, 1H), 1.76-

1.66 (m, 1H), 1.45 (s, 9H), 1.44 (s, 9H), 1.38-1.25 (m, 2H), 1.01 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.51, 155.52, 81.97, 79.59, 53.59, 30.62, 28.48, 28.12, 27.34, 25.34, 19.78; FAB-HRMS calc'd. for $C_{15}H_{27}N_3O_4$ ($M+H^+$) 314.2002, found 314.2013.

*PhotoMet (**5**)* A solution of **4** (100 mg, 0.32 mmol) in THF (10 mL) and 4 M HCl (10 mL) was stirred at r.t. for 20 h. Evaporation of THF and water gave **5** in quantitative yield that was used without further purification. ^{1}H NMR (400 MHz, D_2O) δ 3.79-3.73 (m, 1H), 1.92-1.77 (m, 2H), 1.60-1.50 (m, 2H), 1.11 (s, 3H); FAB-HRMS calc'd. for $C_6H_{11}N_3O_2$ ($M+H^+$) 158.0851, found 158.0927. For ^{1}H NMR, see Figure S1; $[\alpha]_D^{25°}$ = -98.6º (c 4.5, H_2O).

*Fmoc-PhotoMet (**6**).* $NaHCO_3$ (168 mg, 2.00 mmol) and **5** (200 mg, 1.26 mmol) were dissolved in a minimal amount of water (5 mL). *1,4*-dioxane (10 mL) was added followed by a solution of Fmoc-OSu (420 mg, 1.25 mmol) in *1,4*-dioxane over 15 min. The reaction was stirred at r.t. for 20 h, in the dark. At this time, the reaction was acidified to pH 3 using HCl and extracted three times with EtOAc. Organic layers were combined, dried over Na_2SO_4, filtered, and concentrated. Silica gel chromatography (2.5% MeOH in CH_2Cl_2 with 0.1% acetic acid) yielded 370 mg (97%) as an off-white solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.77 (d, 2H), 7.58 (d, 2H), 7.41 (t, 2H), 7,32 (t, 2H), 5.19 (d, 1H), 4.47-4-41 (d, 2H), 4.41-4.34 (m, 1H), 4.25-4.19 (t, 1H), 1.89-1.77 (m, 2H), 1.63-1.35 (m, 2H), 1.02 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 156.51, 144.17, 143.97, 141.76, 128.21, 127.52, 125.41, 120.45, 67.57, 47.55, 30.81, 27.38, 25.55, 20.04; FAB-HRMS calc'd. for $C_{21}H_{21}N_3O_4$ ($M+H^+$) 380.1532, found 380.1606.

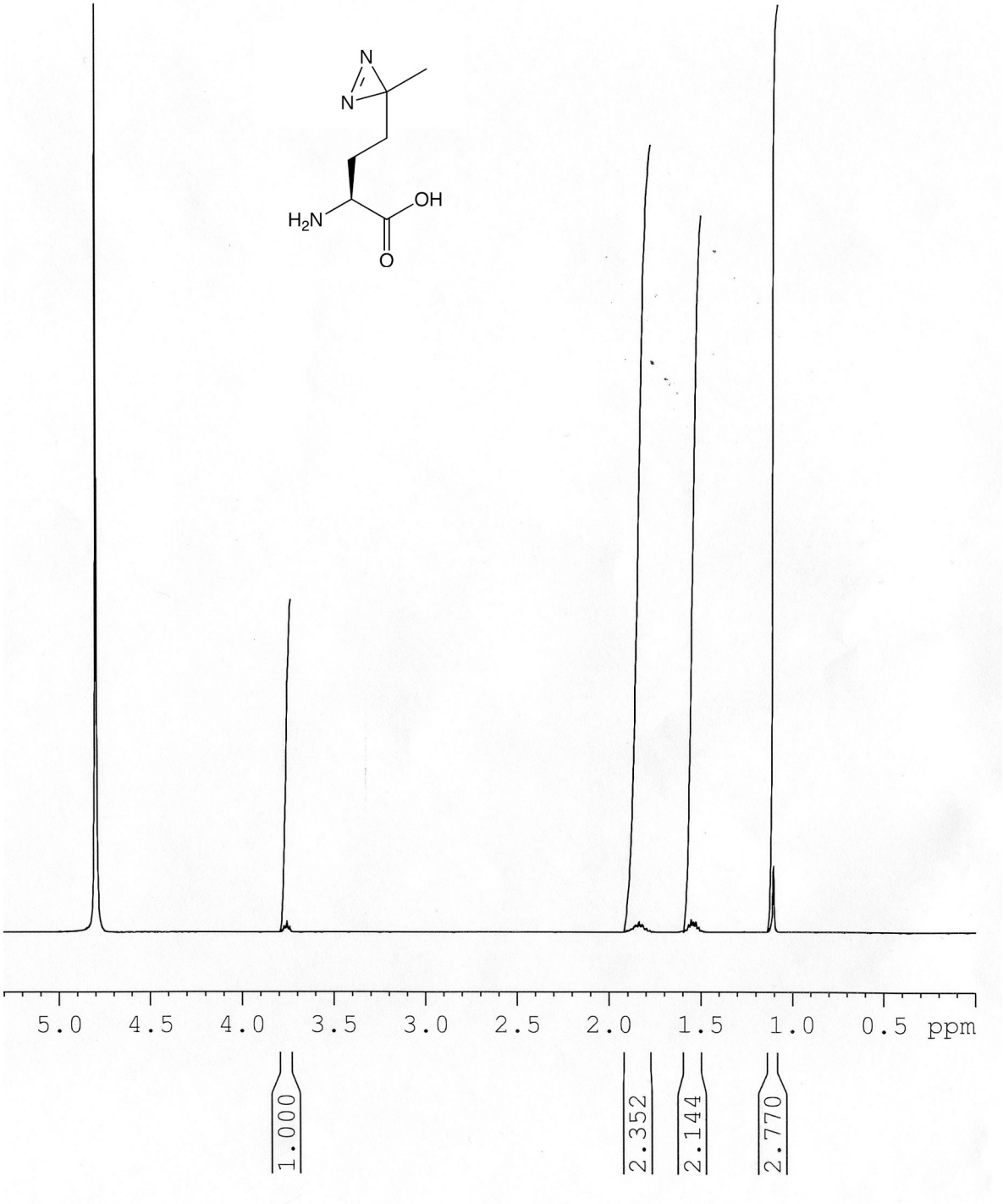

Figure S1: ^{1}H-NMR spectrum of compound **5**

Peptide Synthesis

H-CSpSMpS-OH (**11**) and H-CSpSM-photoMet-OH (**7**) peptides were manually synthesized as C-terminal carboxylic acids on a Wang resin (1.2 mmol/g) on a 0.3 mmol scale using standard Fmoc solid phase peptide synthesis protocols. The following protected amino acid derivatives were used: Fmoc-Cys(Trt)-OH, Fmoc-Ser(*t*Bu)-OH, Fmoc-Ser(PO(OBzl))-OH and Fmoc-photoMet-OH. The first amino acid was coupled as a symmetric anhydride with 0.1 eq of DMAP in DMF and the unreacted hydroxyl groups were capped with acetic anhydride and DIEA as a base. Chain assembly was completed by activation of the amino acids using HBTU and DIEA and couplings were monitored by the Kaiser test [1]. Fmoc removal was carried out by treatment with 20% piperidine / DMF (1 x 1 min followed by 1 x 10 min). Cleavage from the resin was performed by treatment with 95 % TFA, 2.5 % water, 2.5 % ethanedithiol for 2 h at r.t. Peptides were isolated by precipitation with cold Et_2O and centrifugation and taken up in solvent A. RP-HPLC analysis of crude of peptide 7 (Figure S2) revealed a major peak at 23 min with a *m/z* in good agreement with the desired product (found *m/z* 646.0 $M+H^+$, calculated *m/z* 646.2 $M+H^+$) as well as some side products. The two peaks at 28 and 28.5 min correspond to incomplete benzyl deprotection (found *m/z* 736.9 $M+H^+$, calculated *m/z* 736.2 $M+H^+$) and the very minor peak at 31.5 min can be assigned to loss of N_2 from the diazirine (as it does not have the characteristic 350 nm absorption band) and reaction of the resulting carbene species with a benzyl protecting group (found *m/z* 709.9 $M+H^+$, calculated *m/z* 710.2 $M+H^+$). It is unclear whether this last side-product was generated during chain assembly or the cleavage step. Following purification by RP-HPLC, the identities of the two peptides, H-CSpSMpS-OH **11** and H-CSpSM-photoMet-OH **7**, were confirmed by ESI-MS; **11** 673.9 $M+H^+$ (expected 674.1 $M+H^+$) and **7** 646.0 $M+H^+$ (expected 646.2 $M+H^+$) respectively (Figure S3). The stability of the photoMet containing peptide at the ligation conditions was tested by incubating the peptide at 1 mM concentration in ligation buffer (20 mM Tris-HCl, 150 mM

NaCl, 75 mM MESNa, pH 7.8) and monitoring the absorbance of the diazirine moiety at 350 nm. No decrease in absorbance or other changes in the HPLC trace were observed after 24 h indicating the peptide is stable to ligation conditions (Figure S4).

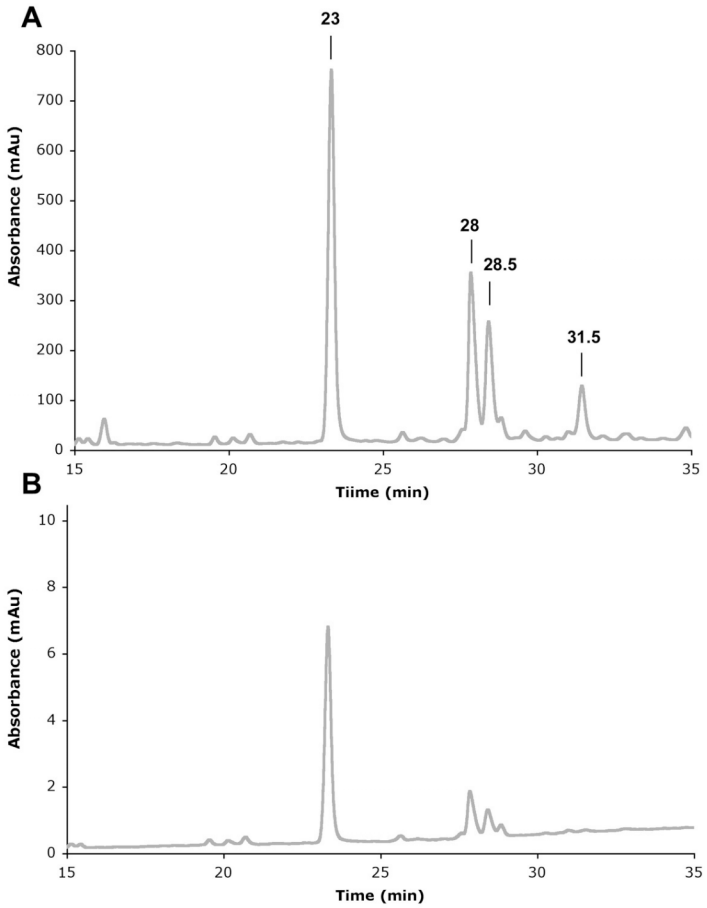

Figure S2: RP-HPLC analysis of crude peptide H-CSpSM-photoMet-OH **7** on a C18 column using a 0 to 20 % B gradient over 30 min. Traces at 214 nm (A) and 350 (B) nm are shown. The peak with a retention time of 23 min corresponds to the desired product **7** (found *m/z* 646.0 $M+H^+$, calculated *m/z* 646.2 $M+H^+$), the two peaks at 28 and 28.5 min correspond to incomplete benzyl deprotection (found *m/z* 736.9 $M+H^+$, calculated *m/z* 736.2 $M+H^+$) and the minor peak at 31.5 min can be assigned to loss of N_2 from the diazirine (as it does not have the characteristic 350 nm absorption band) and the presence of a benzyl protecting group (found *m/z* 709.9 $M+H^+$, calculated *m/z* 710.2 $M+H^+$).

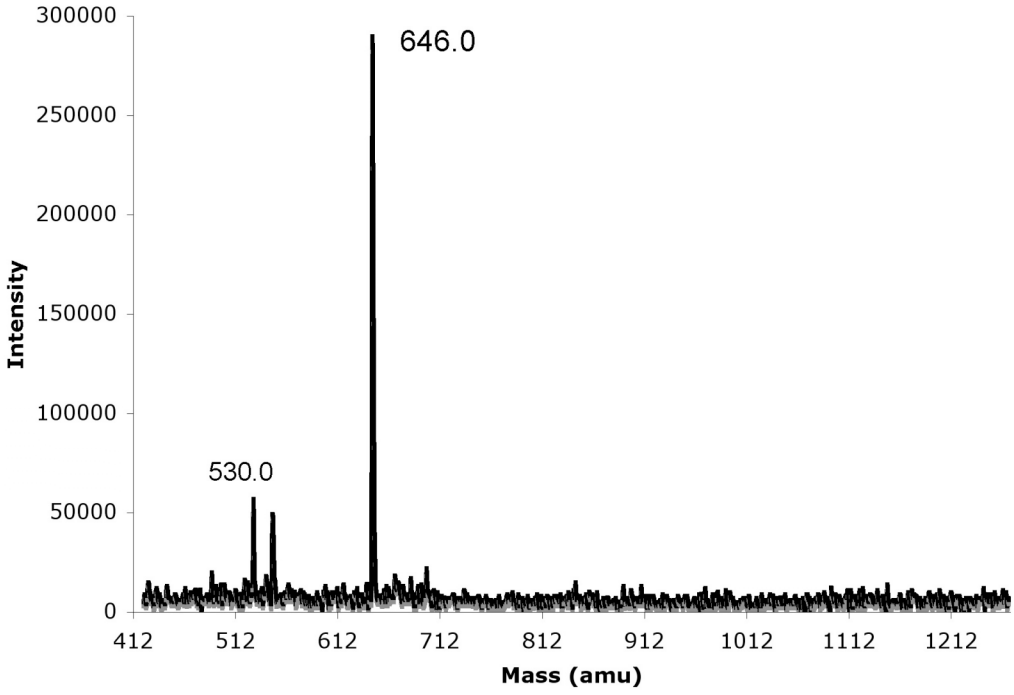

Figure S3: mass spectrum of the purified H-CSpSM-photoMet-OH peptide **7**. The observed mass (646.0 $M+H^+$) is in good agreement with the calculated value (646.2 $M+H^+$).

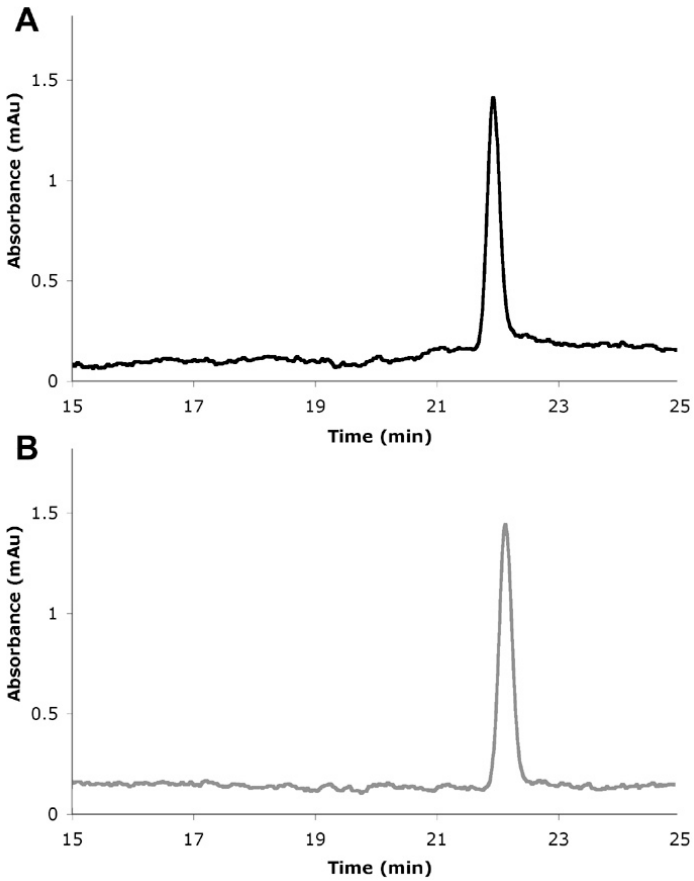

Figure S4: Stability test of peptide H-CSpSM-photoMet-OH **7** in ligation buffer. The purified peptide was analyzed by RP-HPLC as in Figure S2 after 0 (A) and 24 h (B) of incubation in ligation buffer. Traces at 350 nm, characteristic absorbance wavelength of the diazirine group, demonstrate the stability of photoMet containing peptides under EPL reaction conditions.

Protein Expression

HA-MH2 (residues 214-462 from Smad-2) was expressed in *E. coli* as a fusion with the modified intein GyrA and a C-terminal chitin binding domain (CBD) in the pTXB1 vector from New England Biolabs (Beverly, MA) as previously described [2]. Transformed cells were grown in 6 L of LB supplemented with ampicillin (100 μg/mL) at 37 ºC until OD_{600} = 0.6, then expression was induced by addition of 0.5 mM IPTG. Cells were grown for 3 h at 37 ºC and harvested by centrifugation. Cell pellets were resuspended in 30 mL of cold lysis buffer (20 mM Tris, 200 mM NaCl, 200 mM 1,6-hexandiol, pH 7.5) supplemented with Complete protein inhibitor tablets from Roche Diagnostics (Mannheim, Germany). Cells were lysed by means of a French press and the soluble fraction was recovered by centrifugation and further cleared by filtering it through a 5 μm filter. The soluble fraction thus obtained was loaded onto a column containing 6.5 mL of chitin resin, previously equilibrated with lysis buffer, and incubated for 20 h with gentle agitation at 4 ºC. The resin was then washed with 7.5 column volumes (CVs) of lysis buffer, 1.5 CVs of cleavage buffer (100 mM HEPES, 200 mM NaCl, 200 mM 1,6-hexandiol, pH 7.8) and 1 CV of cleavage buffer supplemented with 75 mM 2-mercaptoethansulfonic acid (MESNa). Cleavage of the protein thioester from the GyrA intein and the CBD was accomplished by incubation of the beads with 2 CVs of cleavage buffer with 75 mM MESNa for 24 h. Fractions containing HA-Smad2-MH2 thioester (according to SDS-PAGE) were pooled and analyzed by RP-HPLC and MS to assess the integrity of the thioester. To the Smad2-MH2 thioester were then added 2 molar equivalents of the Smad binding domain of SARA (SARA-SBD, residues 665-721). The Smad2-MH2/SARA-SBD heterocomplex **8** was purified by cation exchange chromatography at pH 6.0 using a 5 mL SP FF column (Amersham Biosciences) using a gradient from 20 mM MES, 10 % glycerol, 10 mM MESNa, pH 6.0 to 20 mM MES, 10 % glycerol, 10 mM MESNa, 1 M NaCl pH 6.0.

Fractions containing the desired complex **8** (as detected by SDS-PAGE and RP-HPLC, Figure S5) were pooled together, exchanged into ligation buffer (20 mM Tris-HCl, 150 mM NaCl, 75 mM MESNa, pH 7.8) and concentrated down to 0.2 mM using a 10 KDa molecular weight cut-off ultrafiltration device from Millipore (Billerica, MA). Protein samples were immediately used in ligation reactions or flash frozen in liquid nitrogen and stored at -80 ºC. This protocol allowed for the production of 8 mg of pure Smad2-MH2 thioester/L of culture in a 1:1 complex with SARA-SBD.

Expressed Protein Ligation (EPL) reactions

EPL reactions were performed at 4 ºC by addition of a 0.2 mM solution of protein complex **8** in ligation buffer to lyophilized peptide **7** or **11** (1 mM final peptide concentration). The reactions were monitored by SDS-PAGE, RP-HPLC and ESI-MS and judged to be complete after 24 - 48 h of incubation (Figure S6). The ligated proteins **9** (MH2-CSpSM-photoMet-OH) and **10** (MH2-CSpSMpS-OH) were then separated from unreacted peptide by size exclusion chromatography (SEC) on a Superdex 200 column from Amersham Biosciences (Uppsala, Sweden) in 20 mM HEPES, 150 mM NaCl, 2 mM DTT, pH 7.5 (Figure S7, S8). Purified proteins were further characterized by ESI-MS (**9**, found *m/z* 26807.16 ± 6.31 $M+H^+$, calculated *m/z* 26803.1 $M+H^+$ and **10**, found *m/z* 26834.0 ± 6.47 $M+H^+$, calculated *m/z* 26830.97 $M+H^+$)

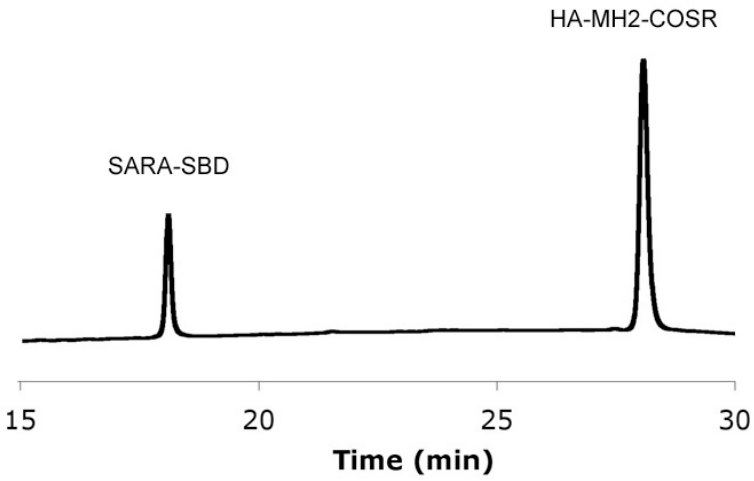

Figure S5: RP-HPLC analysis of the SARA-SBD /HA-MH2 thioester complex **8**.

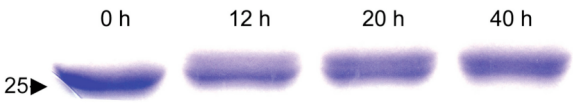

Figure S6: Progress of ligation reaction between **7** and **8** followed by SDS-PAGE. Coomassie stained 12% gel.

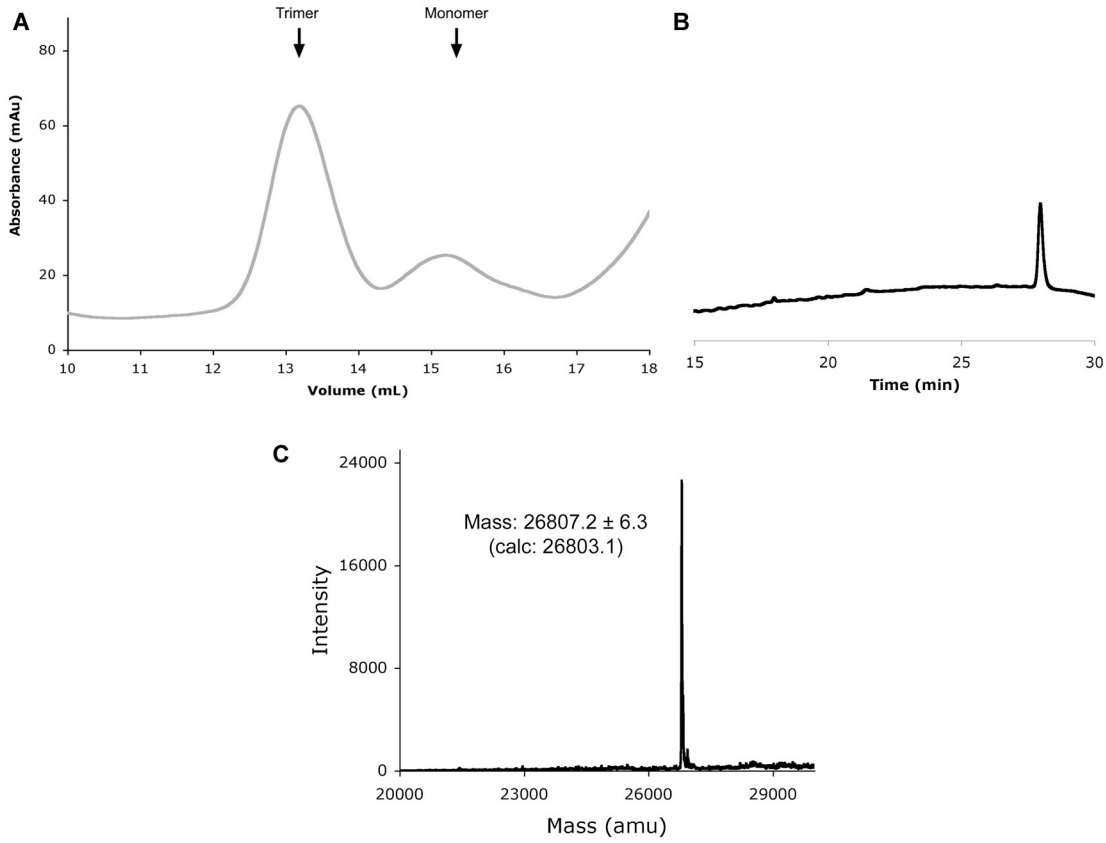

Figure S7: **A**. Size exclusion purification of ligated **9** on a Superdex 200 column. Two main peaks can be observed, one of them elutes at the same volume than the trimeric MH2-CSpSMpS and the other one corresponds to the unligated monomeric protein complexed to SARA-SBD. **B**. RP-HPLC of purified **9** trimer from panel A showing that it no longer contains SARA-SBD. **C**. ESI-MS analysis of protein **9** (found *m/z* 26807.16 ± 6.31 $M+H^+$, calculated *m/z* 26803.1 $M+H^+$).

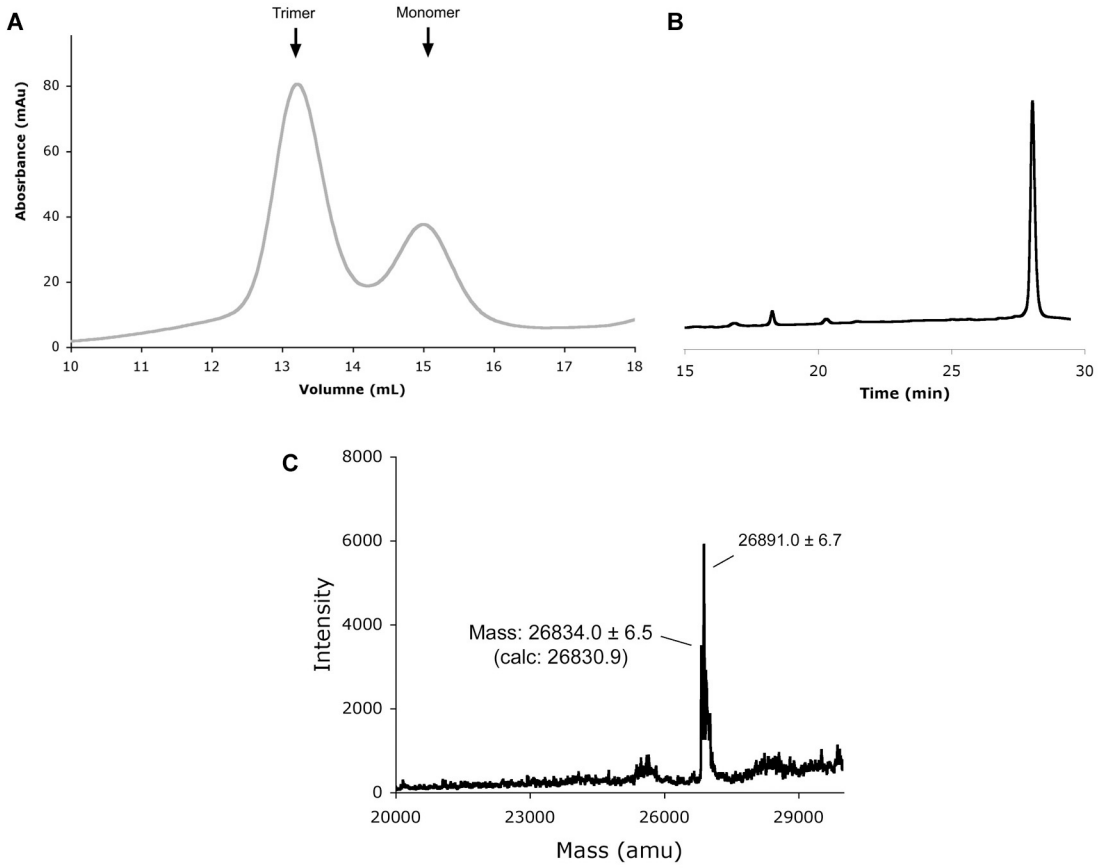

Figure S8: **A**. Size exclusion purification of ligated **10** on a Superdex 200 column. The main peak has the expected elution volume for the homotrimer. **B**. RP-HPLC of purified **10** trimer from panel A showing that it no longer contains SARA-SBD. **C**. ESI-MS analysis of protein **10** (found *m/z* 26834.0 ± 6.47 $M+H^+$, calculated *m/z* 26830.97 $M+H^+$), a peak of *m/z* 26891.0 is also observed and can be assigned to $[M+Fe]^+$ resulting from binding of Fe^{3+} ions to the phosphate groups present in the protein, similar *m/z* peaks have been previously observed in our lab during the preparation of bis-phosphorylated MH2 domains.

Laser induced photo-cross-linking

In order to induce photo-cross-linking, 7.5 μL of each protein sample (at 2 μM) was placed in the beam of a He-Cd laser (325 nm, 4.74 W/cm^2, Kimmon Electric Co., Englewood, CO) and irradiated for 5 seconds at r.t.. Samples were then incubated on ice for 20 min.

SARA-SBD containing samples were prepared by addition of ~ 20 eq. of lyophilized SARA-SBD to protein samples followed by incubation for 2 h at 4 ºC prior laser irradiation. Dephosphorylated samples of **9** and **10** were prepared by treatment of 50 μL of protein solution at 2 μM with 5 μL of Calf Intestinal Phosphatase (New England Biolabs, Beverly, MA) in NEB3 buffer for 30 min at 37 ºC. Complete enzymatic removal of the phosphate group of protein **9** was verified by Western blot analysis using an anti-phospho-Smad2(465/467) rabbit monoclonal antibody (Figure S9) from Cell Signaling Technologies (Denvers, MA).

Western blot analysis and quantification

Protein samples were diluted with 2 X SDS-PAGE loading buffer and resolved by SDS-PAGE before being transferred to Immobilon-P membrane (Millipore) using standard western blotting procedures. Immuno-blotting was performed using anti-HA mouse monoclonal antibody and a fluorescent goat anti-mouse IRDye 800 from Odyssey (Lincoln, NE) as the secondary. Western blot quantification was performed using the Odyssey Infrared Imaging System and the Odyssey v 2.1.10 analysis software from Li-COR.

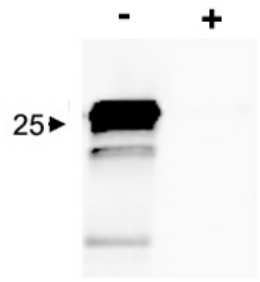

Figure S9: Western blot analysis of protein **9** treated (+) or not treated (-) with Calf Intestinal Phosphatase. An anti-phospho-Smad2 antibody that specifically recognizes phosphorylation at Ser^{465} was used.

REFERENCES

1. Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. Anal Biochem 1970, 34, 595-8.
2. Hahn, M. E.; Muir, T. W. Photocontrol of Smad2, a multiphosphorylated cell-signaling protein, through caging of activating phosphoserines. Angew Chem 2004, 43, 5800-3.