Three Component Mannich-Type Reaction for Selective Tyrosine Bioconjugation

Neel S. Joshi, Leanna R. Whitaker and Matthew B. Francis*

Department of Chemistry, University of California, Berkeley, CA 94720-1460, and Materials Sciences Division, Lawrence Berkeley National Labs, Berkeley, CA 94720-1460

Supporting Information

General Procedures and Materials

Unless otherwise noted, all chemicals were obtained from commercial sources and used without further purification. Analytical thin layer chromatography (TLC) was performed on EM Reagent 0.25 mm silica Gel 60-F_{254} plates with visualization by ultraviolet (UV) irradiation at 254 nm and/or staining with vanillin. Purifications by flash chromatography were performed using EM silica gel 60 (230-400 mesh). The eluting system for each purification was determined by TLC analysis. Chromatography solvents were used without distillation. All reactions were carried out under a nitrogen atmosphere in oven dried glassware unless otherwise noted. All organic solvents were removed under reduced pressure using a rotary evaporator. Pyridine was distilled under an inert atmosphere from calcium hydride. Water (ddH_2O) used in biological procedures or as a reaction solvent was deionized using a NANOpure™ purification system (Barnstead, USA). Ribonuclease A (R 5500) from bovine pancreas, lysozyme (L 6876) from chicken egg white, myoglobin (M 1882) from horse heart, and α-chymotrypsinogen A (P 5380) from bovine pancreas were purchased from Sigma (St. Louis, USA) and used without further purification. Tocinoic acid (T 3149) was purchased from Sigma and used without further purification. Structural analysis of chymotrypsinogen A (entry 2CGAA), myoglobin (entry 1DWR), ribonuclease A (entry 1AFK), and lysozyme (entry 193L) were based on information from the Protein Data Bank.

Instrumentation and Sample Analysis Preparations

IR. Infrared spectra were recorded from thin film samples on a reflecting plate using a Genesis II FTIR™ (ATI Mattson, USA) or from thin film samples on a sodium chloride plate using a Genesis FTIR™ (ATI Mattson, USA).

NMR. ^1H and ^{13}C spectra were measured with a Bruker AVQ-400 (400 MHz) spectrometer or a Bruker DRX-500 (500 MHz) spectrometer as noted. ^1H NMR chemical shifts are reported as δ in units of parts per million (ppm) relative to methanol-d_4 (δ 3.31, septet) or chloroform-d_3 (δ 7.26, singlet). Multiplicities are reported as follows: s (singlet), d (doublet), t (triplet), q (quartet), dd (doublet of

doublets), dt (doublet of triplets), or m (multiplet). Coupling constants are reported as a *J* value in Hertz (Hz). The number of protons (n) for a given resonance is indicated as nH, and is based on spectral integration values. ^{13}C NMR chemical shifts are reported as δ in units of parts per million (ppm) relative to methanol-d_4 (δ 49.0, septet) or chloroform (δ 77.2, triplet).

Mass Spectrometry. Fast Atom Bombardment (FAB) and Electron Impact (EI) mass spectra were obtained at the UC Berkeley Mass Spectrometry Facility. Matrix assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed on a Voyager-DE™ system (PerSeptive Biosystems, USA). All samples were co-crystallized using a sinapinic acid solution (10 mg/mL in 1:1 MeCN:ddH_2O with 0.1% TFA). Electrospray LC/MS analysis was performed using an API 150EX system (Applied Biosystems, USA) equipped with a Turbospray source and an Agilent 1100 series LC pump. Protein chromatography was performed using a Jupiter 5u C5 300Å reversed phase column (2.0 mm x 150 mm) with a MeCN:ddH_2O gradient mobile phase containing 0.1% formic acid (250 μL/min). Protein mass reconstruction was performed on the charge ladder with Analyst software (version 1.3.1, Applied Biosystems). Prior to MS analysis, biological samples were desalted and/or separated from small molecule contaminants using μC18 ZipTip® pipet tips (Millipore, USA), NAP-5™ gel filtration columns (Amersham Biosciences, USA), Strata C-18E™ reversed-phase columns (Phenomenex, USA), or 3,500 molecular weight cutoff Slide-A-Lyzer® Dialysis Cassettes (Pierce Biotechnology, Inc., USA) as indicated below.

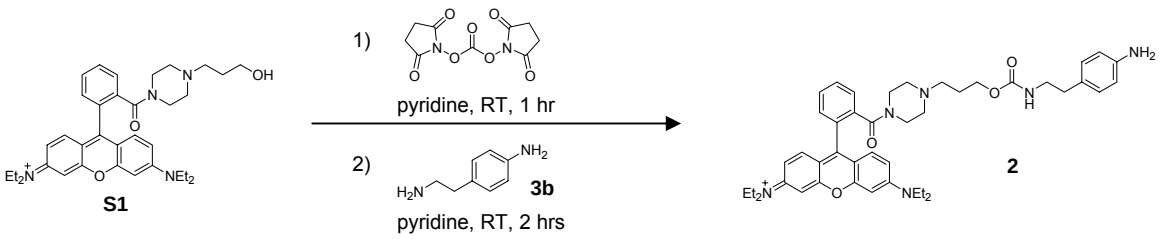
High Perfomance Liquid Chromatography. HPLC was performed on an Agilent 1100 Series HPLC System (Agilent Technologies, USA). Sample analysis for all HPLC experiments was achieved with an inline diode array detector (DAD) and an inline fluorescence detector (FLD).

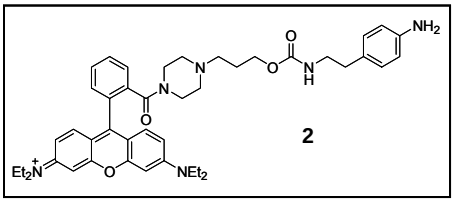
Gel Analyses. For protein analysis, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was accomplished on a Mini-Protean apparatus (Bio-Rad, USA), following the general protocol of Laemmli.[1] Commercially available markers (Bio-Rad, USA) were applied to at least one lane of each gel for calculation of apparent molecular weights. Fluorescence visualization of gels was obtained by UV transillumination at 302 nm. Visualization of protein bands was accomplished by staining with Coomassie® Brilliant Blue R-250 (Bio-Rad, USA). Gel imaging was performed on an EpiChem3 Darkroom system (UVP, USA). Protein recovery was estimated from standard optical density measurements of the observed gel bands with LabWorks™ software (version 4.0.0.8, UVP).

Crude reaction mixtures were diluted by a factor of 5 (v/v) in ddH_2O and combined 1:1 (v/v) with gel loading buffer containing SDS, DTT, and bromophenol blue. Samples were then loaded onto the gel without heating. After removal of the completed gels from their cassettes, the bottom portions containing free dye were excised promptly, after which the gels were submerged in water for rinsing and imaging.

Experimental

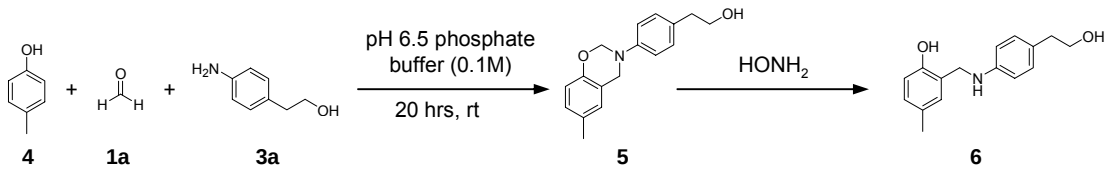
Scheme S1. *Synthesis of Rhodamine B aniline derivative*[2]**.**

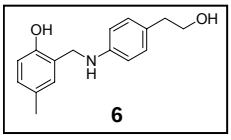




Rhodamine B 4-(3-[2-(4-aminophenyl)-ethyl]-carbamoyloxypropyl)piperazine amide (2). To a solution of alcohol **S1** (50 mg, 0.083 mmol, 1.0 equiv.) in 2.2 mL of pyridine under a N_2 atmosphere was added disuccinimidylcarbonate (DSC) (63 mg, 0.25 mmol, 3.0 equiv.). After 1 hour of stirring at room temperature, the excess DSC was quenched with 15 μL of water for 2 minutes. 2-(4-aminophenyl)-ethanol (16 mg, 0.12 mmol, 1.4 equiv.) was then added to the reaction mixture. After 2 hours of stirring at room temp., the reaction mixture was diluted with an aqueous solution of 1 M K_2CO_3 and extracted 3 times with 2:1 isopropanol/CH_2Cl_2. The combined organic layers were dried over Na_2SO_4, filtered, and the solvent was removed under reduced pressure. The solid was then dissolved in CH_2Cl_2 and filtered to remove salts that had dissolved in the isopropanol/CH_2Cl_2 mixture. The resulting glassy purple solid was used without further purification. ^{1}H NMR (400 MHz, CD3OD): δ, 1.29-1.33 (t, 12H, *J* = 7.2 Hz), 1.70-1.72 (m, 1H), 2.12-2.25 (br s, 4H), 2.30-2.35 (t, 2H, *J*=7.2), 2.40-2.50 (m, 2H), 3.65-3.75 (q, 8H, *J*=7.2), 4.00-4.50 (t, 2H, J=?), 6.97-7.00 (d, 2H, J=2.8), 7.06-7.10 (dd, 2H, J=2.4, 9.6), 7.26-7.30 (d, 2H, *J*=9.6), 7.51-7.53 (m, 1H), 7.60-7.64 (m, 1H), 7.75-7.80 (m, 2H). ^{13}C NMR (400 MHz, CD_3OD): δ, 12.9, 26.1, 30.9, 34.8, 36.4, 36.6, 42.9, 43.5, 47.0, 97.5, 115.0, 115.5, 116.8, 129.0, 130.1, 130.3, 131.5, 131.8, 132.2, 133.3, 136.8, 146.9, 157.0, 157.3, 159.0, 160.4, 169.5, 180.8. IR: 2917, 2849, 1628, 1589, 1463, 1181, cm^{-1}. HRMS (FAB+) calculated for $C_{44}H_{55}N_6O_4$ ($[M]^+$) 731.4279, found 731.4288.

Scheme S2. *Synthesis of p-cresol Mannich-type adduct.*





2-{[4-(2-Hydroxy-ethyl)-phenylamino]-methyl}-4-methyl-phenol (6). To a solution of formaldehyde (0.035g, 1.2 mmol, 5 equiv., added as a 37% (w/w) solution in H_2O) and 2-(4-amino)phenethyl alcohol (0.159 g, 1.16 mmol, 5 equiv.) in phosphate buffer, pH 6.5 (0.1 M), was added *p*-cresol (0.025 g, 0.23 mmol, 1 equiv.). The resulting mixture was stirred at room temperature for 20 hours. After this time, the aqueous phase was separated from the viscous brown precipitate that had formed. The entire mixture was then extracted three times with CH_2Cl_2. The combined organic layers were dried over Na_2SO_4 and concentrated under reduced pressure. This intermediate was determined to be cyclic acetal **5** by ^{1}H NMR and EI-MS characterization. The concentrated mixture was stirred with hydroxylamine hydrochloride (50 mM) in ethanol and then concentrated under reduced pressure. The resulting yellow oil was purified by flash chromatography (hexanes:EtOAc 1:1) yielding compound **6** (0.021g, 18%) as a yellowish solid. ^{1}H NMR (400 MHz, $CDCl_3$): δ, 2.28 (s, 3H), 2.78-2.81 (t, 2H, *J*=6.4), 3.80-3.83 (t, 2H, *J*=6.4), 4.36 (s, 2H), 6.77-6.81 (m, 3H), 6.96 (s, 1H), 7.01-7.03 (d, 2H, *J*=8.0), 7.12-7.10 (d, 2H, *J*=8.4). ^{13}C NMR (400 MHz, $CDCl_3$): δ, 20.5, 38.3, 48.9, 63.8, 116.1, 116.4, 122.5, 129.2, 129.6, 129.9, 130.6, 145.8, 154.4. IR: br 3500-3050, 2924, 2859, 1614, 1517, 1247 cm^{-1} (no carbonyl stretch was observed). HRMS (EI+) calculated for $C_{16}H_{19}NO_2$ ($[M+H]^+$) 258.1494, found 258.1501.

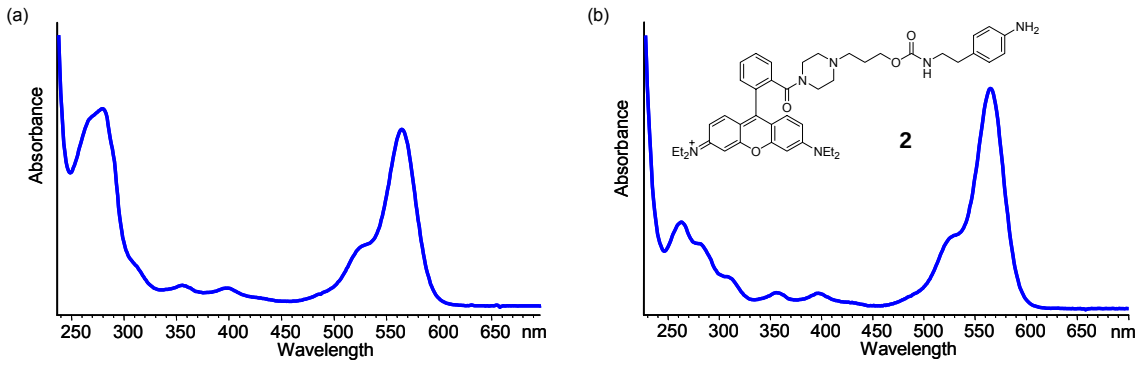


Figure S3. a) UV-vis spectrum of chymotrypsinogen A modified with fluorescent aniline **2**.
b) UV-vis spectrum of fluorescent aniline **2**.

General Procedure for Mannich-Type Modification of Proteins. To a 1.7 mL microcentrifuge tube was added 10 µL of an aniline solution (in 0.1 M phosphate buffer, pH 6.5, 0.5 µmol), 10 µL of an aldehyde solution (in 0.1 M phosphate buffer, pH 6.5, 0.5 µmol), and 2 µL of protein solution (in 0.1 M phosphate buffer, pH 6.5, 0.4 nmol). The final concentration of each reactant was: 25 mM aldehyde, 25 mM aniline, and 20 µM protein. The mixture was vortexed briefly to mix the reaction components, then allowed to stand at room temperature for 18-24 h (some reactions were carried out at 37 °C). The reaction was diluted with 80 µL of water and dialyzed against 0.05 M NH_4HCO_3 buffer, pH 7.8, before further analysis. The UV-Vis spectrum obtained from the inline diode array detector revealed strong

Determination of reaction conversion from UV-Vis data

(2) ε (576 nm) = 96100 $M^{-1}cm^{-1}$, (X) ε (280 nm) = 24911 $M^{-1}cm^{-1}$ (Nguyen, *et. al*)

Chymotrypsinogen A ε (280 nm) = 51240 $M^{-1}cm^{-1}$ (determined from EXPASY Swiss-Prot database)

For labelled protein spectrum:

conc. of **(2)** (according to 576 nm absorbance) = a/(εb) = (30 mAU)/((96100 $M^{-1}cm^{-1}$)(1 cm)) = 0.31mM

280 nm absorbance from **(2)** = εbc = (24911 $M^{-1}cm^{-1}$)(1 cm)(0.00031 M) = 7.78 mAU

280 nm absorbance from chymotrypsiogen A = 32 mAU - 7.78 mAU = 24.2 mAU

conc. of protein = A/(εb) = (24.2 mAU)/((51240 $M^{-1}cm^{-1}$)(1 cm)) = 0.47 mM

percent modification = (conc. of **(2)**)/(conc. of protein) * 100 = (0.31 mM)/(0.47 mM) x 100 = 66%

absorbances at 280 nm and 567 nm, characteristic of dye-labeled protein (**Figure S3**).

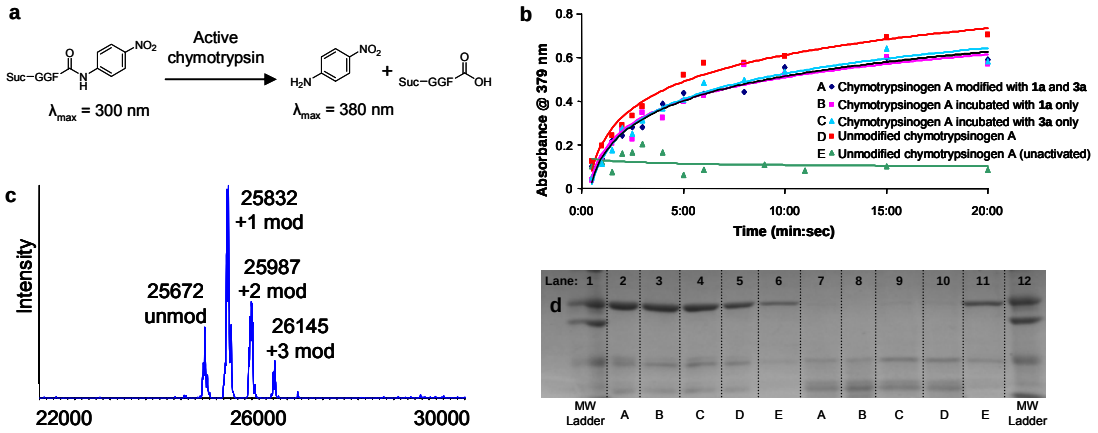
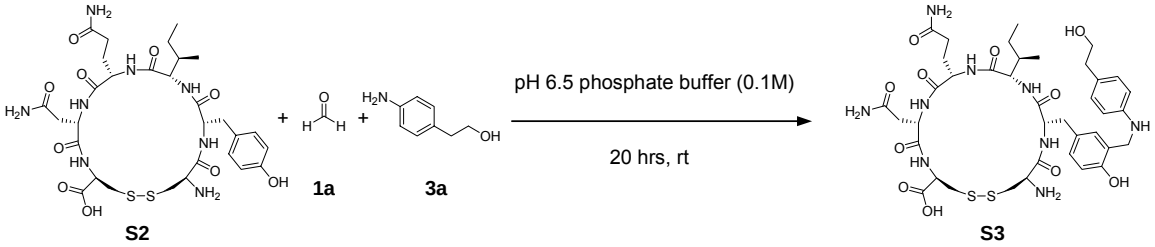


Figure S4. Modified chymotrypsinogen A activity assay. **a**) Activated chymotrypsin catalyzes the hydrolysis of a tripeptide with a *C*-terminal *p*-nitroaniline. **b**) UV-Vis monitoring of the appearance of λmax at 379 nm over 20 minutes for several different samples. **c**) ESI-MS showing extent of modification (sample A) before gel filtration purification. **d**) SDS-PAGE analysis of reaction mixtures before (lanes 2-6) and after (lanes 7-11) addition of trypsin shows that the total protein concentration is the similar in all samples.

Activity Assay of Modified Chymotrypsinogen. An 85 uL aliquot of a reaction mixture for the modification of chymotrypsinogen (5.3 mg/mL, 200µM total protein content) was treated with 11.3 µL of

a solution of sequencing grade modified trypsin (Promega, 20 μg reconstituted with 200 μL of 50 mM acetic acid). The activation of the zymogen was allowed to proceed for 10 minutes at room temperature before being purified by gel filtration (NAP-5) chromatography. The desalted samples were then diluted with 400 μL of 0.1 M Tris buffer (pH 7.6). 100 μL aliquots of the activated protease were then added to 400 μL of 0.5 mM chymotrypsin substrate I, colorimetric (Suc-GGF-pNA, Calbiochem 230912) in 50 mM $CaCl_2$, 20 mM Tris buffer, pH 7.6 (dark blue diamonds). An analogous procedure was followed for chymotripsinogen A samples containing either **1a** (pink squares) or **3a** (light blue triangles) individually and a sample without either component. In a negative control experiment (green triangles), the enzyme was not activated with trypsin before addition of the tripeptide substrate. The progress of the reaction was monitored by UV-Vis spectrophotometry (**Figure S4**).

Scheme S3. *Mannich-type modification of tocinoic acid*



Modification of Tocinoic Acid. To a 1.7 mL microcentrifuge tube was added 50 μL of a 6.8 mg/mL solution of 2-(4-aminophenyl)ethanol (in 0.1 M phosphate buffer, pH 6.5, 2.5 μmol), 50 μL of a 1.5 mg/mL solution of formaldehyde (0.1 M phosphate buffer, pH 6.5, 2.5 μmol), and 10 μL of tocinoic acid (1.48 mg/mL, 20 nmol). The mixture was vortexed briefly to mix the reaction components, then allowed to stand at 37 °C for 18 h. The crude reaction mixture was analyzed directly by ESI-LC/MS. Estimated conversion to singly modified product after this time was approximately 30% as determined by integration of reconstituted mass peaks.

General Procedure for Trypsin Digestion of Modified Proteins. 100 μL of the dialyzed protein reaction mixture was added to 36 mg of solid urea. The resulting solution was briefly vortexed and then heated at 65 ^{o}C in a water bath for 30 min. The denatured protein sample was diluted with 500 μL of 50 mM NH_4HCO_3 buffer, pH 7.8, and then treated with 20 μL of a solution of sequencing grade modified trypsin (Promega, 20 μg reconstituted with 200 μL of 50 mM acetic acid). The digest mixture was then incubated at 37 ^{o}C for 12 h.

The crude digest mixture was desalted using a μC18 ZipTip® pipet tip with a CH_3CN:H_2O, 0.1% trifluoroacetic acid solvent system. The desalted protein mixture was then analyzed by MALDI-TOF MS (**Figure S5**).

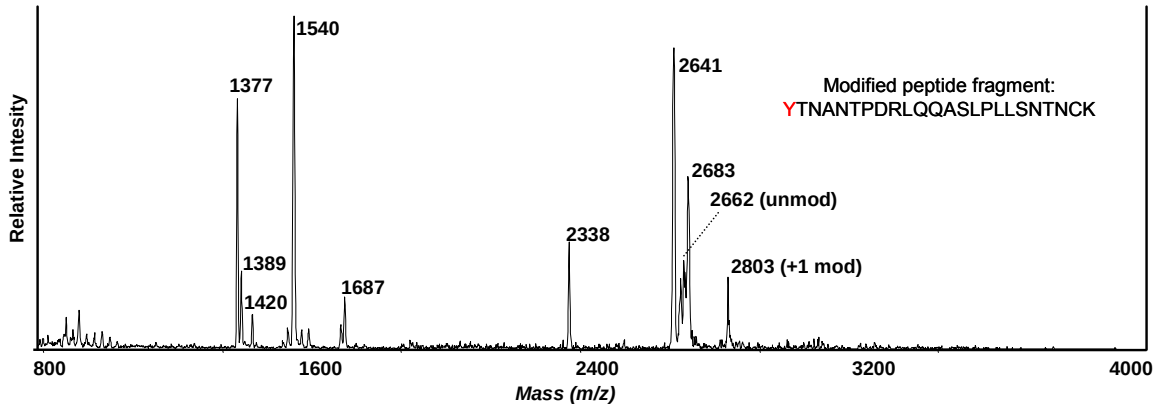


Figure S5. MALDI-TOF MS analysis of modified chymotrypsinogen A trypsin digests. Native peptide fragment containing Y146, expected $[M+H]^+$ = 2664; singly-modified, expected $[M+H]^+$ = 2803.

References

[1] Laemmli, U. K. *Nature* **1970**, *227*, 680.

[2] Synthesis of rhodamine dye starting material has been previously reported. Nguyen, T.; Francis, M. B.; *Org. Lett.* **2003**, *5*, 3245-3248.