Fully Synthetic Granulocyte Colony-Stimulating Factor (G-CSF) Enabled by Isonitrile-Mediated Coupling of Large, Side-Chain Unprotected Peptides

Andrew G. Roberts,[*,‡,Ψ,ǂ] Eric V. Johnston,[‡,Ψ,ǂ,†] Jae-Hung Shieh,[‖] Joseph P. Sondey,[ǂ] Ronald C. Hendrickson,[ǂ] Malcolm A. S. Moore,[‖] and Samuel J. Danishefsky[*,Ψ,ǂ,ø]

[Ψ]Laboratory for Bio-Organic Chemistry, [ǂ]Chemical Biology Program and [‖]Cell Biology Program, Sloan Kettering Institute (SKI) for Cancer Research, New York, New York 10065, United States of America
[ø]Department of Chemistry, Columbia University, New York, New York 10027, United States of America

Supporting Information.

Table of Contents

Part I. Materials and General Procedures.

Materials and General Procedures. All commercially available materials were used without further purification (Aldrich®, Fluka®, Novabiochem®). 2,2'-Azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride (VA-044) was purchased from Wako Pure Chemical Industries. HATU was purchased from Genscript® (Piscataway, New Jersey). Bond-Breaker® solution was purchased from ThermoScientific®. Chitobiose octaacetate was purchased from Carbosynth Limited or prepared from chitin as described. All solvents were reagent grade or HPLC grade (Fisher®). All reactions were performed under an atmosphere of purified dry Argon (Ar) unless indicated otherwise. Low-resolution mass spectral analyses were performed with a JOEL JMS-DX-303-HF mass spectrometer or Waters Micromass ZQ mass spectrometer. Analytical TLC was performed on E. Merck silica gel 60 F254 plates and flash column chromatography was performed on E. Merck silica gel 60 (40–63 mm). Yields refer to chromatographically pure compounds.

Abbreviations: Boc- = *tert*-butyloxycarbonyl, DBU = 1,8-diazobicycloundec-7-ene, DMA = *N,N*-dimethylacetamide, DMF = *N,N*-dimethylformamide, EDC = 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide, Fmoc- = fluorenylmethyloxycarbonyl, HATU = (1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate), HBTU = (2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate), HOOBt = 3-Hydroxy-1,2,3-benzotriazin-4(3H)-one, NMP = *N*-methyl-2-pyrrolidinone, Oxyma Pure® = ethyl (hydroxyimino)cyanoacetate, TCEP = (tris(2-carboxyethyl)phosphine), TFA = trifluoroacetic acid, TFE = 2,2,2-trifluoroethanol.

HPLC-MS, UPLC-MS, HPLC Methods for Analysis and Purification:

HPLC-MS Analysis. All separations involved a mobile phase of 0.05 v/v% TFA in water (Solvent A) and 0.04 v/v% TFA in acetonitrile (Solvent B). Analytical LC-MS analyses were performed using a Waters 2695 Separations Module and a Waters 996 Photodiode Array Detector equipped with Varian Microsorb 100-5, C18 150 × 2.0mm, Varian Microsorb 300-5, C8 250 × 2.0 mm columns and Varian Microsorb 300-5, C4 250 × 2.0 mm columns at a flow rate of 0.2 mL/min. Additional columns used for analysis or semi-preparative purifications were X-Bridge 5 μM, C8 150 × 2.1 mm (0.2 mL/min) and Higgins Analytical PROTO 300 5 μM, C4 250 × 4.6 mm (0.5 mL/min or 0.8 mL/min).

UPLC-MS Analysis. UPLC-MS analyses were performed using a Waters Acquity™ Ultra Performance LC system equipped with Acquity UPLC® BEH C18, 1.7 μl, 2.1 × 100 mm, Acquity UPLC® BEH C8, 1.7μl, 2.1 × 100 mm, Acquity UPLC® BEH 300 C4, 1.7 μl, 2.1 × 100 mm columns at a flow rate of 0.3 mL/min.

Preparative HPLC. All separations involved a mobile phase of 0.05 v/v% TFA in water (Solvent A) and 0.04 v/v% TFA in acetonitrile (Solvent B). Preparative separations were performed using a Rainin HPLC solvent delivery system equipped with a Rainin UV-1 detector and Agilent Dynamax reverse phase HPLC column Microsorb 100-8 C18 (250 × 21.4 mm), or Microsorb 300-5 C8 (250 × 21.4 mm), or Microsorb 300-5 C4 (250 × 21.4 mm) at a flow rate of 16.0 mL/min.

High-Resolution Mass Spectrometry (HR-MS). ***Intact protein LC-MS analysis:*** Intact protein LC-MS analysis was performed on a Waters Nano Acquity LC system coupled to an OrbiElite mass spectrometer (ThermoScientific) via a proxeon-2 nano ESI source. An aliquot of synthetic GCSF **22** or **2a** was injected onto nano ACQUITY UPLC system (Waters) in 20% CH_3CN / water and captured on a C4 IngegraFrit Sample Trap Column 2.5 cm × 150 μm ID (New Objective, Woburn MA)

washed for 10 min at 5 µL / min 20% CH_3CN / water before injection onto a 75-µm inner diameter × 10-cm length C4 analytical column (1.7 µm BEH300; Waters) and elution with a binary linear gradient of 20-80% CH_3CN in water with 0.1 % formic acid over 35 min (400 nL / min). Key parameters for the mass spectrometer were: automatic gain control (AGC) 1×10^6 ions, FT resolution= 240000, scan range 300-2000, capillary temp 275 °C, S-lens RF=64%. For intact protein mode, the gas pressure in the Orbitrap was adjusted by measuring the FT pressure (FT penning gauge) with the HCD gas off and then adjusting the gas needle valve so that the FT pressure was approximately 0.05 - 0.1 $\times 10^{-10}$ Torr above that level with the HCD gas on.

Solid-Phase Peptide Synthesis (SPPS).

Solid-Phase Peptide Synthesis (SPPS) by Fmoc-Strategy.[1,2] Automated SPPS was performed on an Applied Biosystems Pioneer continuous flow peptide synthesizer or a CEM Liberty Microwave Peptide Synthesizer. Peptides were synthesized using automated Fmoc- protocols. The deblock mixtures consisted of a mixture of 100:2:2 of DMF / DBU / piperidine or 4:1 DMF / piperidine with 0.1 M Oxyma Pure®.[3] The appropriate Boc- or Fmoc- amino acids and Fmoc-protected pseudoproline dipeptides[4] from Novabiochem® were employed as indicated. C-terminal thioesters were prepared with using a modification of elongation methods.[5]

Standard Amino Acid Residues Used.

Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Asn(Trt)- OH, Fmoc-Asp(OtBu)-OH, Fmoc-Cys(S*t*-Bu)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Asp(OAll)-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Met-OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Thz-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Val-OH.

Representative Microwave SPPS Protocol.[2]

Initial deprotection (add deblock solution, Power 30 W, 75 °C, 30 secs) / Wash with DMF

Deprotection - add deblock solution (Power 30 W, 75 °C, 180 secs) / Wash with DMF

Coupling - add amino acid (5 equiv, 0.2 M in DMF), add activator (5 equiv, 0.4 M Pyclock in DMF), add base (10 equiv, 2.0 M *i*-Pr_2NEt in NMP), coupling (Power 0 W, RT, 120 secs; Power 15 W, 50 °C, 240 secs) / Wash with DMF

Peptide Cleavage and Deprotection.

Mild Cleavage from Resin to Access Fully Protected Peptides. Upon completion of automated synthesis on a 0.1 mmol scale, the peptide resin was washed into a peptide synthesis vessel with MeOH. After drying, the resin was subjected to a cleavage cocktail (1:1:8 of acetic acid / TFE / methylene chloride) for 30 min (3 × 8 mL) filtering after each treatment. The resulting cleavage solutions were pooled and concentrated. The oily residue was resuspended in minimum amount of acetic acid, precipitated with water, shell frozen and lyophilized.

Acido Labile Protecting Group Removal, Cocktail B. Peptides were subjected to Cocktail B (3 mL / 100 mg of peptide) consisting of TFA (88% by volume), water (5% by volume), phenol (5% by weight), and *i*-Pr_3SiH (2% by volume). The resulting solution was concentrated to half of the original volume with an Ar stream, followed by induced precipitation with ice-cold diethyl ether (3 × 40 mL) and centrifugation to give a white precipitate upon decantation. The precipitate was solubilized in water/acetonitrile (1:1, 0.05% TFA) and lyophilized. The resulting solid was purified by HPLC.

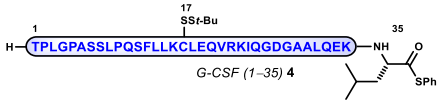
Native Chemical Ligation.

Kinetic Chemical Ligation (KCL) Buffer. The buffer required for kinetic chemical ligation (KCL) was freshly prepared prior to the reaction. Na_2HPO_4 (56.6 mg, 0.4 mmol) was solubilized in water (1 mL), Guanidine·HCl (1.146 g, 12 mmol), and TCEP·HCl (10.8 mg, 0.04 mmol) were then added, solubilzed, the volume adjusted to 2 mL and the pH was brought to the desired pH with a solution of NaOH (5 M). After 15 min of degassing with Ar while sonicating the solution was ready for use.

Native Chemical Ligation Buffer. The buffer required for native chemical ligation (NCL) was freshly prepared prior to the reaction. Na_2HPO_4 (56.6 mg, 0.4 mmol) was solubilized in water (1 mL), Guanidine·HCl (1.146 g, 12 mmol), and TCEP·HCl (10.8 mg, 0.04 mmol) were then added, solubilzed, the volume adjusted to 2 mL and the pH was brought to 7 with a solution of NaOH (5 M, 20 µL). After 15 min degassing with argon, 4-mercaptophenylacetic acid (MPAA) (67 mg, 0.4 mmol) was added and the pH was brought to the desired pH with a solution of NaOH (5 M). After 15 min of degassing with Ar while sonicating the solution was ready for use.

Part II. First Synthetic Approach to G-CSF 1–174 Aglycone (2a).

Preparation of G-CSF 1–35 (4).



Sequence (4):

H-**T**PLGP**AS**SLP**QS**FLLK**C**(S*t*-Bu)LEQVRKIQGDGAALQE**K**L-(SPh)

Chemical Formula: $C_{176}H_{291}N_{45}O_{49}S_3$ **Molecular Weight (g / mol):** 3,917.71

Following the general procedure for microwave-assisted Fmoc-based SPPS with deblock 4:1 DMF/piperidine, peptide (G-CSF 1-34) was prepared on 0.1 mmol scale from Fmoc-Lys(Boc)-NovaSyn® TGT resin (residue 34), Fmoc-Cys(S*t*-Bu)-OH (residue 17), Fmoc-Gln(Trt)-Ser($\Psi^{Me,Me}$Pro)-OH (residues 11-12), Fmoc-Ala-Ser($\Psi^{Me,Me}$Pro)-OH (residues 6-7) and Boc-Thr(*t*-Bu)-OH (residue 1). The peptide resin was subjected to mild cleavage protocol.

After lyophilization, the resultant fully protected peptide (0.1 mmol, G-CSF 1-34) was dissolved in chloroform (20 mL) and cooled to 0 °C with an ice-water bath. To the solution was added H-Leu-SPh•HCl **S-1** (104 mg, 4 equiv), HOOBt (65 mg, 4 equiv) and EDC (71 µL, 4 equiv). The resultant yellow solution was stirred for 3 h at 0 °C, allowed to reach RT, and stirred overnight (12 h). The solution was diluted with 5 vol % HOAc in chloroform (20 mL) and washed with H_2O (2 × 5 mL). The organic layer was washed with a saturated aqueous brine solution (5 mL), dried over Na_2SO_4, filtered and concentrated en vacuo to yield fully protected (G-CSF 1-35) as a crude yellow oil. The oil was diluted with HOAc (2 mL) and precipitated with H_2O (30 mL). The resultant precipitous solution was shell-frozen and lyophilized. The resultant powder was treated with cocktail B (20 mL, 2 h) following the general procedure. The obtained crude pellet was dissolved in H_2O/CH_3CN (1:1, 0.05% TFA) and purified by HPLC (Microsorb 300-5 C8 column, 35% to 60% solvent B over 30 min). Product **4** eluted at 11.0 min. Lyophilization of the collected fractions provided peptide **4** (129 mg, 33%) as a white solid.

See Figure 1.1A and Figure 1.1B ESI-MS (+) of peptide **4**. M = $C_{176}H_{291}N_{45}O_{49}S_3$. ESI calculated for $[M+2H^+]^{2+}$ *m/z*: 1959.86, found: 1959.42; $[M+3H^+]^{3+}$ *m/z*: 1306.91, found: 1306.40; $[M+4H^+]^{4+}$ *m/z*: 980.44, found: 979.96.

Preparation of G-CSF 36–73 (5).

Sequence (5):

H-**C**(S*t*-Bu)ATYKL**C**(S*t*-Bu)HPEELVLLGHSLGIPWAP**LS**S**C**(S*t*-Bu)PSQALQL**A**G-(SEt)

Chemical Formula: $C_{194}H_{312}N_{46}O_{50}S_7$ **Molecular Weight (g / mol):** 4,313.32

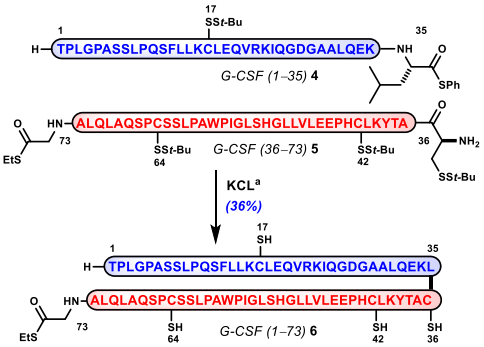
Following the general procedure for microwave-assisted Fmoc-based SPPS with deblock 4:1 DMF/piperidine, peptide (G-CSF 36-72) was prepared on 0.1 mmol scale from Fmoc-Ala-NovaSyn® TGT resin (residue 72), Fmoc-Cys(S*t*-Bu)-OH (residues 42, 64), Fmoc-Leu-Ser($\Psi^{Me,Me}$Pro)-OH (residues 61-62) and Boc-Cys(S*t*-Bu)-OH (residue 36). The peptide resin was subjected to mild cleavage protocol.

After lyophilization, the resultant fully protected peptide (0.1 mmol, G-CSF 36-72) was dissolved in chloroform (20 mL) and cooled to 0 °C with an ice-water bath. To the solution was added H-Gly-SEt•HCl **S-2** (63 mg, 4 equiv), HOOBt (65

mg, 4 equiv) and EDC (71 µL, 4 equiv). The resultant yellow solution was stirred for 3 h at 0 °C, allowed to reach RT, and stirred overnight (12 h). The solution was diluted with 5 vol % HOAc in chloroform (20 mL) and washed with H_2O (2 × 5 mL). The organic layer was washed with a saturated aqueous brine solution (5 mL), dried over Na_2SO_4, filtered and concentrated en vacuo to yield fully protected (G-CSF 36-73) as a crude yellow oil. The oil was diluted with HOAc (2 mL) and precipitated with H_2O (30 mL). The resultant precipitous solution was shell-frozen and lyophilized. The resultant powder was treated with cocktail B (20 mL, 2 h) following the general procedure. The obtained crude pellet was dissolved in H_2O/CH_3CN (1:1, 0.05% TFA) and purified by HPLC (Microsorb 300-5 C8 column, 35% to 50% solvent B over 30 min). Product **5** eluted at 12.4 min. Lyophilization of the collected fractions provided peptide **5** (168 mg, 39%) as a white solid.

See Figure 1.2A and Figure 1.2B ESI-MS (+) of peptide **5**. M = $C_{194}H_{312}N_{46}O_{50}S_7$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1438.78, found: 1439.05; $[M+4H^+]^{4+}$ *m/z*: 1079.34, found: 1079.51; $[M+5H^+]^{5+}$ *m/z*: 863.67, found: 863.82.

Synthesis of G-CSF 1–73 (6).



Sequence (6):

H-TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKL**—**CHPEELVLLGHSLGIPWAPLSSCPSQALQLAG-(SEt)

Chemical Formula: $C_{348}H_{565}N_{91}O_{99}S_5$ **Molecular Weight (g / mol):** 7,768.19

To a solid mixture of pre-weighed peptide **4** (10 mg, 1 equiv) and peptide **5** (14 mg, 1.3 equiv) was added freshly prepared kinetic ligation buffer (1000 µL, pH = 6.8). To this mixture was added neutral TCEP solution (0.5 M, 100 µL). The reaction vial was flushed with Ar and stirred at RT for 12h. After completion, the mixture was diluted with H_2O/CH_3CN (1:1, 0.05% TFA, 2 mL) and neutral TCEP solution (0.5 M, 100 uL). The solution was stirred for 10 min and purified directly by HPLC (Microsorb 300-5 C4 column, 15% to 90% solvent B over 30 min). Product **6** eluted at 13.3 min. Lyophilization of the collected fractions provided peptide **6** (7.5 mg, 38%) as a white solid.*

See Figure 1.3A and Figure 1.3B ESI-MS (+) of peptide **6**. M = $C_{348}H_{565}N_{91}O_{99}S_5$. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1943.06, found: 1943.32; $[M+5H^+]^{5+}$ *m/z*: 1554.65, found: 1555.05; $[M+6H^+]^{6+}$ *m/z*: 1295.71, found: 1295.78; $[M+7H^+]^{7+}$ *m/z*: 1110.75, found: 1110.87; $[M+8H^+]^{8+}$ *m/z*: 972.03, found: 972.32; $[M+9H^+]^{9+}$ *m/z*: 864.14, found: 864.60.

*In a batch and time dependent manner, the presumed macrocyclic product resulting from the intramolecular cyclization of **6** could be observed as a minor product. The presumed macrocycle could be used without consequence in subsequent ligations.

Sequence (macrocycle):

H-TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLSSCPSQALQLAG-(**cyclic**)

Chemical Formula: $C_{346}H_{559}N_{91}O_{99}S_4$ **Molecular Weight (g / mol):** 7,706.06

See Figure 1.4A and Figure 1.4B ESI-MS (+) of peptide **6** (macrocycle). M = $C_{346}H_{559}N_{91}O_{99}S_4$. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1927.42, found: 1927.54; $[M+5H^+]^{5+}$ *m/z*: 1542.22, found: 1542.59; $[M+6H^+]^{6+}$ *m/z*: 1285.35, found: 1285.77; $[M+7H^+]^{7+}$ *m/z*: 1101.87, found: 1102.24.

Preparation of G-CSF 74–126 (7).

Sequence (7):

H-**C**(Thz)LSQLHSGLFLYQGLLQALEGISPELGPTL**DT**LQLDVADF**AT**TIWQQMEEL**G**M-(SEt)

Chemical Formula: $C_{265}H_{414}N_{62}O_{81}S_4$ **Molecular Weight (g / mol):** 5,892.82

Following the general procedure for microwave-assisted Fmoc-based SPPS with deblock 4:1 DMF/piperidine, peptide (G-CSF 74-125) was prepared on 0.2 mmol scale from Fmoc-Gly-NovaSyn® TGT resin (residue 125), Fmoc-Asp(O*t*-Bu)-Thr($\Psi^{Me,Me}$Pro)-OH (residues 104-105), Fmoc-Ala-Thr($\Psi^{Me,Me}$Pro)-OH (residues 114-115) and Boc-Cys(Thz)-OH (residue 74). The peptide resin was subjected to mild cleavage protocol.

After lyophilization, the resultant fully protected peptide (0.2 mmol, G-CSF 74-125) was dissolved in chloroform (20 mL) and cooled to 0 °C with an ice-water bath. To the solution was added H-Met-SEt•HCl **S-3** (138 mg, 3 equiv), HOOBt (98 mg, 3 equiv) and EDC (106 µL, 3 equiv). The resultant yellow solution was stirred for 3 h at 0 °C, allowed to reach RT, and stirred overnight (12 h). The solution was diluted with 5 vol % HOAc in chloroform (40 mL) and washed with H_2O (2 × 5 mL). The organic layer was washed with a saturated aqueous brine solution (5 mL), dried over Na_2SO_4, filtered and concentrated en vacuo to yield fully protected (G-CSF 74-126) as a crude yellow oil. The oil was diluted with HOAc (4 mL) and precipitated with H_2O (60 mL). The resultant precipitous solution was shell-frozen and lyophilized. The resultant powder was treated with cocktail B (40 mL, 2 h, twice) following the general procedure. The obtained crude pellet was dissolved in TFA (3 mL) and H_2O/CH_3CN (3:7, 0.05% TFA, 30-mL) and purified by HPLC (Microsorb 300-5 C4 column, 60% to 95% solvent B over 30 min). Product **7** eluted at 17.0 min. Lyophilization of the collected fractions provided peptide **7** (50 mg, 4%) as a white solid.*

*Precipitation of **7** from crude solutions was common during purification. The product could be re-isolated by centrifugation and re-dissolution prior to HPLC purification. Elution of product **7** was with broadened peak shape and in some cases chromatography was difficult to reproduce. The corresponding CO_2 adduct of Trp118 was a minor side product that complicates purification.

See Figure 1.5A and Figure 1.5B ESI-MS (+) of peptide **7**. M = $C_{265}H_{414}N_{62}O_{81}S_4$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1965.28, found: 1965.52; $[M+4H^+]^{4+}$ *m/z*: 1474.21, found: 1474.41; $[M+5H^+]^{5+}$ *m/z*: 1179.57, found: 1179.66.

Preparation of G-CSF 127–174 (8).

Sequence (8):

H-**C**(S*t*-Bu)PALQPTQGAMPAF**AS**AFQRRAGGVLV**AS**HL**QS**FLE**VS**YRVLRHLAQ**P**-(OH)

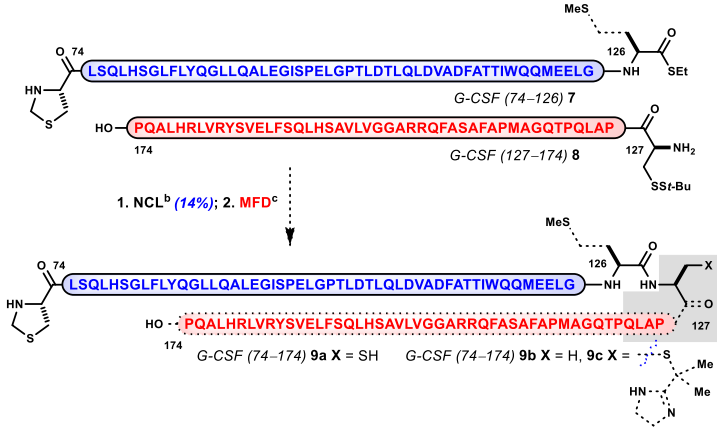
Chemical Formula: $C_{236}H_{375}N_{69}O_{62}S_3$ **Molecular Weight (g / mol):** 5,267.20

Following the general procedure for microwave-assisted Fmoc-based SPPS with deblock 4:1 DMF/piperidine, peptide (G-CSF 127-174) was prepared on 0.2 mmol scale from Fmoc-Pro-NovaSyn® TGT resin (residue 174), Fmoc-Val-Ser($\Psi^{Me,Me}$Pro)-OH (residues 163-164), Fmoc-Gln(Trt)-Ser($\Psi^{Me,Me}$Pro)-OH (residues 158-159), Fmoc-Ala-Ser($\Psi^{Me,Me}$Pro)-OH (residues 154-155, 141-142) and Fmoc-Cys(S*t*-Bu)-OH (residue 127). The peptide resin was subjected to mild cleavage protocol.

After lyophilization, the resultant fully protected peptide (0.2 mmol, G-CSF 127-174) was treated with cocktail B (20 mL, 2h, twice) following the general procedure. The obtained crude pellet was dissolved in H_2O/CH_3CN (1:1, 0.05% TFA) and purified by HPLC (Microsorb 300-5 C8 column, 40% to 65% solvent B over 30 min). Product **8** eluted at 11.5 min. Lyophilization of the collected fractions provided peptide **8** (320 mg, 30%) as a white solid.

See Figure 1.6A and Figure 1.6B ESI-MS (+) of peptide **8**. M = $C_{236}H_{375}N_{69}O_{62}S_3$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1756.74, found: 1756.51; $[M+4H^+]^{4+}$ *m/z*: 1317.81, found: 1317.52; $[M+5H^+]^{5+}$ *m/z*: 1054.45, found: 1054.08.

Synthesis of G-CSF 74–174 (9a) and Experimental Discussion.



Sequence (9a):

H-C(Thz)LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGM**—**CPALQPTQGAMPAFASAFQ RRAGGVLVASHLQSFLEVSYRVLRHLAQP-(OH)

Chemical Formula: $C_{495}H_{775}N_{131}O_{143}S_5$ **Molecular Weight (g / mol):** 11,009.72

Step 1:

To a solid mixture of pre-weighed peptide **7** (3.6 mg, 1 equiv) and peptide **8** (14 mg, 1.1 equiv) was added freshly prepared ligation buffer (400 µL, pH = 7.6). To this mixture was added neutral TCEP solution (0.5 M, 50 µL). The reaction vial was flushed with Ar and stirred at RT for 12h. After completion, the mixture was diluted with H_2O/CH_3CN (3:7, 0.05%

TFA, 2 mL) and neutral TCEP solution (0.5 M, 100 µL). The solution was stirred for 10 min and precipitated with ice cold H_2O (3 mL). The precipitous solution was centrifugated and the resultant pellet was redissolved with TFA (100 µL), diluted with H_2O/CH_3CN (3:7, 0.05% TFA, 2 mL) and neutral TCEP solution (0.5 M, 100 µL) and purified directly by HPLC (Microsorb 300-5 C4 column, 60% to 90% solvent B over 30 min). Product **9a** eluted at 17.0 min. Lyophilization of the collected fractions provided peptide **9a** (0.91 mg, 14%) as a white solid.

For crude reaction analysis, see **Figure 1.7A and Figure 1.7B**.

See Figure 1.8A and Figure 1.8B ESI-MS (+) of peptide **9a**. M = $C_{495}H_{775}N_{131}O_{143}S_5$. ESI calculated for $[M+6H^+]^{6+}$ *m/z*: 1835.96, found: 1836.49; $[M+7H^+]^{7+}$ *m/z*: 1573.83, found: 1573.91; $[M+8H^+]^{8+}$ *m/z*: 1377.22, found: 1377.29; $[M+9H^+]^{9+}$ *m/z*: 1224.31, found: 1224.74; $[M+10H^+]^{10+}$ *m/z*: 1101.98, found: 1102.08; $[M+11H^+]^{11+}$ *m/z*: 1001.89, found: 1002.02.

Attempted Preparation of G-CSF 74–174 (9b).

General Metal-Free Dethiylation Procedure. To a solution of the purified ligation product (1.0 mg) in 0.2 ml of degassed buffer (6 M Gnd·HCl, 200 mM Na_2HPO_4) was added 0.2 ml of 0.5 M bond-breaker® TCEP solution (Pierce), 0.05 ml of 2-methyl-2-propanethiol, and 0.1 ml of radical initiator VA-044 (0.1 M in H_2O). The reaction mixture was stirred at 37 °C and monitored by UPLC-MS following aliquot quench with the addition of acetonitrile / water / TFA.

Step 2:
Attempted Procedure: To a suspension of peptide **9a** (3.0 mg) in Ar sparged DMF (1.5 mL) was added neutral TCEP solution (0.5 M, 500 µL) and 2-methyl-2-propanethiol (120 µL). This solution was further sparged with Ar for 15 min followed by the addition of VA-044 (240 µL of 33 mg/mL H_2O solution). The mixture was sealed under Ar and stirred at RT or elevated temperature and monitored by UPLC-MS following aliquot quench.

Alternative conditions including the substitution of VA-044 with VA-055 or azobisisobutyronitrile (AIBN) and or the substitution of DMF for NMP or DMSO were unsuccessful. Species identified from the attempted desulfurization of **9a** include the desired product **9b** and the radical initiator adduct of **9a** referred to as **9c** as a mostly inseparable mixture.

Products observed by UPLC-MS:

For crude reaction analysis, see **Figure 1.9A, Figure 1.9B, Figure 1.9C and Figure 1.9D.**

Sequence (9b):
H-C(Thz)LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGM**—**
APALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP-(OH)
Chemical Formula: $C_{495}H_{775}N_{131}O_{143}S_4$ **Molecular Weight (g / mol):** 10,977.66

9b, ESI calculated for $[M+8H^+]^{8+}$ *m/z*: 1373.22, found: 1373.62, $[M+7H^+]^{7+}$ *m/z*: 1569.25, found: 1569.68; $[M+6H^+]^{6+}$ *m/z*: 1830.62, found: 1831.00.

Sequence (9c):

H-C(Thz)LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGM—

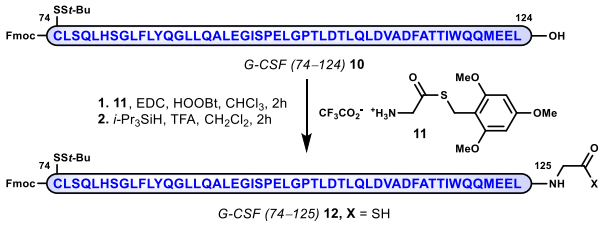
C(adduct)PALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP-(OH)

Chemical Formula: $C_{501}H_{785}N_{133}O_{143}S_5$ **Molecular Weight (g / mol):** 11,119.88

9c, ESI calculated for $[M+8H^+]^{8+}$ *m/z*: 1390.99, found: 1391.18, $[M+7H^+]^{7+}$ *m/z*: 1589.56, found: 1589.86; $[M+6H^+]^{6+}$ *m/z*: 1854.32, found: 1854.63.

Part III. Total Synthesis of G-CSF 1–174 Aglycone (2a).

Preparation of G-CSF 74–125 Thioacid (12).



Sequence (12):

Fmoc-**C**(S*t*-Bu)LSQLHSGLFLYQGLLQALEGISPELGPTL**DT**LQL**D**VA**D**F**AT**TI**W**QQMEELG-(SH)

Chemical Formula: $C_{276}H_{419}N_{61}O_{82}S_4$ **Molecular Weight (g / mol):** 6,031.97

Following the general procedure for microwave-assisted Fmoc-based SPPS with deblock 4:1 DMF/piperidine, peptide (G-CSF 74-124) was prepared on 0.2 mmol scale from Fmoc-Leu-NovaSyn® TGT resin (residue 124), Fmoc-Trp-OH (residue 118), Fmoc-Ala-Thr($\Psi^{Me,Me}$Pro)-OH (residues 114-115), Fmoc-Asp(OMpe)-OH (residues 108, 111), Fmoc-Asp(O*t*-Bu)-Thr($\Psi^{Me,Me}$Pro)-OH (residues 104-105) and Fmoc-Cys(S*t*-Bu)-OH (residue 74). The peptide resin was subjected to mild cleavage protocol.

After lyophilization, the resultant fully protected peptide (assumed 0.14 mmol by mass recovered, G-CSF 74-124) was dissolved in chloroform (14 mL) and cooled to 0 °C with an ice-water bath. To the solution was added **11** (158 mg, 3 equiv), HOOBt (70 mg, 3 equiv) and EDC (75 µL, 3 equiv). The resultant yellow solution was stirred for 3 h at 0 °C, allowed to reach RT, diluted with 5 vol % HOAc in chloroform (40 mL) and washed with H_2O (2 × 5 mL). The organic layer was washed with a saturated aqueous brine solution (5 mL), dried over Na_2SO_4, filtered and concentrated en vacuo to yield fully protected (G-CSF 74-125) as a crude yellow oil. The oil was diluted with HOAc (4 mL) and precipitated with H_2O (60 mL). The resultant precipitous solution was shell-frozen and lyophilized. The resultant powder was treated with a modified cocktail B (24 mL, 2 h, twice) following the general procedure (Modified Cocktail B, water removed and solution is diluted: 200 mg phenol dissolved in methylene chloride (12 mL) and TFA (12 mL), add 600 µL *i*-Pr_3SiH).

Purification of (12) by HPLC. From a similar 0.2 mmol scale batch, the obtained crude pellet was dissolved in TFA (3 mL) and H_2O/CH_3CN (3:7, 0.05% TFA, 30 mL) and purified by HPLC (Microsorb 300-5 C8 column, 70% to 95% solvent B over 30 min). Product **12** eluted at 22.0 min. Lyophilization of the collected fractions provided peptide **12** (30 mg, 2%) as a white solid. Low yield is reflective of broadened peak shape during chromatography as well as observed hydrolysis of **12** to the corresponding carboxylic acid during chromatography or prolonged storage of the crude solution of **12**.

See Figure 1.10A and Figure 1.10B ESI-MS (+) of peptide **12** from LC-MS analysis. M = $C_{276}H_{419}N_{61}O_{82}S_4$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 2011.67, found: 2012.30; $[M+4H^+]^{4+}$ *m/z*: 1509.00, found: 1509.26; $[M+5H^+]^{5+}$ *m/z*: 1207.40, found: 1207.92. * indicates presence of corresponding carboxylic acid in sample or due to hydrolysis during chromatography.

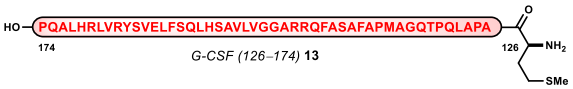
See Figure 1.10C, and Figure 1.10D ESI-MS (+) of peptide **12** from UPLC-MS analysis. M = $C_{276}H_{419}N_{61}O_{82}S_4$. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1509.00, found: 1509.95; $[M+5H^+]^{5+}$ *m/z*: 1207.40, found: 1208.03.

Purification of (12) by C_2 Silica Gel Flash Chromatography. An alternative purification affords **12** with comparable purity suitable for subsequent use in isonitrile-mediated ligations. The hydrophobic nature of **12** makes standard HPLC chromatography (C4 or C8) challenging and difficult to reproduce. The use of Biotage Isolute® C_2 silica gel on a flash chromatography system allows **12** to be purified in fewer injections and in less time on the column. Column dimensions 30 × 100 mm, approximately 15 g C_2 silica gel. From the above preparation (0.2 mmol), a portion of crude **12** (approximately 1/6 of material) was dissolved in 100 µL TFA and 500 µL H_2O/CH_3CN (3:7, 0.05% TFA) and loaded onto an equilibrated C_2 silica gel column (equilibrated successively with CH_3CN → H_2O → H_2O/CH_3CN (1:1, 0.05% TFA)). A stepwise gradient was employed at a flow rate of 12 mL/min, H_2O/CH_3CN (1:1, 0.05% TFA, 6 CV) → H_2O/CH_3CN (1:3, 0.05% TFA, 7 CV) → CH_3CN (0.05% TFA, 7 CV). The product **12** was eluted in several 8 mL fractions during H_2O/CH_3CN (1:3, 0.05% TFA) elution (at approximately 12 min of a 20 min run). The entire batch of crude **12** was purified in 6 injections. Lyophilization of the collected fractions provided peptide **12** (240 mg, 20%) as a white solid. Purified **12** obtained from this method contains a minor amount of trifluoroacetylated (-$COCF_3$) derivatives.

For UV-chromatogram from flash chromatography purification, see **Figure 1.11A**.

See Figure 1.11B and Figure 1.11C Top: ESI-MS (+) of peptide **12** + **-$COCF_3$**. M = $C_{278}H_{418}F_3N_{61}O_{83}S_4$. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1533.00, found: 1532.96; $[M+5H^+]^{5+}$ *m/z*: 1226.60, found: 1226.40. Bottom: ESI-MS (+) of peptide **12**. M = $C_{276}H_{419}N_{61}O_{82}S_4$. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1509.00, found: 1508.86; $[M+5H^+]^{5+}$ *m/z*: 1207.40, found: 1207.23.

Preparation of G-CSF 126–174 (13).



G-CSF (126–174) **13**

Sequence (13):

H-**M**APALQPTQGAMPAF**AS**AFQRRAGGVLV**AS**HL**QS**FLE**VS**YRVLRHLAQ**P**-(OH)

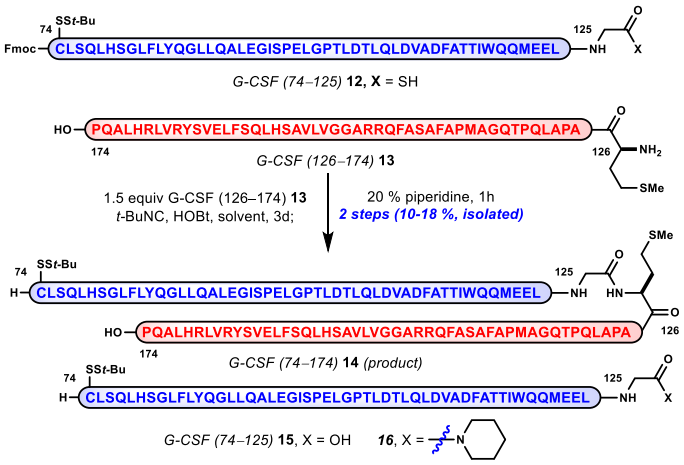
Chemical Formula: $C_{237}H_{376}N_{70}O_{63}S_2$ **Molecular Weight (g / mol):** 5,278.16

Following the general procedure for microwave-assisted Fmoc-based SPPS with deblock 4:1 DMF/piperidine, peptide (G-CSF 126-174) was prepared on 0.2 mmol scale from Fmoc-Pro-NovaSyn® TGT resin (residue 174), Fmoc-Val-Ser($\Psi^{Me,Me}$Pro)-OH (residues 163-164), Fmoc-Gln(Trt)-Ser($\Psi^{Me,Me}$Pro)-OH (residues 158-159), Fmoc-Ala-Ser($\Psi^{Me,Me}$Pro)-OH (residues 154-155, 141-142) and Fmoc-Met-OH (residue 126). The peptide resin was subjected to mild cleavage protocol.

After lyophilization, the resultant fully protected peptide (0.2 mmol, G-CSF 127-174) was treated with cocktail B (20 mL, 2h, twice) following the general procedure. The obtained crude pellet was dissolved in H_2O/CH_3CN (1:1, 0.05% TFA) and purified by HPLC (Microsorb 300-5 C8 column, 40% to 65% solvent B over 30 min). Product **13** eluted at 11.0 min. Lyophilization of the collected fractions provided peptide **13** (314 mg, 30%) as a white solid.

See Figure 1.12A and Figure 1.11B ESI-MS (+) of peptide **13**. M = $C_{237}H_{376}N_{70}O_{63}S_2$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1760.39, found: 1761.75; $[M+4H^+]^{4+}$ *m/z*: 1320.55, found: 1321.55; $[M+5H^+]^{5+}$ *m/z*: 1056.64, found: 1057.38; $[M+6H^+]^{6+}$ *m/z*: 880.70, found: 881.37.

Synthesis of G-CSF 74–174 (14) and Experimental Discussion.



Sequence (14):

H-C(S*t*-Bu)LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELG**—**MAPALQPTQGAMPAFASAFQ RRAGGVLVASHLQSFLEVSYRVLRHLAQP-(OH)

Chemical Formula: $C_{498}H_{783}N_{131}O_{143}S_5$ **Molecular Weight (g / mol):** 11,053.82

General Procedure. To an oven-dried 4-mL vial was added thioacid G-CSF 74–125 **12** (5.0 mg, 0.83 μmol, 1 equiv), N-terminal amine nucleophile G-CSF 126–174 **13** (6.6 mg, 1.24 μmol, 1.5 equiv), and HOBt (1.3 mg, 8.3 μmol, 10 equiv). A stir bar was added, followed by the addition of anhydrous solvent in accordance with the table (1.4 mM = 600 μL, 2.8 mM = 300 μL, 5.6 mM = 150 μL). The heterogeneous mixture was stirred at RT for 5 min followed by the addition of *tert*-butyl isonitrile via syringe from a 86.6 mM stock-solution prepared in DMF (29 μL, 2.49 μmol, 3.0 equiv). The reactions were capped and stirred for 72h. At t = 72 h, reactions were diluted with 300 μL deblock solution (20 vol % piperidine in DMF) and stirred vigorously at RT for 1h, cooled in dry ice, followed by precipitation with ice cold diethyl ether (6 mL). The precipitate was centrifugated (5 min, 5 °C, 5000 rpm) and the supernatant was decanted (repeated twice). The resultant pellet was frozen in liquid nitrogen and lyophilized prior to dissolution and purification by semi-preparative HPLC.

Entry	Solvent	Conc.	*t*-BuNC	HOBt	Ratio (14 : 15 : 16)[a]	% of 14[a]	Yield[b]
1	DMF	2.8 *mM*	3 equiv	10 equiv	**1.9 : 3.1 : 1**	32	**11%**
2	NMP	2.8 *mM*	3 equiv	10 equiv	**1.6 : 2.5 : 1**	31	**-[c]**
3	DMA	2.8 *mM*	3 equiv	10 equiv	**3.1 : 2.6 : 1**	46	**11%**
4	DMF	1.4 *mM*	3 equiv	10 equiv	**1.5 : 2.1 : 1**	33	**16%**
5	DMA	5.6 *mM*	3 equiv	10 equiv	**3.7 : 3.8 : 1**	44	**18%**
6	DMF	2.8 *mM*	*0 equiv*	10 equiv	**1 : 13.1 : 15.2**	3	**ND**
7	DMA	2.8 *mM*	*0 equiv*	10 equiv	**1 : 9.9 : 10.7**	5	**ND**
8	DMA	2.8 *mM*	3 equiv	*0 equiv*	**1 : 4.7 : 2.5**	12	**ND**
9	DMA	2.8 *mM*	3 equiv	10 equiv	**2.4 : 1.4 : 1**	50	**10%**

TABLE S-1. Optimization of Coupling Conditions for the Preparation of G-CSF 74–174 (14). a) Determined by relative integration (area under the curve, AUC) of the corresponding UV-peaks at 280 nm; b) Yields refer to HPLC purified material and account for material loss due to reaction monitoring, ND = not determined; c) sample oxidation and aggregation upon work-up.

Reaction monitoring. At the indicated timepoints (t = 20h, 44h, 72h) a 20 μL aliquot was taken and diluted with 100 μL of deblock solution (20 vol % piperidine in DMF). The Fmoc- removal reaction was allowed to stir at RT for 1h and subsequently cooled in dry ice followed by precipitation with of cold diethyl ether (1 mL). The precipitate was centrifugated (5 min, 5 °C,

5000 rpm) and the supernatant was decanted. The resultant pellet was frozen in liquid nitrogen and lyophilized prior to UPLC-MS analysis. The dried pellet was redissolved in 40 µL TFA, 50 µL acetonitrile, 10 µL water (in that order) followed by immediate UPLC-MS analysis. Relative product ratios were determined by integration of corresponding UV peaks analyzed at 280 nm.[6,7]

Reaction purification. (Entries 1, 3-5, 9) The dried pellet was re-dissolved in 100 µL TFA and 200 µL acetonitrile. This solution was precipitated by the addition of 10 mL water. The resultant cloudy solution was frozen in liquid nitrogen and lyophilized prior to re-dissolution for purification. ***This re-lyophilization process was essential to minimize product loss due to significant aggregation.*** The resultant crude powder was dissolved in 100 µL TFA, 200 µL acetonitrile, and 20 µL water (in that order). The crude solution was purified in 3 × 100 µL injections by semi-preparative HPLC. Isolated yields are reflective of pure products obtained after lyophilization of corresponding fractions. Theoretical yield calculations account for material lost to timepoint analyses (For example: Entry 1, 2.8 mM DMF = 300 µL, removed 60 µL for analyses, theoretical yield of product: 7.30 mg (reduced from 9.16 mg).

Sample Entry (A):

Entry 4. From thioacid G-CSF 74–125 **12** (5.0 mg, 0.83 µmol, 1 equiv), *N*-terminal amine nucleophile G-CSF 126–174 **13** (6.6 mg, 1.24 µmol, 1.5 equiv), and HOBt (1.3 mg, 8.3 µmol, 10 equiv) in DMF (600 µL) at RT for 72h. Crude material (containing excess **13**, product **14**, carboxylic acid **15** and piperidide **16**) was dissolved as described and purified by HPLC (Higgins Analytical PROTO 300 5 µM, C4 250 × 4.6 mm, 40% to 95% solvent B over 30 min, flow = 0.8 mL/min). Product **14** (1.3 mg, 16%) elutes at 15.3 min and the carboxylic acid **15** (0.6 mg, 13%) elutes at 20.8 min followed by an impure mixture (0.8 mg) containing piperidide **16**.

Sample Entry (B):

Entry 5. From thioacid G-CSF 74–125 **12** (5.0 mg, 0.83 µmol, 1 equiv), *N*-terminal amine nucleophile G-CSF 126–174 **13** (6.6 mg, 1.24 µmol, 1.5 equiv), and HOBt (1.3 mg, 8.3 µmol, 10 equiv) in DMA (150 µL) at RT for 72h. Crude material (containing excess **13**, product **14**, carboxylic acid **15** and piperidide **16**) was dissolved as described and purified by HPLC (Higgins Analytical PROTO 300 5 µM, C4 250 × 4.6 mm, 40% to 95% solvent B over 30 min, flow = 0.8 mL/min). Product **14** (1.0 mg, 18%) elutes at 15.3 min and the carboxylic acid **15** (0.6 mg, 20%) elutes at 20.8 min followed by an impure mixture (0.7 mg) containing piperidide **16**.

For crude reaction analysis, see **Figure 1.13A, Figure 1.13B, and Figure 1.13C.**
For graphical representation of timepoint analyses, see **Figure 1.14A, Figure 1.14B and Figure 1.14C**.
For sample purification of **14** from crude reaction mixture, see **Figure 1.15**.

Sequence (14):

H-C(S*t*-Bu)LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELG**—**MAPALQPTQGAMPAFASAFQ RRAGGVLVASHLQSFLEVSYRVLRHLAQP-(OH)

Chemical Formula: $C_{498}H_{783}N_{131}O_{143}S_5$ **Molecular Weight (g / mol):** 11,053.82

See Figure 1.16A, Figure 1.16B and Figure 1.14C ESI-MS (+) of peptide **14**. M = $C_{498}H_{783}N_{131}O_{143}S_5$. ESI calculated for $[M+6H^+]^{6+}$ *m/z*: 1843.31, found: 1842.97; $[M+7H^+]^{7+}$ *m/z*: 1580.12, found: 1579.72; $[M+8H^+]^{8+}$ *m/z*: 1382.74, found: 1382.45; $[M+9H^+]^{9+}$ *m/z*: 1229.21, found: 1228.98; $[M+10H^+]^{10+}$ *m/z*: 1106.39, found: 1106.28; $[M+11H^+]^{11+}$ *m/z*: 1005.90, found: 1005.92; $[M+12H^+]^{12+}$ *m/z*: 922.16, found: 921.85; $[M+13H^+]^{13+}$ *m/z*: 851.30, found: 850.97; $[M+14H^+]^{14+}$ *m/z*: 790.57, found: 790.38; $[M+15H^+]^{15+}$ *m/z*: 737.93, found: 737.80.

Sequence (15):

H-C(S*t*-Bu)LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELG-(OH)

Chemical Formula: $C_{261}H_{409}N_{61}O_{81}S_3$ **Molecular Weight (g / mol):** 5,793.67

See Figure 1.17A and Figure 1.17B ESI-MS (+) of peptide **15**. M = $C_{261}H_{409}N_{61}O_{81}S_3$. ESI calculated for $[M+3H^+]^{5+}$ *m/z*: 1932.23, found: 1923.05; $[M+4H^+]^{5+}$ *m/z*: 1449.43, found: 1449.28; $[M+5H^+]^{5+}$ *m/z*: 1159.74, found: 1159.60.

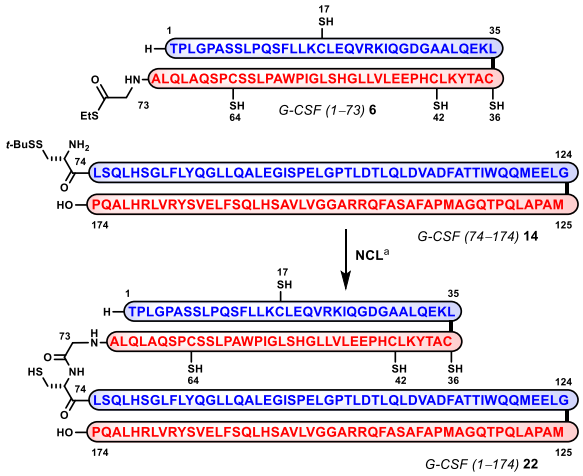
Sequence (16):

H-C(S*t*-Bu)LSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELG-(piperidine)

Chemical Formula: $C_{266}H_{418}N_{62}O_{80}S_3$ **Molecular Weight (g / mol):** 5,860.80

See Figure 1.13A and Figure 1.13B Top: Product identified as piperidide (**16**) at 3.94 min, found $[M+3H^+]^{3+}$, $[M+4H^+]^{5+}$, $[M+5H^+]^{5+}$.

Synthesis of G-CSF 1–174 (22).



Sequence (22):

H-TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLSS
CPSQALQLAG—CLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGM
APALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP-(OH)

Chemical Formula: $C_{840}H_{1334}N_{222}O_{242}S_8$ **Molecular Weight (g / mol):** 18,671.4735

To a pre-weighed mixture of G-CSF 1–73 **6** (2.4 mg, 0.31 μmol, 1.7 equiv) and G-CSF 74–174 **14** (2.0 mg, 0.18 μmol, 1 equiv) in a conical vial with a stir bar was added modified native chemical ligation buffer (200 μL, pH = 6.8). (Modified Buffer: Same as general procedure with 2 × TCEP quantity (21.6 mg TCEP, 0.08 mmol, total volume 2-mL)) The vial was

flushed with Ar, capped and stirred at RT for 18h until judged complete by UPLC-MS analysis indicating consumption of **14**. Sample aliquot preparation: (t = 18h, 10 µL reaction diluted with neutral TCEP solution (0.5 M, 10 µL), inject 10 µL). The reaction was protected from oxidation by dilution with neutral TCEP solution (0.5 M, 20 µL) and purified from the modified native chemical ligation buffer at 22h by direct injection (2 × ~100 µL) onto semi-preparative HPLC (Higgins Analytical PROTO 300 5 µM, C4 250 × 4.6 mm, 40% to 95% solvent B over 30 min, flow = 0.8 mL/min). Product **22** eluted from 16.0 to 21.0 min. All fractions were reanalyzed by UPLC-MS and fractions containing pure **22** (17.0 to 20.5 min) were pooled, shell frozen and lyophilized to yield **22** (550 µg, 16%) as a white powder.

For LC-MS trace and MS analysis of **22**, see **Figure 1.18A and Figure 1.18B**.
For HRMS analysis and tabulation of MS data for **22**, see **Figure 1.26 and Figure 1.28.**

The conversion of purified **6** and **14** to G-CSF 1–174 **22** was conducted several times (Entry 1: isolated **22**: 650 µg lyophilized weight, 19% from 0.181 µmol scale). Precipitation from ligation buffer was a common occurrence. The two step procedure involving direct precipitation of crude G-CSF 1–174 **22** and direct folding to afford folded G-CSF 1–174 **2a** addressed this issue (see below).

Synthesis of G-CSF 1–174 Aglycone (2a).
Sequence (2a) (Cys36Cys42, Cys64Cys74):
H-TPLGPASSLPQSFLLKCLEQVRKIQGDGAALQEKLCATYKLCHPEELVLLGHSLGIPWAPLSS
CPSQALQLAGCLSQLHSGLFLYQGLLQALEGISPELGPTLDTLQLDVADFATTIWQQMEELGM
APALQPTQGAMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP-(OH)
Chemical Formula: $C_{840}H_{1330}N_{222}O_{242}S_8$ **Molecular Weight (g / mol):** 18,667.4418

To a pre-weighed mixture of G-CSF 1–73 **6** (3.0 mg, 0.38 µmol, 1.5 equiv) and G-CSF 74–174 **14** (2.8 mg, 0.25 µmol, 1 equiv) in a conical vial with a stir bar was added modified native chemical ligation buffer (300 µL, pH = 6.8). (Modified Buffer: Same as general procedure with 2 × TCEP quantity (21.6 mg TCEP, 0.08 mmol, total volume 2-mL)) The vial was flushed with Ar, capped and stirred at RT for 20h until judged complete by UPLC-MS analysis indicating consumption of **14**. Sample aliquot preparation: (t = 17h, 10 µL reaction diluted with neutral TCEP solution (0.5 M, 10 µL), inject 10 µL). The reaction was diluted at 20h with neutral TCEP solution (0.5 M, 100 µL) followed by the addition of ice cold degassed H_2O (1000 µL) to induce precipitation of the desired crude product **22**. The vial was cooled in an ice-water bath for 5 min followed by centrifugation and decantation (repeated 3 × 1000 µL with cold H_2O). The final pellet was shell frozen and lyophilized to yield crude G-CSF **22** (assumed 0.25 µmol) as a white solid.

The crude material **22** (assumed 0.25 µmol) was suspended in buffer-1 (900 µL) and stirred for 30 min in a preheated oil bath at 25 °C. At this point, buffer-2 (90 µL) was added and the now homogeneous mixture was stirred loosely capped at 25 °C for 20h.[8] The reaction was cooled to RT and purified directly by semi-preparative HPLC (Higgins Analytical PROTO 300 5 µM, C4 250 × 4.6 mm, 40% to 95% solvent B over 30 min, flow = 0.5 mL/min). The batch was purified in eight injections (8 x ~100 µL crude). Product **2a** eluted from 33.0 to 35.0 min. All fractions were reanalyzed by UPLC-MS and pure fractions were pooled (~8 mL total volume, estimated at 75% solvent B). The combined fractions were diluted with H_2O (12 mL) containing HOAc (40 µL) and concentrated using 10,000 MWCO Amicon centrifugation filters at 3000 rpm, 15 °C (3 × 10 min cycles, with re-dilution each cycle). The combined solution was reduced to total volume of 1500 µL and re-

diluted with buffer-3 (2 mL). This diluted solution was concentrated using 10,000 MWCO Amicon centrifugation filters at 3000 rpm, 15 °C (3 × 10 min cycles, with re-dilution each cycle). The combined solution was reduced to total volume of 3750 µL buffer-3. The concentration of this solution was determined to be 92 µg/mL **2a** (calc. 345 µg, 7%, 2 steps) by UV-peak integration following UPLC-MS analysis relative to an internal standard.[7] Concentration analysis by Bradford protein assay determined sample **2a** to be 80 µg/mL (calc. 300 µg, 6%, 2 steps).[9] This concentration (80 µg/mL) was used in all subsequent experiments.

Buffer 1: Take 1000 µL of TRIS•HCl solution (1 M, pH = 8.0) and sodium lauroyl sarcosinate (400 mg) and dilute to 20 mL with degassed H_2O to make a 50 mM TRIS•HCl buffer containing 2 wt% sarkosyl.
Buffer 2: Take 10 mL of the above 'Buffer-1' and add $CuSO_4$ pentahydrate (1.0 mg) to make a 400 µM $CuSO_4$ solution.
Buffer 3: A 1-L solution was prepared by dissolution of D-sorbitol (50 g) and NaOAc (820 mg) in pyrogen-free H_2O (945 mL) The pH was adjusted by the addition of glacial HOAc (300 µL, to pH = 4.6). A solution of Tween-20 (60 mg dispersed in 5 mL pyrogen-free H_2O) was added and the solution was thoroughly mixed. The solution was allowed to rest for 12h.

Sequence (2a) (Cys36Cys42, Cys64Cys74):
Chemical Formula: $C_{840}H_{1330}N_{222}O_{242}S_8$ **Molecular Weight (g / mol):** 18,667.4418

Alternatively, G-CSF **2a** could be prepared from HPLC purified and lyophilized material G-CSF **22** in a similar manner. Spectroscopic characterization data, biological activity and recovery of G-CSF **2a** in this way (2 steps, 2 purifications) were comparable to the one-step protocol described. The conversion of purified **22** to folded **2a** was conducted several times (Entry 1: isolated **2a**: 180 µg lyophilized weight, 28% from 0.034 µmol scale).

For crude reaction analysis, folding of **22** to provide **2a** see **Figure 1.19A, Figure 1.19B and Figure 1.19C**.
For low resolution MS analysis of **22** and **2a**, see **Figure 1.20A and Figure 1.20B**.
For LC-MS data of standard recombinant G-CSF **2b** see **Figure 1.21A**.
For LC-MS data comparing recombinant G-CSF **2b** with crude G-CSF **2a** and sample purification, see **Figure 1.21B**.
For LC-MS data comparing recombinant G-CSF **2b** with crude G-CSF **2a** and purified G-CSF **2a**, see **Figure 1.21C**.
For UV analysis from purification of G-CSF **2a**, see **Figure 1.22 and Figure 1.23.**
For UPLC-MS analysis of purified G-CSF **2a**, see **Figure 1.24A**.
For UPLC-MS analysis comparing **22** and G-CSF **2a**, see **Figure 1.24B.**
For LC-MS analysis comparing **22** and G-CSF **2a**, see **Figure 1.25.**
For HRMS analysis and tabulation of MS data for G-CSF **2a**, see **Figure 1.27 and Figure 1.28.**

Part IV. Synthesis of G-CSF 126–174 (13) and Glycopeptide G-CSF 126–174 (21).

Preparation of G-CSF 126–149 (17).

Sequence (17):

Fmoc-MAPALQPTQGAMPAF**AS**AFQRRAG-(SH)

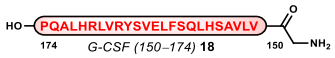
Chemical Formula: $C_{123}H_{181}N_{33}O_{31}S_3$ **Molecular Weight (g / mol):** 2,714.18

Following the general procedure for Fmoc-based SPPS on an Applied Biosystems Pioneer with deblock 4:1 DMF/piperidine containing 0.1 M Oxyma Pure, peptide (G-CSF 126-148) was prepared on 0.1 mmol scale from Fmoc-Ala-NovaSyn® TGT resin (residue 148), Fmoc-Ala-Ser($\Psi^{Me,Me}$Pro)-OH (residues 154-155, 141-142) and Fmoc-Met-OH (residue 126). The peptide resin was subjected to mild cleavage protocol.

After lyophilization, the resultant fully protected peptide (assumed 0.08 mmol by mass recovered, G-CSF 126-148) was dissolved in chloroform (3 mL) and cooled to 0 °C with an ice-water bath. To the solution was added **11** (90 mg, 3 equiv), HOOBt (39.4 mg, 3 equiv) and EDC (44 µL, 3 equiv). The resultant yellow solution was stirred for 3 h at 0 °C, allowed to reach RT, stirred at RT for 12h, diluted with 5 vol % HOAc in chloroform (20 mL) and washed with H_2O (2 × 5 mL). The organic layer was washed with a saturated aqueous brine solution (5 mL), dried over Na_2SO_4, filtered and concentrated en vacuo to yield fully protected (G-CSF 126-149) as a crude yellow oil. The oil was diluted with HOAc (2 mL) and precipitated with H_2O (60 mL). The resultant precipitous solution was shell-frozen and lyophilized. The resultant powder was treated with a modified cocktail B (6 mL, 2 h, twice) following the general procedure (Modified Cocktail B, water removed and solution is diluted: 200 mg phenol dissolved in methylene chloride (12 mL) and TFA (12 mL), add 600 µL *i*-Pr_3SiH). The obtained crude pellet was dissolved in H_2O/CH_3CN (1:1, 0.05% TFA) and purified by HPLC (Microsorb 300-5 C4 column, 30% to 70% solvent B over 30 min). Product **17** eluted at 16.0 min. Lyophilization of the collected fractions provided peptide **17** (24 mg, 9%) as a white solid.

See Figure 1.29A and Figure 1.29B ESI-MS (+) of peptide **17**. M = $C_{123}H_{181}N_{33}O_{31}S_3$. ESI calculated for $[M\times2+3H^+]^{3+}$ *m/z*: 1810.5, found: 1810.4; $[M+2H^+]^{2+}$ *m/z*: 1358.1, found: 1358.0; $[M+3H^+]^{3+}$ *m/z*: 905.7, found: 905.7.

Preparation of G-CSF 150–174 (18).



Sequence (18):

H-GVLV**AS**HL**QS**FLE**VS**YRVLRHLAQ**P**-(OH)

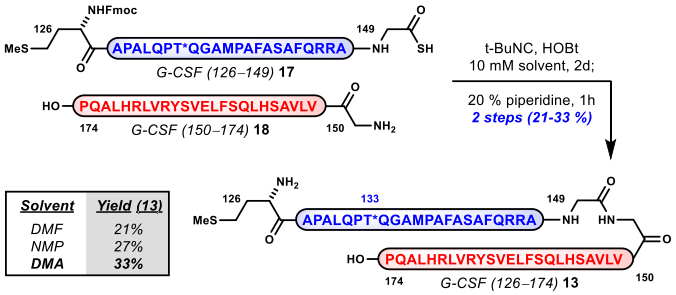
Chemical Formula: $C_{129}H_{207}N_{37}O_{34}$ **Molecular Weight (g / mol):** 2,820.30

Following the general procedure for microwave-assisted Fmoc-based SPPS with deblock 4:1 DMF/piperidine, peptide (G-CSF 150-174) was prepared on 0.1 mmol scale from Fmoc-Pro-NovaSyn® TGT resin (residue 174), Fmoc-Val-Ser($\Psi^{Me,Me}$Pro)-OH (residues 163-164), Fmoc-Gln(Trt)-Ser($\Psi^{Me,Me}$Pro)-OH (residues 158-159), and Fmoc-Ala-Ser($\Psi^{Me,Me}$Pro)-OH (residues 154-155). The peptide resin was subjected to mild cleavage protocol.

After lyophilization, the resultant fully protected peptide (0.1 mmol, G-CSF 136-174) was treated with cocktail B (20 mL, 2h, twice) following the general procedure. A portion (0.05 mmol) of obtained crude pellet was dissolved in H_2O/CH_3CN (1:1, 0.05% TFA) and purified by HPLC (Microsorb 300-5 C8 column, 40% to 65% solvent B over 30 min). Product **18** eluted at 12.0 min. Lyophilization of the collected fractions provided peptide **18** (36 mg, 26%) as a white solid.

See Figure 1.30A and Figure 1.30B ESI-MS (+) of peptide **18**. M = $C_{129}H_{207}N_{37}O_{34}$. ESI calculated for $[M\times2+3H^+]^{3+}$ *m/z*: 1881.20, found: 1881.82; $[M+2H^+]^{2+}$ *m/z*: 1411.16, found: 1411.41; $[M+3H^+]^{3+}$ *m/z*: 941.11, found: 941.21; $[M+4H^+]^{4+}$ *m/z*: 706.08, found: 706.08.

Synthesis of G-CSF 126–174 (13).



Sequence (13):

H-MAPALQPTQGAMPAFASAFQRRAG—GVLVASHLQSFLEVSYRVLRHLAQP-(OH)

Chemical Formula: $C_{237}H_{376}N_{70}O_{63}S_2$ **Molecular Weight (g / mol):** 5,278.16

To an oven-dried 4-mL vial was added thioacid G-CSF 126–149 **17** (5.0 mg, 1.8 µmol, 1.0 equiv), N-terminal amine **18** (7.8 mg, 2.8 µmol, 1.5 equiv), and HOBt (2.8 mg, 18.4 µmol, 10 equiv). A stir bar was added, followed by the addition of anhydrous solvent (DMF, NMP or DMA) to a final concentration of 10 mM **17** (180 µL). The mixture was stirred at RT for 5 min followed by the addition of *tert*-butyl isonitrile (5.5 µmol, 3.0 equiv) via syringe from a 100 mM stock-solution prepared in the corresponding solvent. The reactions were capped and stirred for 72h. At t = 72 h, reactions were diluted with 100 µL deblock solution (20 vol % piperidine in DMF) and stirred vigorously at RT for 1h, cooled in dry ice, followed by precipitation with 6 mL of cold diethyl ether. The precipitate was centrifugated (5 min, 5 °C, 5000 rpm) and the supernatant was decanted. The resultant pellet was frozen in liquid nitrogen and lyophilized prior to purification by semi-preparative HPLC. The crude material (containing excess **18** and product **13**) was dissolved in H_2O/CH_3CN (1:1, 0.05% TFA) and purified by HPLC (Microsorb 300-5 C8 column, 40% to 65% solvent B over 30 min). Product **13** elutes at 11.0 min and followed by recovered **18**.

Entry 1 (DMF) – Isolated product **13** (2.1 mg, 21%) and recovered **18** (3.6 mg).
Entry 2 (NMP) – Isolated product **13** (2.6 mg, 27%) and recovered **18** (1.6 mg).
Entry 3 (DMA) – Isolated product **13** (3.2 mg, 33%) and recovered **18** (1.3 mg).

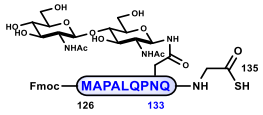
Sequence (13):

H-MAPALQPTQGAMPAFASAFQRRAG—GVLVASHLQSFLEVSYRVLRHLAQP-(OH)

Chemical Formula: $C_{237}H_{376}N_{70}O_{63}S_2$ **Molecular Weight (g / mol):** 5,278.16

See Figure 1.31A and Figure 1.24B ESI-MS (+) of peptide **13**. M = $C_{237}H_{376}N_{70}O_{63}S_2$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1760.4, found: 1760.7; $[M+4H^+]^{4+}$ *m/z*: 1320.6, found: 1320.8; $[M+5H^+]^{5+}$ *m/z*: 1056.6, found: 1056.8; $[M+6H^+]^{6+}$ *m/z*: 880.7, found: 880.8; $[M+7H^+]^{7+}$ *m/z*: 750.0, found: 750.3. * indicates presence of TFA adduct in spectrum.

Preparation of Glycopeptide G-CSF 126–135 Thioacid (19).



G-CSF (126–135) **19**

Sequence (19a) (thiocarboxylic acid)**:**

H-MAPALQPN(chitobiose)QG-(SH)

Chemical Formula: $C_{74}H_{107}N_{15}O_{25}S_2$ **Molecular Weight (g / mol):** 1,670.87

Sequence (19b) (carboxylic acid)**:**

H-MAPALQPN(chitobiose)QG-(OH)

Chemical Formula: $C_{74}H_{107}N_{15}O_{26}S$ **Molecular Weight (g / mol):** 1,654.81

Fmoc-**M**APALQP**D**(OAll)**Q**-(OH)

Following the general procedure for Fmoc-based SPPS on an Applied Biosystems Pioneer with deblock 4:1 DMF/piperidine containing 0.1 M Oxyma Pure, peptide (G-CSF 126-134) was prepared on 0.19 mmol scale from Fmoc-Gln-NovaSyn® TGT resin (residue 134), Fmoc-Asp(OAll)-OH (residues 133) and Fmoc-Met-OH (residue 126). The peptide resin was subjected to mild cleavage protocol.

The lyophilized material (assumed 0.19 mmol) was dissolved in chloroform (12 mL) and cooled to 0 °C with an ice-water bath. To the solution was added **11** (106 mg, 1.5 equiv), HOOBt (93 mg, 3 equiv) and EDC (101 µL, 3 equiv). The resultant yellow solution was stirred for 3 h at 0 °C, allowed to reach RT, stirred at RT for 1h, diluted with 5 vol % HOAc in chloroform (20 mL) and washed with H_2O (2 × 5 mL). The organic layer was washed with a saturated aqueous brine solution (5 mL), dried over Na_2SO_4, filtered and concentrated en vacuo to yield fully protected (G-CSF 126-135) as a crude yellow oil. The oil was diluted with HOAc (2 mL), precipitated with H_2O (60 mL), shell frozen and lyophilized.

Fmoc-**M**APALQP**D**(OH)**Q**G-(SCH_2Ar)

After lyophilization, the resultant fully protected peptide (assumed 0.19 mmol by mass recovered, G-CSF 126-135) was dissolved in chloroform (12 mL) followed by the addition of phenyl silane (936 µL, 40 equiv) and $Pd(PPh_3)_4$ (6.6 mg, 3 mol %). The light yellow reaction was stirred at RT shielded from light for 60 min until judged complete by UPLC-MS analysis of a reaction aliquot indicating complete deallylation at residue Asp133. The reaction was concentrated at RT with a stream of Ar, diluted with HOAc (1 mL) and centrifugated to remove insoluble salts by decantation. The decanted layer was precipitated with H_2O (20 mL), shell frozen and lyophilized to provide G-CSF 126-135 as an off-white solid.

Fmoc-**M**APALQP**N**(chitobiose)**Q**G-(SCH$_2$Ar)

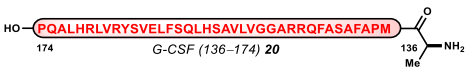
The resultant lyophilized powder, G-CSF 126-135 (assumed 0.19 mmol) was dissolved in anhydrous DMSO (19 mL) followed by the addition of chitobiose amine (**S-5**, 241 mg, 3 equiv) and PyAOP (297 mg, 3 equiv).[10] The homogeneous reaction was stirred at RT for 1 min followed by the addition of *i*-Pr$_2$NEt (100 µL, 3 equiv). The reaction was stirred for 45 min at RT at which point the reaction was judged complete by UPLC-MS analysis of the protected peptide indicating consumption of starting carboxylic acid. The material was precipitated with cold H$_2$O (60 mL), centrifugated and the resultant pellet was redissolved in H$_2$O/CH$_3$CN (1:1, 0.05% TFA), shell frozen and lyophilized.

The resultant powder was treated with a modified cocktail B (12 mL, 2 h) following the general procedure (Modified Cocktail B, water removed and solution is diluted: 200 mg phenol dissolved in methylene chloride (12 mL) and TFA (12 mL), add 600 µL *i*-Pr$_3$SiH). The obtained crude pellet was dissolved in H$_2$O/CH$_3$CN (1:1, 0.05% TFA) and purified by HPLC (Microsorb 300-5 C8 column, 10% to 60% solvent B over 30 min). Product **19** eluted at 13.0 min. Lyophilization of the collected fractions provided peptide **19** (67 mg, 21%, 3:1 **19a**:**19b**) as a white solid.

Desired glycopeptide **19a** was isolated and characterized as a presumed 3:1 mixture with its corresponding carboxylic acid **19b** (relative integration of UV-trace for peaks corresponding to **19a** and **19b**). This hydrolysis presumably occurs during HPLC purification and prolonged storage of crude **19** in aqueous solution. This mixture was used without further purification in subsequent isonitrile-mediated ligations at RT. Subsequent yields are provided based upon this ratio (3:1) as well as total mass-based yield in parentheses (See **Figure 1.32D**)

See Figure 1.32A, Figure 1.32A, Figure 1.32A and Figure 1.32D Top: ESI-MS (+) of peptide **19b**. M = $C_{74}H_{107}N_{15}O_{26}S_1$. ESI calculated for $[M+H^+]^{1+}$ *m/z*: 1655.82, found: 1656.24; $[M+2H^+]^{2+}$ *m/z*: 828.41, found: 828.62. **Bottom:** ESI-MS (+) of peptide **19a.** M = $C_{74}H_{107}N_{15}O_{25}S_2$. ESI calculated for $[M+H^+]^{1+}$ *m/z*: 1671.88, found: 1672.45; $[M+2H^+]^{2+}$ *m/z*: 836.44, found: 836.66.

Preparation of G-CSF 136–174 (20).



Sequence (20):

H-AMPAF**AS**AFQRRAGGVLV**AS**HL**QS**FLE**VS**YRVLRHLAQ**P**-(OH)

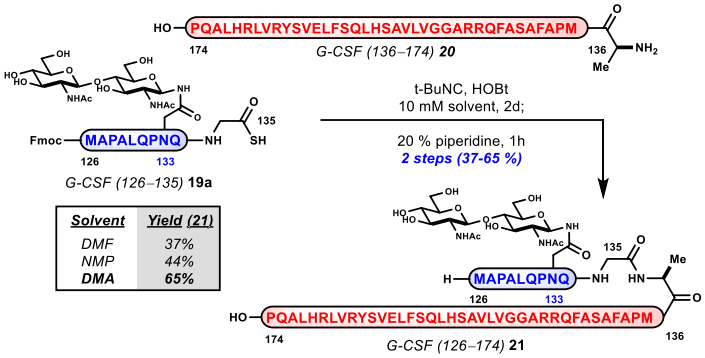
Chemical Formula: $C_{194}H_{306}N_{58}O_{50}S$ **Molecular Weight (g / mol):** 4,282.99

Following the general procedure for microwave-assisted Fmoc-based SPPS with deblock 4:1 DMF/piperidine, peptide (G-CSF 126-174) was prepared on 0.1 mmol scale from Fmoc-Pro-NovaSyn® TGT resin (residue 174), Fmoc-Val-Ser($\Psi^{Me,Me}$Pro)-OH (residues 163-164), Fmoc-Gln(Trt)-Ser($\Psi^{Me,Me}$Pro)-OH (residues 158-159), and Fmoc-Ala-Ser($\Psi^{Me,Me}$Pro)-OH (residues 154-155, 141-142). The peptide resin was subjected to mild cleavage protocol.

After lyophilization, the resultant fully protected peptide (0.1 mmol, G-CSF 136-174) was treated with cocktail B (20 mL, 2h, twice) following the general procedure. A portion of obtained crude pellet (0.05 mmol) was dissolved in H$_2$O/CH$_3$CN (1:1, 0.05% TFA) and purified by HPLC (Microsorb 300-5 C8 column, 40% to 65% solvent B over 30 min). Product **20** eluted at 11.0 min. Lyophilization of the collected fractions provided peptide **20** (81 mg, 38%) as a white solid.

See Figure 1.33A and Figure 1.33B ESI-MS (+) of peptide **20**. M = $C_{194}H_{306}N_{58}O_{50}S$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1428.67, found: 1428.21; $[M+4H^+]^{4+}$ *m/z*: 1071.76, found: 1072.16; $[M+5H^+]^{5+}$ *m/z*: 857.61, found: 857.99; $[M+6H^+]^{6+}$ *m/z*: 714.84, found: 715.12.

Synthesis of Glycopeptide G-CSF 126–174 (13).



Sequence (21):

H-MAPALQPN(chitobiose)QG—AMPAFASAFQRRAGGVLVASHLQSFLEVSYRVLRHLAQP-(OH)

Chemical Formula: $C_{253}H_{401}N_{73}O_{73}S_2$ **Molecular Weight (g / mol):** 5697.55

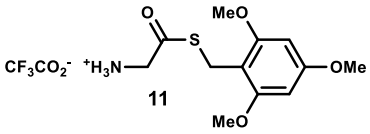
To an oven-dried vial was added thioacid G-CSF 126–135 **19** (10.0 mg, 6.0 µmol, 1.0 equiv), N-terminal amine **20** (38.4 mg, 9.0 µmol, 1.5 equiv), and HOBt (9.2 mg, 59.9 µmol, 10 equiv). A stir bar was added, followed by the addition of anhydrous DMF to a final concentration of 10 mM **19** (600 µL). The heterogeneous mixture was stirred at RT for 5 min followed by the addition of *tert*-butyl isonitrile via syringe from a 100 mM stock-solution prepared in the corresponding solvent (17.9 µmol, 3.0 equiv). The reactions was capped and stirred for 48h. At t = 48 h, the reaction was diluted with deblock solution (20 vol % piperidine in DMF, 1 mL) and stirred vigorously at RT for 1h, cooled in dry ice, followed by precipitation with cold diethyl ether. The precipitate was centrifugated (5 min, 5 °C, 5000 rpm) and the supernatant was decanted. The resultant pellet was frozen in liquid nitrogen and lyophilized prior to purification by semi-preparative HPLC. The crude material (containing excess **20** and product **21**, see **Figure 1.34A and Figure 1.34B**) was dissolved in H_2O/CH_3CN (1:1, 0.05% TFA) and purified by HPLC (Microsorb 300-5 C18 column, 20% to 70% solvent B over 30 min). Product **21** elutes first at 21.0 min and partially co-elutes with recovered **20** (22 to 24 min, 4.4 mg mixed). Desired fractions were pooled and lyophilized to yield **21** (9.4 mg, 37%* (28%), steps) as a white powder.

See Figure 1.35A and Figure 1.35B ESI-MS (+) of peptide **21**. M = $C_{253}H_{401}N_{73}O_{73}S_2$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1900.19, found: 1900.34; $[M+4H^+]^{4+}$ *m/z*: 1425.40, found: 1425.65; $[M+5H^+]^{5+}$ *m/z*: 1140.52, found: 1140.55; $[M+6H^+]^{6+}$ *m/z*: 950.60, found: 950.87; $[M+7H^+]^{7+}$ *m/z*: 814.94, found: 815.20.

Entry 1 (12.5 mM **19** in DMF, 3 µmol scale) – Isolated product **21** (5.6 mg, 44%* (33%)), recovered **20** (11.7 mg).
Entry 2 (12.5 mM **19** in NMP, 3 µmol scale) – Isolated product **21** (5.6 mg, 44%* (33%)), recovered **20** (6.0 mg).
Entry 3 (12.5 mM **19** in DMA, 3 µmol scale) – Isolated product **21** (8.3 mg, 65%* (49%)), recovered **20** (6.0 mg).
Large scale (1) (10 mM **19** in DMF, 6 µmol scale) – Isolated product **21** (9.4 mg, 37%* (28%)).
Large scale (2) (10 mM **19** in DMA, 12 µmol scale) – Isolated product **21** (25.0 mg, 49%* (37%)), recovered **20** (26.0 mg).
* Calculation based on the presumed mixture of **19** (**19a**:**19b**, 3:1) as described in the above preparation.

Part V. Preparation of Thioesters (11, S-1, S-2, S-3) and Chitobiose Derivatives (S-4, S-5).

Preparation of Gly Thioester Trifluoroacetate (11).



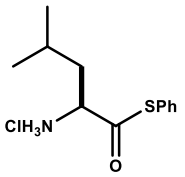
11

Chemical Formula (freebase)**:** $C_{12}H_{17}NO_4S$ **Molecular Weight (g / mol):** 271.33

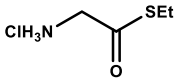
To a solution containing Boc-Gly-OH (339 mg, 1.94 mmol, 1 equiv) and HBTU (736 mg, 1.94 mmol, 1 equiv) dissolved in acetonitrile (5.9 mL) was added *i*-Pr_2NEt (0.78 mL, 4.46 mmol, 2.3 equiv). The light yellow solution was stirred for 2 min at RT followed by the addition of (2,4,6-trimethoxyphenyl)methanethiol (500 mg, 2.33 mmol, 1.2 equiv). The reaction was stirred at RT until judged complete by TLC analysis. The solution was concentrated en vacuo onto 7 g Celite and purified on a 40 g SiO_2 flash chromatography cartridge. The product was eluted with a stepwise gradient of hexanes/ethyl acetate (hexanes, 100%, 2 CV → hexanes/ethyl acetate, 9:1, 3 CV → hexanes/ethyl acetate, 7:3, 9 CV). The desired eluted during hexanes/ethyl acetate, 7:3, 9 CV and has R_f = 0.4 (hexanes/ethyl acetate, 7:3), stains with $KMnO_4$ solution. Desired fractions were combined and concentrated to provide the corresponding Boc-protected thioester as a white solid. This material (700 mg) was dissolved in methylene chloride (14 mL) and treated with TFA (5.5 mL) for 15 min at RT. Analysis by UPLC-MS indicates complete Boc- removal. The reaction was concentrated en vacuo, redissolved in H_2O/CH_3CN (7:3, 0.05% TFA), shell frozen and lyophilized to provide pure **11** (746 mg, 97%) as a white solid.

See Figure 1.36A ^{1}H NMR (600 MHz, D_2O): 6.28 (s, 2H), 4.28 (s, 2H), 4.14 (s, 2H), 3.83 (s, 9H).
See Figure 1.36B ^{1}H NMR (600 MHz, 6:1 $CDCl_3$ / DMSO-d_6): 7.24 (s, 2H), 5.45 (s, 2H), 5.12 (s, 2H), 4.95 (s, 3H), 4.93 (s, 6H).
See Figure 1.36C ^{13}C NMR (150 MHz, 6:1 $CDCl_3$ / DMSO-d_6): 193.4, 161.0, 158.9, 103.0, 90.2, 55.5, 55.1, 46.8, 22.1.
LRMS (ESI): calculated for M = $C_{12}H_{17}NO_4S$, $[M+H^+]^+$: 272.3, found: 272.2.

Preparation of Leu Thioester Hydrochloride (S-1) and Gly Thioester Hydrochloride (S-2).



S-1
H-Leu-SPh•HCl



S-2
H-Gly-Et•HCl

Leu Thioester Hydrochloride (**S-1**) and Gly Thioester Hydrochloride (**S-2**) were prepared from Boc-Leu-OH and thiophenol and Boc-Gly-OH and ethane thiol respectively according to Stuhr-Hansen *et al.*[5b] The corresponding Boc-protected derivatives were treated in methylene chloride with 4 M HCl in 1,4-dioxane at RT followed by concentration, dilution of the incipient residues H_2O/CH_3CN (7:3, 0.05% TFA), shell frozen and lyophilized to provide **S-1** and **S-2** of suitable purity for subsequent reactions.

Leu Thioester Hydrochloride (S-1).

Chemical Formula (freebase)**:** $C_{12}H_{17}NOS$ **Molecular Weight (g / mol):** 223.33

See Figure 1.37A **^{1}H NMR** (600 MHz, D_2O): 7.62-7.54 (m, 1H), 4.54-4.48 (m, 1H), 2.04-1.97 (m, 2H), 1.06 (d, *J* = 6.3 Hz, 3H), 1.02 (d, *J* = 6.3 Hz, 3H).

See Figure 1.37B **^{13}C NMR** (150 MHz, D_2O): 199.0, 134.8, 130.9, 130.0, 124.5, 57.4, 40.1, 24.1, 21.8, 20.8.

LRMS (ESI): calculated for M = $C_{12}H_{17}NOS$, $[M+H^+]^+$: 224.34, found: 224.30.

Gly Thioester Hydrochloride (S-2).

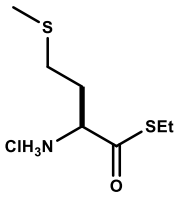
Chemical Formula (freebase)**:** C_4H_9NOS **Molecular Weight (g / mol):** 119.18

See Figure 1.38A **^{1}H NMR** (600 MHz, D_2O): 4.17 (s, 2H), 3.05 (q, *J* = 7.4 Hz, 2H), 1.29 (t, *J* = 7.4 Hz, 3H).

See Figure 1.38B **^{13}C NMR** (150 MHz, D_2O): 195.4, 47.0, 23.5, 13.7.

LRMS (ESI): calculated for M = $C_{12}H_{17}NOS$, $[M+H^+]^+$: 120.19, found: 120.10.

Preparation of Met Thioester Hydrochloride (S-3).



S-3
H-Met-SEt•HCl

Chemical Formula (freebase)**:** $C_7H_{15}NOS_2$ **Molecular Weight (g / mol):** 193.32

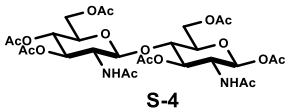
To a suspension containing Boc-Met-OH (1.56 g, 1 equiv) and HOBt (2.50 g, 1 equiv) in methylene chloride (32 mL) was added EDC (1.12 mL, 1 equiv). The heterogeneous reaction was stirred at RT for 5 min followed by the addition of ethane thiol (0.93 mL, 2.0 equiv). The reaction becomes homogeneous and light yellow. The reaction was stirred at RT for 12h, concentrated onto Celite and purified on a 80 g SiO_2 flash chromatography cartridge. The product was eluted with a stepwise gradient of hexanes/ethyl acetate (hexanes, 100%, 2 CV → hexanes/ethyl acetate, 9:1, 2 CV → hexanes/ethyl acetate, 7:3, 5 CV). The desired eluted during hexanes/ethyl acetate, 7:3, 5 CV and has R_f = 0.6 (hexanes/ethyl acetate, 7:3), stains with $KMnO_4$ solution. Desired fractions were combined and concentrated to provide the corresponding Boc-protected thioester as a white solid. This material (904 mg) was dissolved in methylene chloride (1.7 mL), cooled in an ice water bath and treated with 4.0 M HCl in 1,4-dioxane (1.7 mL) for 30 min at 0 °C and 60 min at RT. Analysis by UPLC-MS indicates complete Boc- removal. The reaction was concentrated en vacuo, redissolved in H_2O/CH_3CN (7:3, 0.05% TFA), shell frozen and lyophilized to provide pure **S-3** (640 mg, 45%) as an off-white solid.

See Figure 1.39A **^{1}H NMR** (500 MHz, D_2O): 4.50-4.44 (m, 1H), 3.12-2.97 (m, 2H), 2.73-2.66 (m, 2H), 2.38-2.17 (m, 2H), 2.15 (s, 3H), 1.32-1.25 (m, 3H).

See Figure 1.39B **^{13}C NMR** (125 MHz, D_2O): 198.4, 58.0, 30.1, 28.2, 23.8, 13.9, 13.5.

LRMS (ESI): calculated for M = $C_7H_{15}NOS_2$, $[M+H^+]^+$: 194.32, found: 194.20.

Preparation of Chitobiose Octaacetate (S-4).



S-4

Chemical Formula: $C_{28}H_{40}N_2O_{17}$ **Molecular Weight (g / mol):** 676.63

Chitobiose octaacetate was prepared with minor modifications to the reported procedure.[11] Chitin (10g, coarse flakes from shrimp shells) was suspended in 73-mL conc. HCl in a 1-L Erlenmeyer flask. The resultant light brown viscous solution was sonicated at 30 °C for 90 min followed by the addition of ice (300 g) and ice cold H_2O (500 mL) to provide a momentarily clear solution. After 2-5 min, the solution precipitates a white powder. The precipitous solution was stored overnight at 4 °C. This procedure was repeated with another batch of chitin (10 g) and the combined precipitates were isolated by vacuum filtration, washed with 1 M aqueous NaOH (40 mL) and H_2O. The resultant white powder was suspended in H_2O (100 mL), shell frozen and lyophilized to provide 14.6 g of crude chitin after acidic treatment.

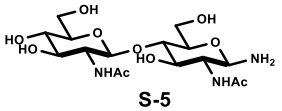
A portion of this material (6.3 g) was suspended in phosphate buffer (0.1 M Na_2HPO_4, pH = 6.5) and the final pH was adjusted to pH = 6.4 with 6 M HCl. The reaction was heated in a 37 °C oil bath followed by the immediate addition of 30 mg chitinase (539 units/mg, ~16 units). After stirring at 37 °C for 15 days (reaction progress can be monitored by crude aliquots analyzed by ^{1}H NMR in D_2O). The crude reaction was cooled to RT, filtered, shell-frozen and lyophilized to provide 7.06 g of chitobiose.

A portion of this material (6.39 g) and NaOAc (7.8 g) was suspended in neat Ac_2O (72 mL) at RT. The reaction was heated in an 80 °C oil bath for 48h, cooled to RT and stirred overnight (10 h). The reaction was diluted with chloroform (250 mL) and carefully washed with saturated aqueous $NaHCO_3$ solution (2 × 20 mL) and saturated aqueous brine solution (10 mL). Layers were allowed to separate and the organic layer was dried over Na_2SO_4, filtered and concentrated en vacuo to provide crude **S-4** (5.70 g) as an off-white solid. The crude material was recrystallized from boiling methanol (55 mL), cooled to RT, allowed to sit at RT for 5h followed by storage at -20 °C for 48h. Filtration of recrystallized material provided **S-4** (3.1 g) as a white solid of purity suitable for subsequent reactions.

See Figure 1.40A ^{1}H NMR (500 MHz, $CDCl_3$): 6.10 (d, *J* = 3.6 Hz, 1H), 5.92 (d, *J* = 8.8 Hz, 1H), 5.62 (d, *J* = 9.2 Hz, 1H), 5.22 (dd, *J* = 9.2, 11.2 Hz, 1H), 5.13 (app. t, *J* = 9.6 Hz, 1H), 5.05 (app. t, *J* = 9.6 Hz, 1H), 4.40 (m, 4H), 4.18 (dd, *J* = 1.6, 12.0 Hz, 1H), 4.02 (dd, *J* = 1.6, 12.0 Hz, 1H), 3.92 (m, 2H), 3.73 (app. t, *J* = 9.2 Hz, 1H), 3.62 (m, 1H), 2.18-1.92 (8 × s, 24H).

See Figure 1.40B ^{13}C NMR (125 MHz, $CDCl_3$): 171.7, 171.5, 171.0, 170.7, 170.6, 170.3, 169.5, 169.1, 102.0, 90.7, 76.2, 72.8, 72.2, 71.0, 70.9, 68.1, 61.9, 61.6, 54.7, 51.4, 23.4, 23.3, 21.22, 21.20, 20.9, 20.8, 20.8.

Preparation of Chitobiose Amine (S-5).



S-5

Chemical Formula: $C_{16}H_{29}N_3O_{10}$ **Molecular Weight (g / mol):** 423.42

Chitobiose octaacetate (**S-4**, 2.00 g, 1 equiv) was suspended in methylene chloride (32 mL) followed by the addition of 7 M NH_3 in methanol (96 mL).[11] The clear solution was stirred for 48h at RT, concentrated to approximately 30 mL of methanol at which point substantial precipitation occurs. The precipitous solution was diluted with ethyl acetate (10 mL) and

equally portioned as a slurry into 4 centrifugation tubes (~10 mL each, 50-mL volume tubes). The tubes were each diluted with ice cold ethyl acetate (30 mL) and centrifugated. This precipitation process was repeated and resultant pellets were recombined, dissolved in water (120mL), shell frozen and lyophilized to provide crude chitobiose (1.3 g).

The crude chitobiose (1.3 g) was dissolved in saturated aqueous NH_4HCO_3 solution (prepared by dissolution of 30 g NH_4HCO_3 in H_2O (25 mL)). The precipitous solution was stirred for 72h at RT, filtered through a cotton plug (washed 2 × 10 mL H_2O) and combined filtrates were further diluted with H_2O (160 mL), shell frozen and lyophilized three times from pure H_2O until a stable mass was obtained. This procedure affords chitobiose amine (**S-5**, 1.26 g, 94%, 2 steps) as a white solid of purity suitable for subsequent reactions.

See Figure 1.41A ^{1}H NMR (600 MHz, D_2O): 4.60 (d, *J* = 8.5 Hz, 1H), 4.17 (d, J = 8.7 Hz, 1H), 3.98-3.48 (m, 13H), 2.09 (s, 3H), 2.06 (s, 3H).

See Figure 1.41B ^{13}C NMR (155 MHz, D_2O): 174.63, 174.57, 101.5, 84.1, 79.7, 75.9, 75.3, 73.4, 73.1, 69.7, 60.5, 60.3, 55.7, 55.6, 22.2, 22.1.

LRMS (ESI): calculated for M = $C_{16}H_{29}N_3O_{10}$, $[M+H^+]^+$: 424.43, found: 424.40.

Part VI. Characterization Data for G-CSF 1–174 (22) and G-CSF 1–174 Aglycone (2a).

Circular Dichroism Spectra for G-CSF 1–174 Aglycone (2a).

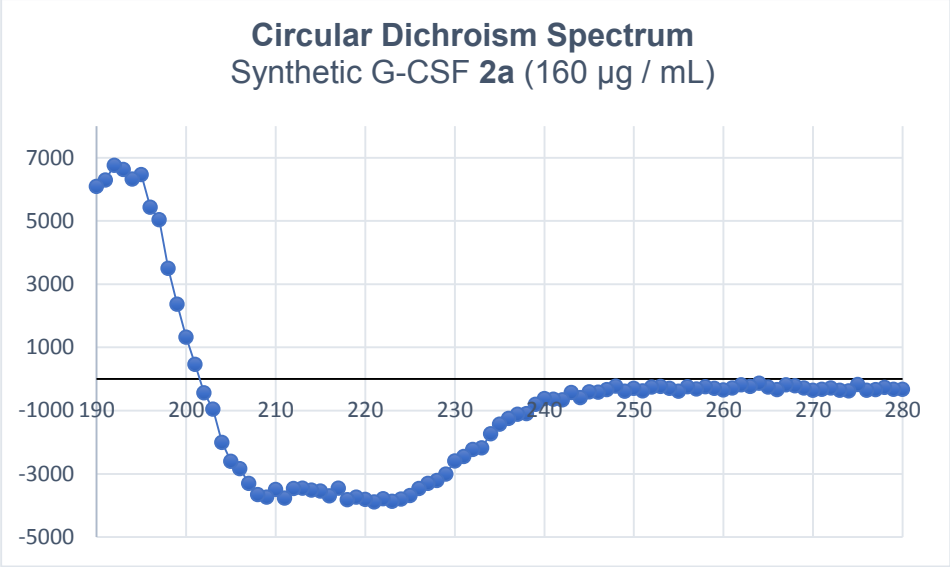


FIGURE S-1. Circular Dichroism Spectrum for Synthetic G-CSF 1–174 (2a) Acquired with Aviv Circular Dichroism Model 62DS. The sample is G-CSF (**2a**) at est. 160 µg / mL in pH = 7.0 water. Data was acquired at 25 °C and is plotted as the collected average of 3 scans. y axis = [Θ] deg cm^2/dmol, x axis = λ nm (190 to 280 nm)

The sample synthetic G-CSF (**2a**) was prepared by exchange from buffer 3 (**2a**, 80 µg / mL, 1000 µL) into water (**2a**, est. 160 µg / mL, 500 µL). The water used was pH = 7.0. Solvent exchange and concentration were made using a 10,000 MWCO Amicon centrifugation filter at 3000 rpm, 15 °C (3 × 10 min cycles, with re-dilution each cycle). The reduced signal intensity for synthetic G-CSF **2a** (molar elipticity, [Θ], reduced from anticipated signal of ~20,000 deg cm^2/dmol) has been previously observed in reports where recombinant G-CSF and pegylated G-CSF samples aggregate during thermal experiments or during formulation.[12,13] Further supportive characterization was performed with this sample of synthetic G-CSF **2a** by additional methods including SDS-PAGE and proteolytic mapping using recombinant G-CSF **2b** as a standard control.[14]

SDS-PAGE Data for G-CSF 1–174 (22) and G-CSF 1–174 Aglycone (2a).

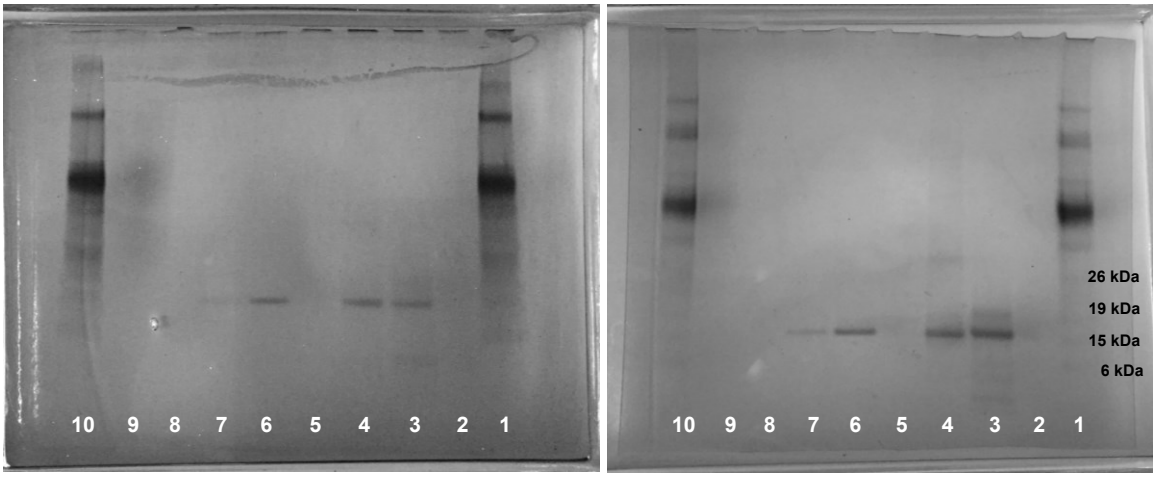


FIGURE S-2. SDS-PAGE analysis for synthetic G-CSF 1–174 (22), synthetic G-CSF 1–174 Aglycone (2a) and recombinant G-CSF 0–174 (2b). Left: Samples prepared under reducing conditions. **Right:** Samples prepared under non-reducing conditions. Gels were developed for 4 min 35 secs. **Lanes.** 1 – MW Marker, Benchmark™ Protein Ladder, 2 – blank (no sample), 3 – Non-folded G-CSF **22** (26 µL from stock estimated at 200 µg/mL, AGR-IV-69), 4 – Folded G-CSF **2a** (26 µL from stock calc. at 92 µg/mL, AGR-III-250, synthetic, February 2015), 5 – blank (no sample), 6 – Amgen G-CSF **2b** (20 µL of 450 µg/mL = 9 µg loaded), 7 – Amgen G-CSF **2b** (20 µL of 150 µg/mL = 3 µg loaded), 8 – Amgen G-CSF **2b** (20 µL of 50 µg/mL = 1 µg), 9 – blank (no sample), 10 – MW Marker, Benchmark™ Protein Ladder.

Enzymatic Digestion and Disulfide Mapping of Synthetic G-CSF (2a) Compared with Recombinant G-CSF (2b). Data was acquired with a full MS spectrum (*m/z* 300 to 1650, AGC: 1e6) at resolution R = 120000 followed by 10 MS/MS scans (isolation width: 2.0) of the most abundant ions from the full MS scan.

Enzymatic Digestion and Disulfide Mapping. Disulfide mapping was performed using the method of Mo *et al*. with slight modification.[14] Non-reduced recombinant human granulocyte-colony stimulating factor (**2b**, rhG-CSF) and non-reduced synthetic GCSF **2a** were prepared by alkylating the sample with N-ethylmaleimide (NEM) at a 1:100 in excess molar ratio and incubating at 37 °C for 2 h at pH = 7.0. The samples were then digested overnight with Glu-C (25:1 protein / enzyme ratio) at 37 °C. A 1 µL aliquot of chymotrypsin (1 µg / µL) was added to the samples and incubated at 37 °C for an additional 4 h. The reduced samples were prepared by adding 1:1 (v/v) 100mM dithiothreitol (DTT) and incubating at 65 °C for 30 min. All the samples were desalted using a C18 disc (Empore). Digested and desalted peptides were then analyzed using a reverse phase C18 column (ACQUITY UPLC Peptide BEH C18 nanoACQUITY Column, 10K psi, 130Å, 1.7 µm, 100 µm × 100 mm; Waters Corp, Milford, MA) on a Waters nanoACQUITY UPLC with a gradient of 0 to 50 % Buffer B (0.1% formic acid in CH_3CN) over 90 min at a flow rate of 300 nL / min. High-resolution MS (LC-MS) data was collected on an OrbiTrap Elite mass spectrometer (Thermo Scientific). Data was acquired with a full MS spectrum (*m/z* 300 to 1650, AGC: 1e6) at resolution R = 120000 followed by 10 MS/MS scans (isolation width: 2.0) of the most abundant ions from the full MS scan.

Cleavage peptides were identified in accord with observations described by Mo *et al.*[14] Sequence numbering is consistent with the manuscript, where the N-terminal Met in recombinant G-CSF **2b** is position 0 (Met0).

See FIGURE S-3. Data for sequence Lys34–Glu46 which comprises intact disulfide (non-reduced) Cys36–Cys42, $[M+2H^+]^{2+}$ and $[M+3H^+]^{3+}$ were observed in both samples **2a** and **2b** under non-reducing conditions.

Ion search for $[M+2H^+]^{2+}$ calculated value: 766.8654 and $[M+3H^+]^{3+}$ calculated value: 511.5794.

See FIGURE S-4. Data for sequence Lys34–Glu46 which comprises free sulfhydryls (reduced) at Cys36 and Cys42, $[M+2H^+]^{2+}$ and $[M+3H^+]^{3+}$ were observed in both samples **2a** and **2b** under reducing conditions.

Ion search for $[M+2H^+]^{2+}$ calculated value: 767.8733 and $[M+3H^+]^{3+}$ calculated value: 512.2513.

See FIGURE S-5. Data for sequence Ser62–Leu75 which comprises intact disulfide (non-reduced) Cys64–Cys74, $[M+H^+]^{1+}$ and $[M+2H^+]^{2+}$ were observed in both samples **2a** and **2b** under non-reducing conditions.

Ion search for $[M+H^+]^{1+}$ calculated value: 1375.6345 and $[M+2H^+]^{2+}$ calculated value: 688.3209.

See FIGURE S-6. Data for sequence Ser62–Leu75 which comprises free sulfhydryls (reduced) at Cys64 and Cys74, $[M+H^+]^{1+}$ and $[M+2H^+]^{2+}$ were observed in both samples **2a** and **2b** under reducing conditions.

Ion search for $[M+H^+]^{1+}$ calculated value: 1377.6501 and $[M+2H^+]^{2+}$ calculated value: 689.3287.

Part VII. Biological Assay Data for G-CSF 1–174 Aglycone (2a).

A proliferation assay of murine myeloblastic NFS-60 cells and a Colony-Forming assay of human primary cord blood $CD34^+$ cells were used to validate the biological activity of synthetic G-CSF 1–174 aglycone (**2a**) compared with control, recombinant G-CSF 0–174 (**2b**) (Neupogen®, Amgen Inc., CA).[15,16]

Assay I. Murine Myeloblastic NFS-60 Cells.

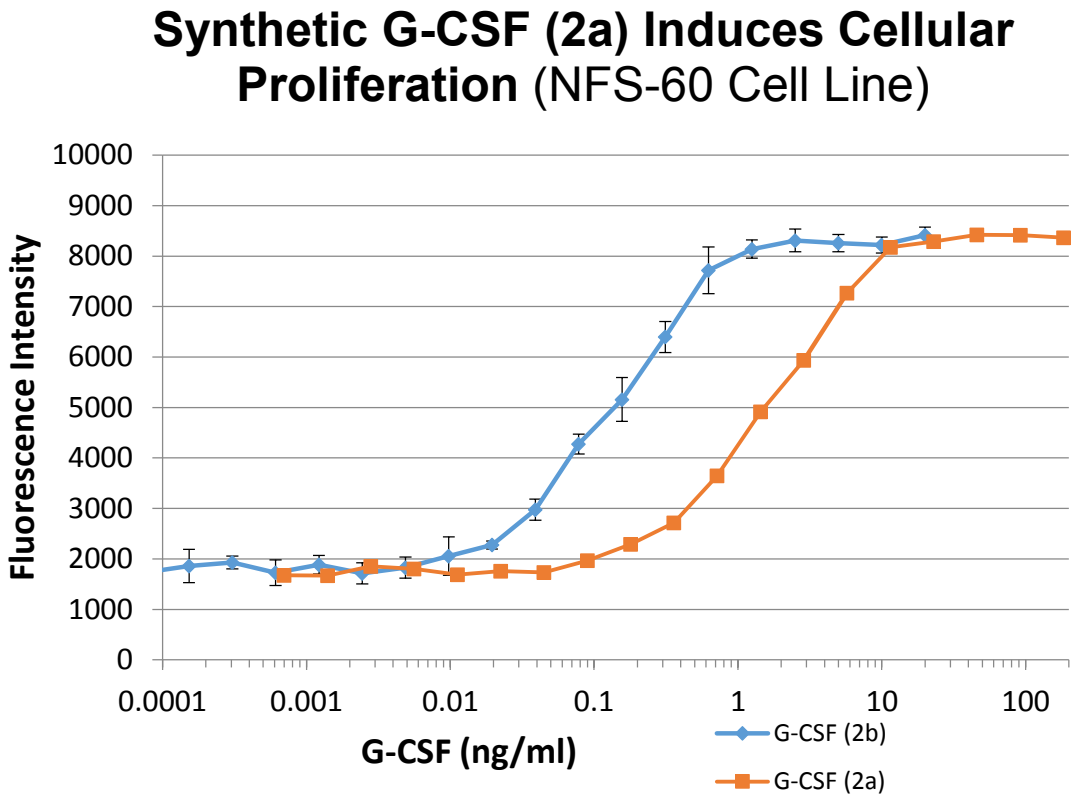


FIGURE S-7. Murine Myeloblastic NFS-60 Cellular Proliferation Assay: In-vitro assay measuring the effect of synthetic G-CSF **2a** and recombinant G-CSF **2b** (Amgen) on the proliferation of NFS-60 cells. Results are expressed as average relative fluorescent intensity ± SD, run in triplicate.

The murine myeloblastic NFS-60 cells were maintained in Iscove's modified Dulbecco's medium (IMDM) containing 10% Fetal Calf Serum, 20% WEHI-3 conditioned medium, 0.05 mM monothiolglycerol, 2 mM glutamine, 50 units/mL penicillin, and 50 µg/mL streptomycin. When the NFS-60 cells were in the log-phase of growth, the cultures were harvested and washed with three times with 10 mL cold IMDM. The 4,000 washed NFS-60 cells/well were cultured in 50 µL IMDM containing 20% Knockout™ serum replacement (KOSR) (Life Technologies Inc., NY), 0.05 mM monothiolglycerol, 2 mM glutamine, 50 units/ml penicillin, 50 µg/ml streptomycin with or without various doses of synthetic G-CSF **2a** or recombinant G-CSF **2b** in 5% CO_2, at 37 °C in triplicate. After two days, the cultures were pulsed with AlamarBlue® overnight. A H1 synergy plate reader was used to measure fluorescence intensity. As shown in Figure S-1, the biological activity of synthetic G-CSF **2a** is approximately 10% that of recombinant G-CSF **2b**. The half maximal effective concentration (EC_{50}) to induce proliferation of NFS-60 cell line is ~2 ng/ml for synthetic G-CSF **2a** and ~0.2 ng/ml for recombinant G-CSF **2b**.

Assay II. Colony Forming Assay.

	G-CSF (ng/ml)	KL (ng/ml)	large colony (>50 cells / colony)	small colony (<50 cells / colony)
	0.00	0.00	0±0	0±0
	0.00	20.00	0±0	6±2
G-CSF **2a** (synthetic)	20.00	0.00	31±4	11±1
	0.02	20.00	9±2	33±3
	0.08	20.00	21±2	49±6
	0.31	20.00	21±5	42±8
	1.25	20.00	50±4	35±4
	5.00	20.00	80±2	34±1
	20.00	20.00	110±10	44±3
G-CSF **2b** (recombinant)	20.00	0.00	45±7	25±6
	0.02	20.00	10±1	29±4
	0.08	20.00	55±4	46±8
	0.31	20.00	92±23	63±3
	1.25	20.00	102±5	38±10
	5.00	20.00	107±3	43±4
	20.00	20.00	121±12	40±3

TABLE S-1. Effect of Synthetic G-CSF (**2a**) and Recombinant G-CSF (**2b**) on Colony-Formation of Human Cord Blood $CD34^+$ Cells.

To further characterize the biological activity of synthetic G-CSF **2a** in normal human hematopoietic progenitor cells, a procedure of colony-forming assay in serum-free media was carried out as described previously with a minor modification.[15,16] One thousand purified human primary cord blood $CD34^+$ cells were cultured in 1 mL of IMDM containing 1.2 % methylcellulose, 30 % KOSR, 1 × Insulin / Transferrin / Selenium solution (Life Technologies Inc.), 0.05 mM monothiolglycerol, 2 mM glutamine, 50 units/mL penicillin, 50 µg/mL streptomycin, and with or without various doses of synthetic G-CSF **2a**, recombinant **2b** or human klotho protein (KL) (20 ng/mL) alone or in combination in 5% CO_2 at 37°C in triplicate. After fourteen days, numbers of clusters (< 50 cells/colony) and colony (> 50 cells/colony) were scored under a microscope. As shown in Table S-1, Figure S-2 and Figure S-3, synthetic G-CSF **2a** alone is capable of stimulating proliferation of normal human $CD34^+$ cells to form clusters and colonies. It also synergizes with human KL to form clusters and colonies. An EC_{50} of recombinant G-CSF **2b** synergized with human KL is ~0.08 ng/ml and ~1.25 ng/ml for synthetic G-CSF **2a**, indicating the specific biological activity of synthetic G-CSF **2a** is about ~6.4% of that of the recombinant.

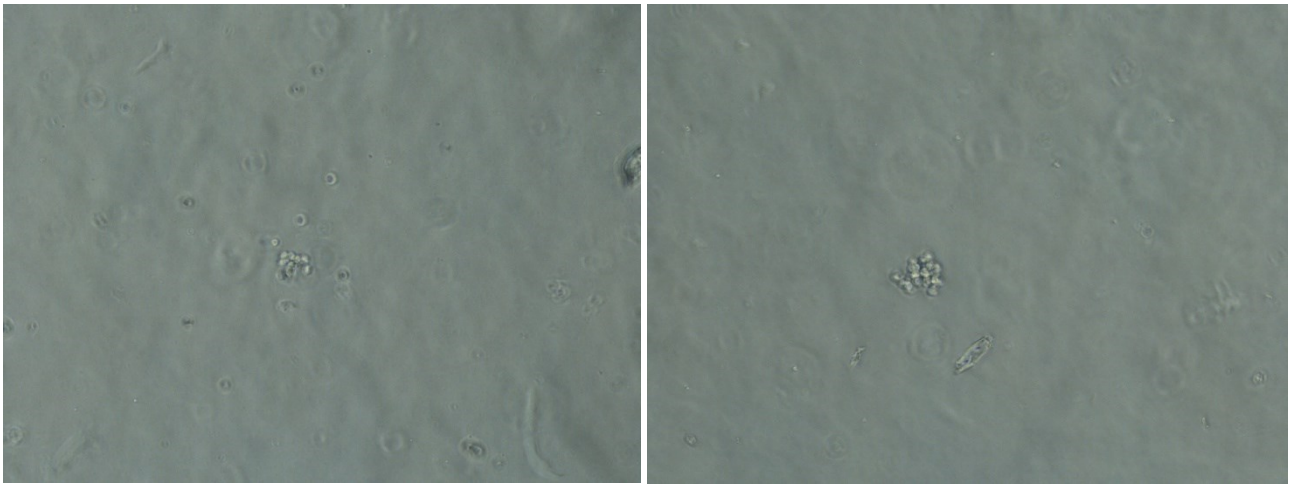


FIGURE S-8. Representative Small Colony Formation – **Left**: Image 0009; **Right**: Image 0012.

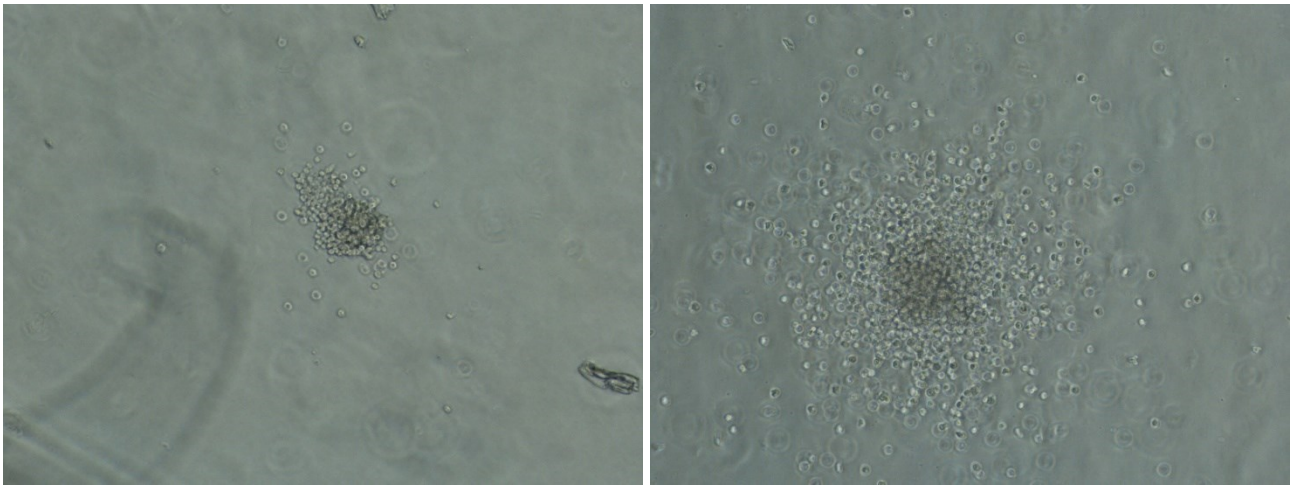


FIGURE S-9. Representative Large Colony Formation – **Left**: Image 0013; **Right**: Image 0007.

Part VIII. Supporting Information References.

References.

1. For Fmoc- based SPPS, see: Fields, G. B.; Noble, R. L. *Int. J. Pept. Prot. Res.* **1990**, *35*, 161−214.
2. For microwave-assisted SPPS and protocols, see: Palasek, S. A.; Cox, Z. J.; Collins, J. M. *J. Pept. Sci.* **2007**, *13*, 143–148 and references therein.
3. For the use of Oxyma Pure as a pH modulatory reagent, see: Subirós-Funosas, R.; El-Faham, A.; Albericio, F. *Biopolymers* **2012**, *98*, 89−97.
4. For the use of pseudoproline dipeptides, see: Wöhr, T.; Wahl, F.; Nefzi, A.; Rohwedder, B.; Sato, T.; Sun, X.; Mutter, M. *J. Am. Chem. Soc.* **1996**, *118*, 9218–9227.
5. (a) Sakakibara, S. *Biopolymers* **1995**, *37*, 17−28.; (b) Stuhr-Hansen, N.; Wilbekand, T.S.; Strømgaard, K. *Eur. J. Org. Chem.* **2013**, *24*, 5290−5294.
6. Edelhoch, H. *Biochemistry* 1967, 6, 1948–1954.
7. Moffatt, F.; Senkans, P.; Ricketts, D. *J. Chromatogr., A* **2000**, 891, 235–242.
8. Lu, H. S.; Clogston, C. L.; Narhi, L. O.; Merewether, L. A.; Pearl, W. R.; Boone, T. C. *J. Biol. Chem.* **1992**, *267*, 8770−8777.
9. Bradford, M. M. *Anal. Biochem.* **1976**, 72, 248−254.
10. Wang, P.; Aussedat, B.; Vohra, Y.; Danishefsky, S. J. *Angew. Chem. Int. Ed.* **2012**, *51*, 11571−11575 and references cited therein.
11. Hackenberger, C. P. R.; O'Reilly, M. K.; Imperiali, B. *J. Org. Chem.* **2005**, *70*, 3574−3578.
12. Veronese, F. M.; Mero, A.; Caboi, F.; Sergi, M.; Marongiu, C.; Pasut, G. *Bioconjugate Chem.* **2007**, 18, 1824−1830.
13. Natalello, A.; Ami, D.; Collini, M.; D'Alfonso, L.; Chirico, G.; Tonon, G.; Scaramuzza, S.; Schrepfer, R.; Doglia, S. M. *PLoS ONE* **2012**, *7*, e42511.
14. Mo, J.; Tymiak, A. A.; Chen, G. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 940–946.
15. Shirafuji, N.; Asano, S.; Matsuda, S.; Watari, K.; Takaku, F.; Nagata, S. *Exp Hematol.* **1989**, *17*, 116−119.
16. Wang, P.; Dong, S.; Shieh, J.-H.; Peguero, E.; Hendrickson, R.; Moore, M. A. S.; Danishefsky, S. J. *Science* **2013**, 342, 1357–1360.

Compound (**4**).

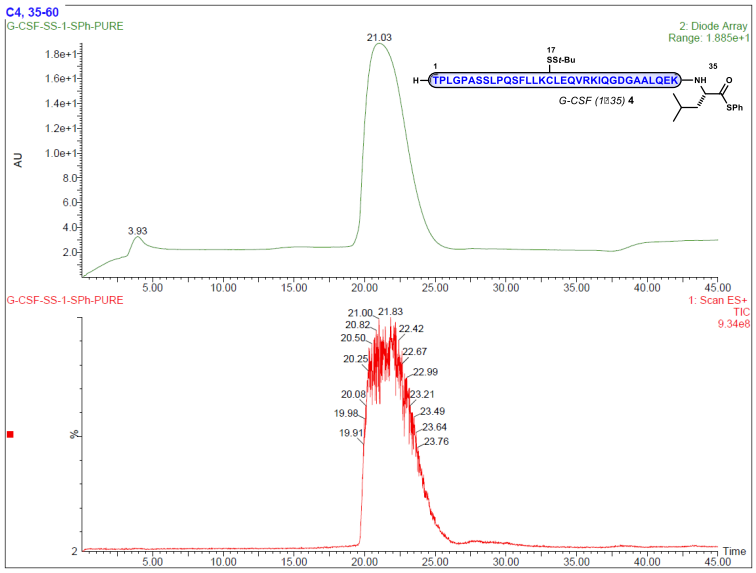


Figure 1.1A Top: UV Trace from LC-MS analysis of purified peptide **4**; gradient of 35% to 60% solvent B over 30 min at a flow rate of 0.2 mL/min, Agilent Microsorb C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**4**).

Figure 1.1B ESI-MS (+) of peptide **4**. M = $C_{176}H_{291}N_{45}O_{49}S_3$. ESI calculated for $[M+2H^+]^{2+}$ *m/z*: 1959.86, found: 1959.42; $[M+3H^+]^{3+}$ *m/z*: 1306.91, found: 1306.40; $[M+4H^+]^{4+}$ *m/z*: 980.44, found: 979.96.

Compound (**5**).

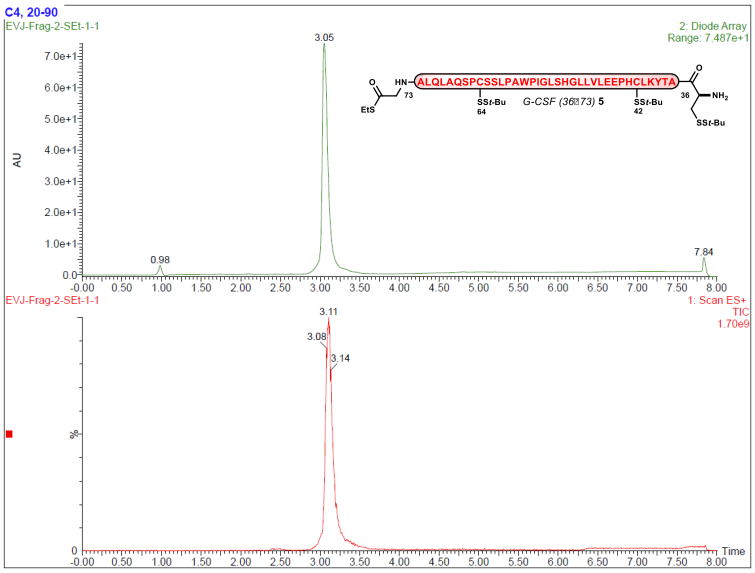


Figure 1.2A Top: UV Trace from UPLC-MS analysis of purified peptide **5**; gradient of 20% to 90% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**5**).

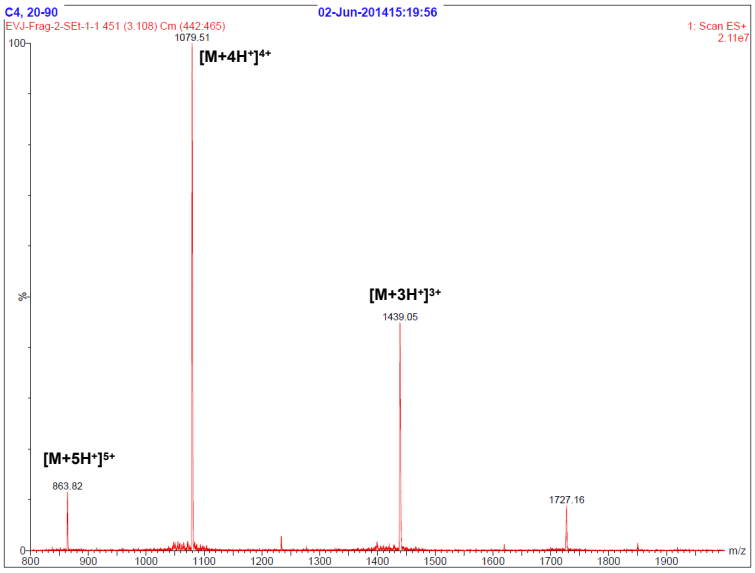


Figure 1.2B ESI-MS (+) of peptide **5**. M = $C_{194}H_{312}N_{46}O_{50}S_7$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1438.78, found: 1439.05; $[M+4H^+]^{4+}$ *m/z*: 1079.34, found: 1079.51; $[M+5H^+]^{5+}$ *m/z*: 863.67, found: 863.82.

Compound (**6**).

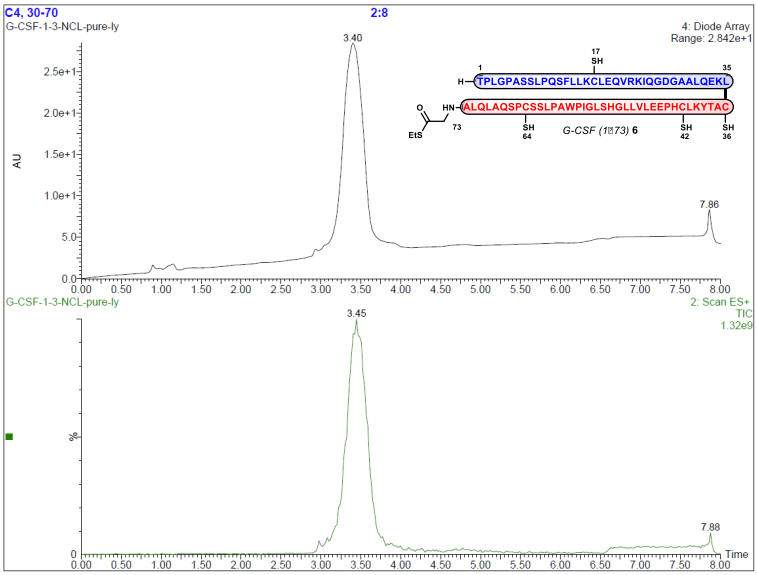


Figure 1.3A Top: UV Trace from UPLC-MS analysis of purified peptide **6**; gradient of 30% to 70% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**6**).

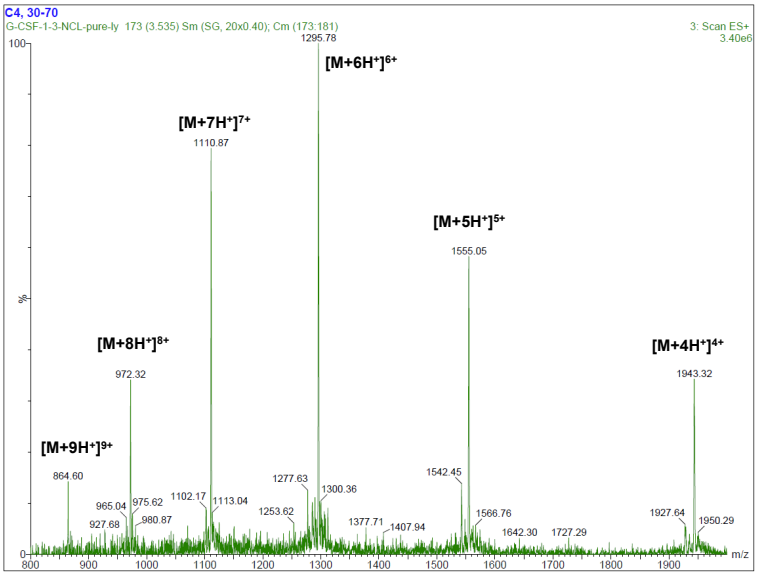


Figure 1.3B ESI-MS (+) of peptide **6**. M = $C_{348}H_{565}N_{91}O_{99}S_5$. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1943.06, found: 1943.32; $[M+5H^+]^{5+}$ *m/z*: 1554.65, found: 1555.05; $[M+6H^+]^{6+}$ *m/z*: 1295.71, found: 1295.78; $[M+7H^+]^{7+}$ *m/z*: 1110.75, found: 1110.87; $[M+8H^+]^{8+}$ *m/z*: 972.03, found: 972.32; $[M+9H^+]^{9+}$ *m/z*: 864.14, found: 864.60.

Compound (**6**). (macrocyle)

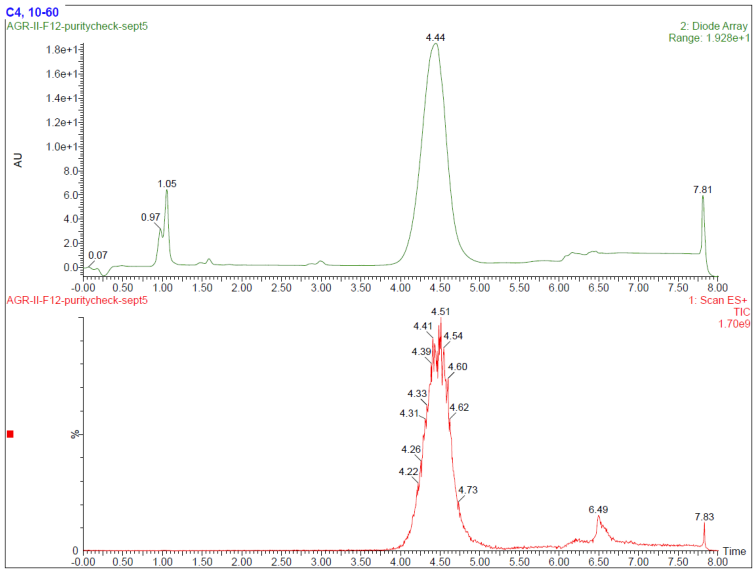


Figure 1.4A Top: UV Trace from UPLC-MS analysis of purified peptide **6** (with macrocycle); gradient of 10% to 60% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**6**). (macrocycle)

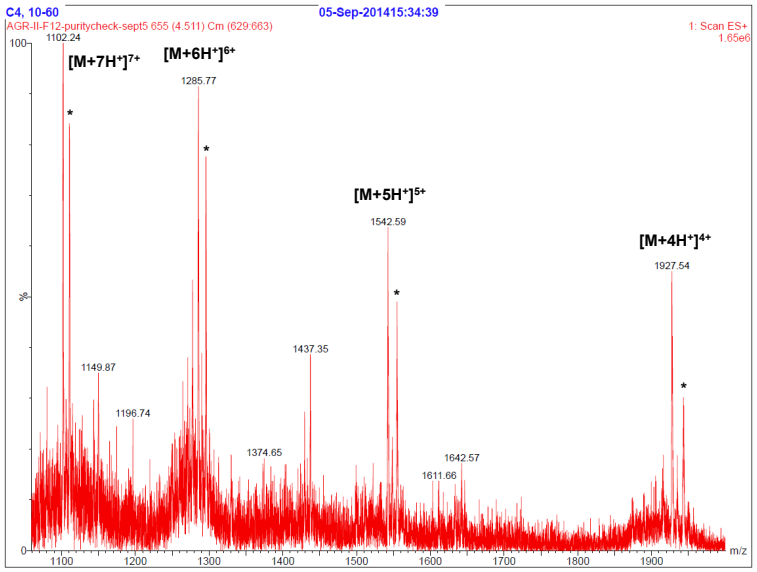


Figure 1.4B ESI-MS (+) of peptide **6** (macrocycle). M = $C_{346}H_{559}N_{91}O_{99}S_4$. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1927.42, found: 1927.54; $[M+5H^+]^{5+}$ *m/z*: 1542.22, found: 1542.59; $[M+6H^+]^{6+}$ *m/z*: 1285.35, found: 1285.77; $[M+7H^+]^{7+}$ *m/z*: 1101.87, found: 1102.24. *indicates presence of **6** in sample spectrum.

Compound (**7**).

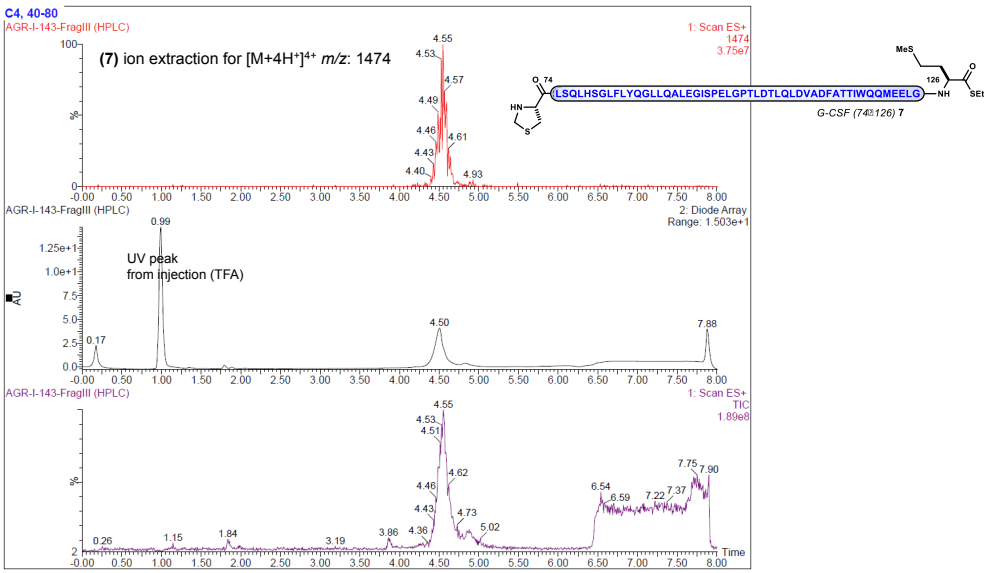


Figure 1.5A Top: ESI (+), ion extraction for $[M+4H^+]^{4+}$ *m/z*: 1474. **Middle:** UV Trace from UPLC-MS analysis of purified peptide **7**; gradient of 40% to 80% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**7**).

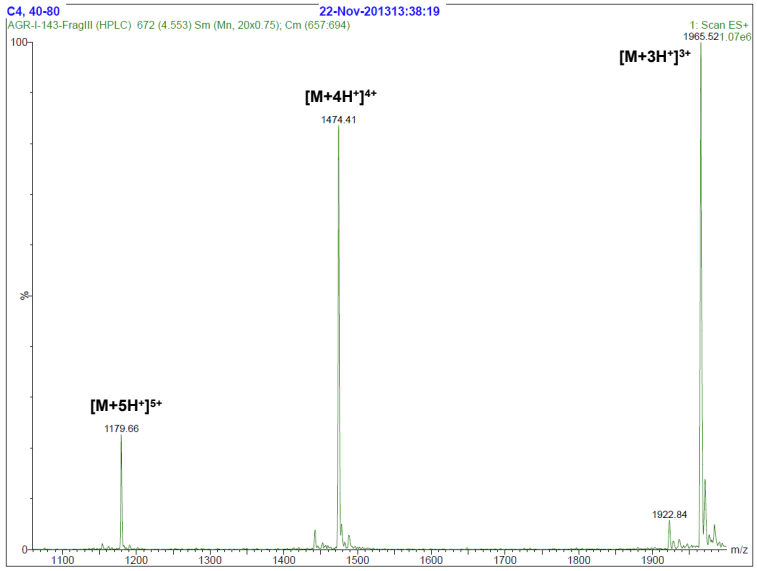


Figure 1.5B ESI-MS (+) of peptide **7**. M = $C_{265}H_{414}N_{62}O_{81}S_4$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1965.28, found: 1965.52; $[M+4H^+]^{4+}$ *m/z*: 1474.21, found: 1474.41; $[M+5H^+]^{5+}$ *m/z*: 1179.57, found: 1179.66.

Compound (**8**).

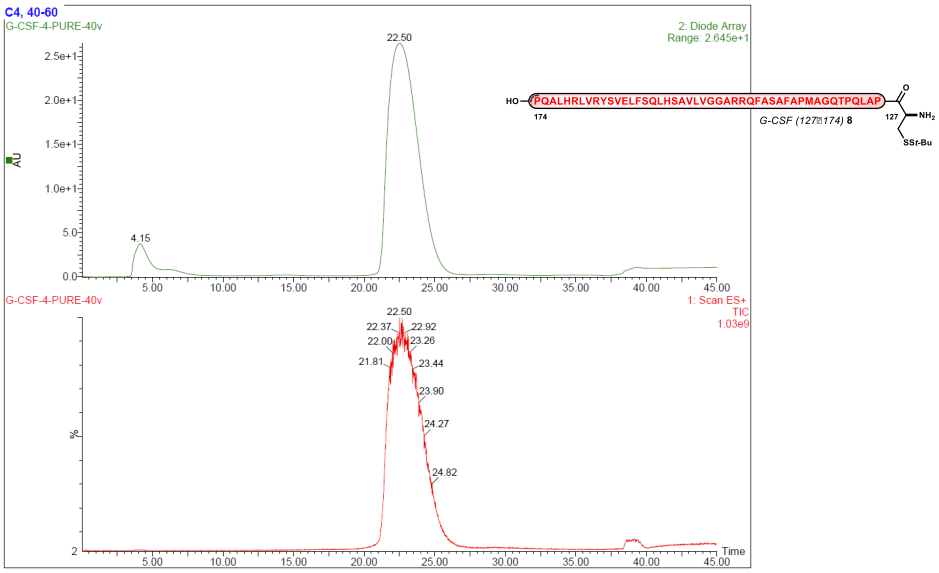


Figure 1.6A Top: UV Trace from LC-MS analysis of purified peptide **8**; gradient of 40% to 60% solvent B over 30 min at a flow rate of 0.2 mL/min, Agilent Microsorb C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**8**).

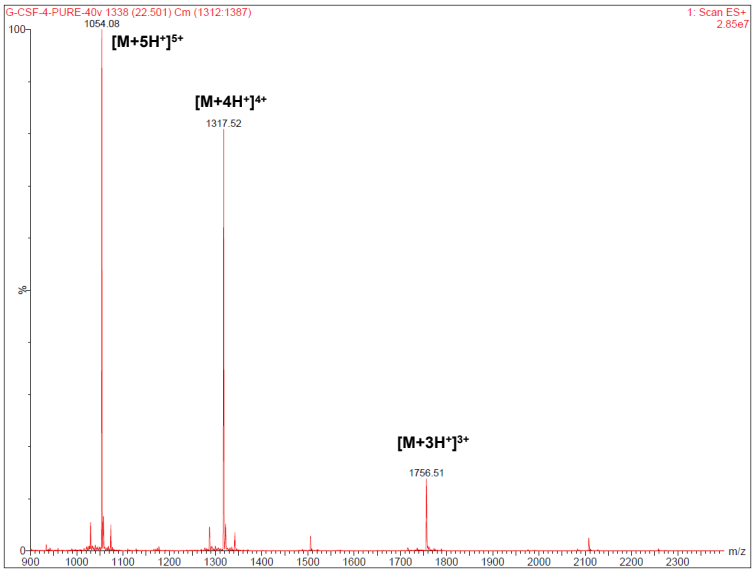


Figure 1.6B ESI-MS (+) of peptide **8**. M = $C_{236}H_{375}N_{69}O_{62}S_3$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1756.74, found: 1756.51; $[M+4H^+]^{4+}$ *m/z*: 1317.81, found: 1317.52; $[M+5H^+]^{5+}$ *m/z*: 1054.45, found: 1054.08.

Compound (**9a**). (crude reaction)

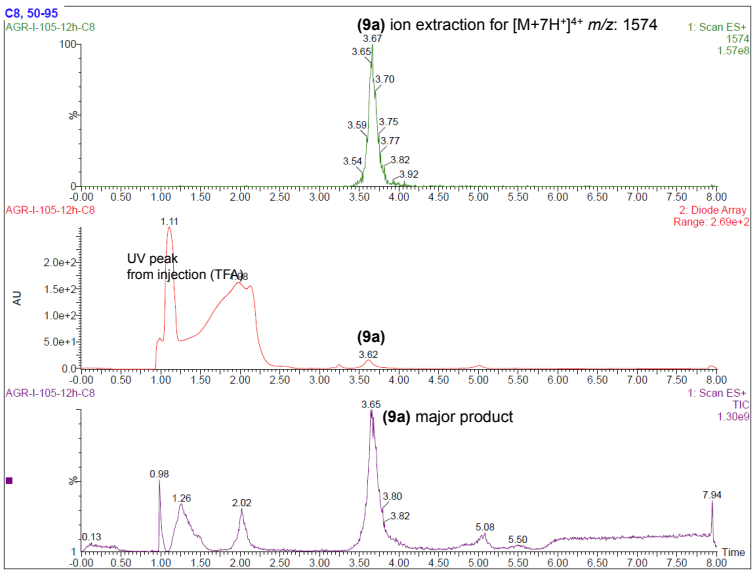


Figure 1.7A Top: ESI (+), ion extraction for **9a**, $[M+7H^+]^{7+}$ *m/z*: 1574 **Middle:** UV Trace from UPLC-MS analysis of crude reaction to prepare **9a**; gradient of 50% to 95% solvent B over 6 min at a flow rate of 0.3 mL/ min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**9a**). (crude reaction)

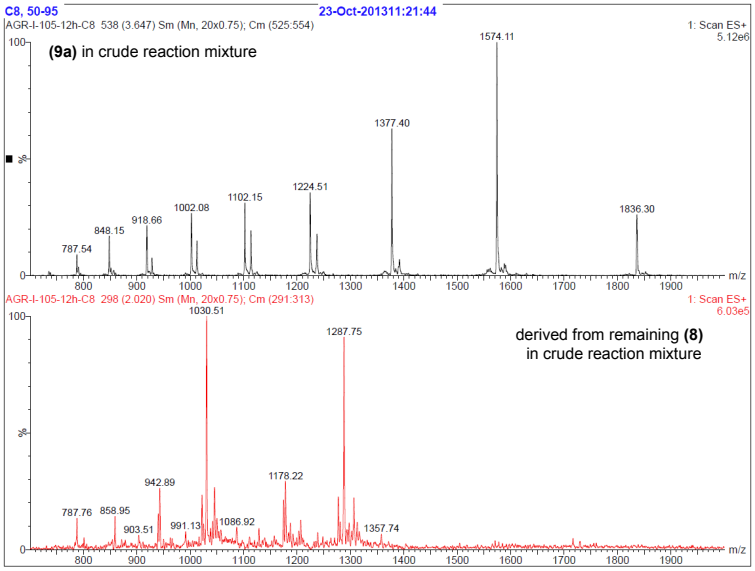


Figure 1.7B Top: ESI-MS (+) for crude product **9a**. **Bottom:** ESI-MS (+) for peptide derived from excess peptide **8**.

Compound (**9a**).

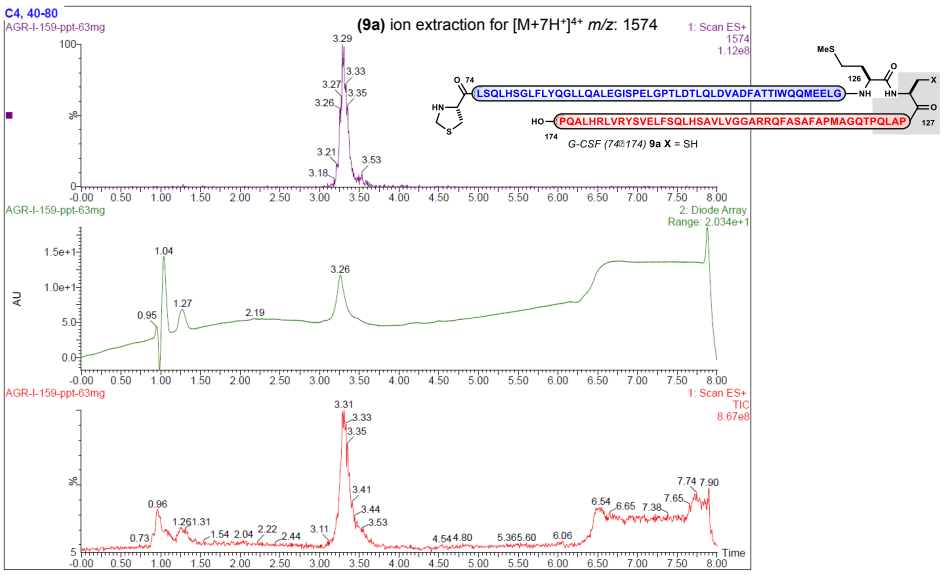


Figure 1.8A Top: ESI (+), ion extraction for **9a**, $[M+7H^+]^{7+}$ *m/z*: 1574 **Middle:** UV Trace from UPLC-MS analysis of purified peptide **9a** (precipitation); gradient of 40% to 80% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**9a**).

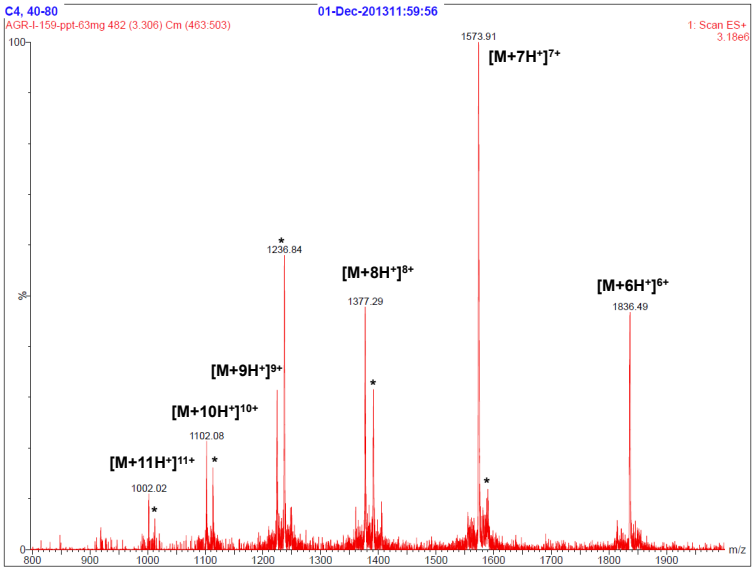


Figure 1.8B ESI-MS (+) of peptide **9a**. M = $C_{495}H_{775}N_{131}O_{143}S_5$. ESI calculated for $[M+6H^+]^{6+}$ *m/z*: 1835.96, found: 1836.49; $[M+7H^+]^{7+}$ *m/z*: 1573.83, found: 1573.91; $[M+8H^+]^{3+}$ *m/z*: 1377.22, found: 1377.29; $[M+9H^+]^{9+}$ *m/z*: 1224.31, found: 1224.74; $[M+10H^+]^{10+}$ *m/z*: 1101.98, found: 1102.08; $[M+11H^+]^{11+}$ *m/z*: 1001.89, found: 1002.02. * indicates the presence of a TFA adduct.

Compound (**9abc**)
(crude reaction)

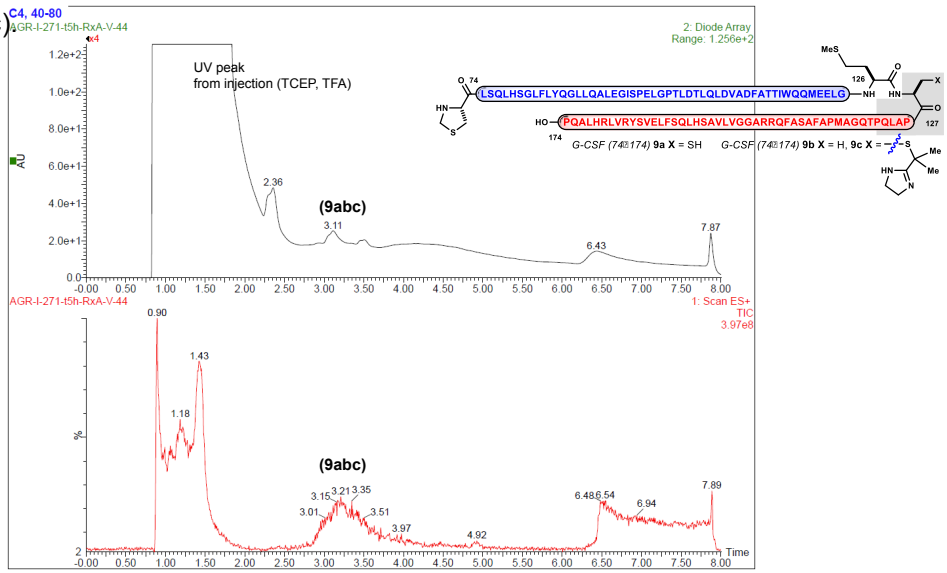


Figure 1.9A Top: UV Trace from UPLC-MS analysis of crude reaction to convert **9a** into **9b**; gradient of 40% to 80% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**9abc**)
(crude reaction)

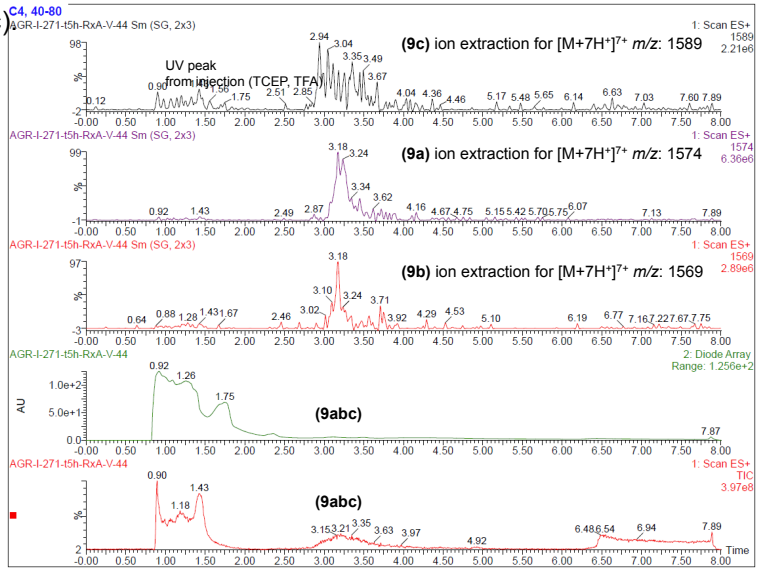


Figure 1.9B Top: ESI (+), ion extraction for **9c**, $[M+7H^+]^{7+}$ *m/z*: 1589 **Middle (1):** ESI (+), ion extraction for **9a**, $[M+7H^+]^{7+}$ *m/z*: 1574 **Middle (2):** ESI (+), ion extraction for **9b**, $[M+7H^+]^{7+}$ *m/z*: 1569 **Middle (3):** UV Trace from UPLC-MS analysis of crude reaction to convert **9a** into **9b**; gradient of 40% to 80% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**9abc**)
(crude reaction)

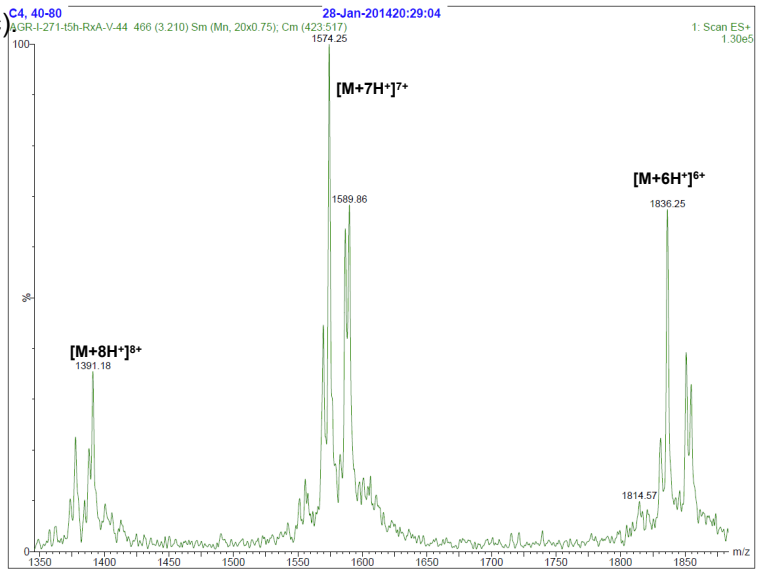


Figure 1.9C ESI-MS (+) of peptide **9abc**. Extracted mass from t = 2.8 to 3.8 min.
9a, ESI calculated for $[M+8H^+]^{8+}$ *m/z*: 1377.22, found: 1377.75, $[M+7H^+]^{7+}$ *m/z*: 1573.83, found: 1574.25; $[M+6H^+]^{6+}$ *m/z*: 1835.96, found: 1836.25.
9b, ESI calculated for $[M+8H^+]^{8+}$ *m/z*: 1373.22, found: 1373.62, $[M+7H^+]^{7+}$ *m/z*: 1569.25, found: 1569.68; $[M+6H^+]^{6+}$ *m/z*: 1830.62, found: 1831.00.
9c, ESI calculated for $[M+8H^+]^{8+}$ *m/z*: 1390.99, found: 1391.18, $[M+7H^+]^{7+}$ *m/z*: 1589.56, found: 1589.86; $[M+6H^+]^{6+}$ *m/z*: 1854.32, found: 1854.63.

Compound (**9abc**)
(crude reaction)

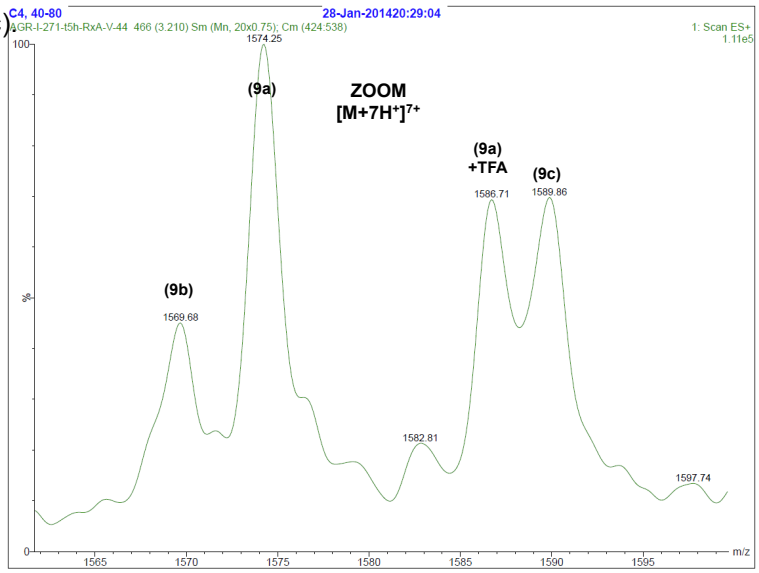


Figure 1.9D ESI-MS (+) of peptide **9abc**. Extracted mass from t = 2.8 to 3.8 min.
9a, ESI calculated for $[M+7H^+]^{7+}$ *m/z*: 1573.83, found: 1574.25.
9b, ESI calculated for $[M+7H^+]^{7+}$ *m/z*: 1569.25, found: 1569.68.
9c, ESI calculated for $[M+7H^+]^{7+}$ *m/z*: 1589.56, found: 1589.86.

Compound (**12**). (LC-MS analysis)

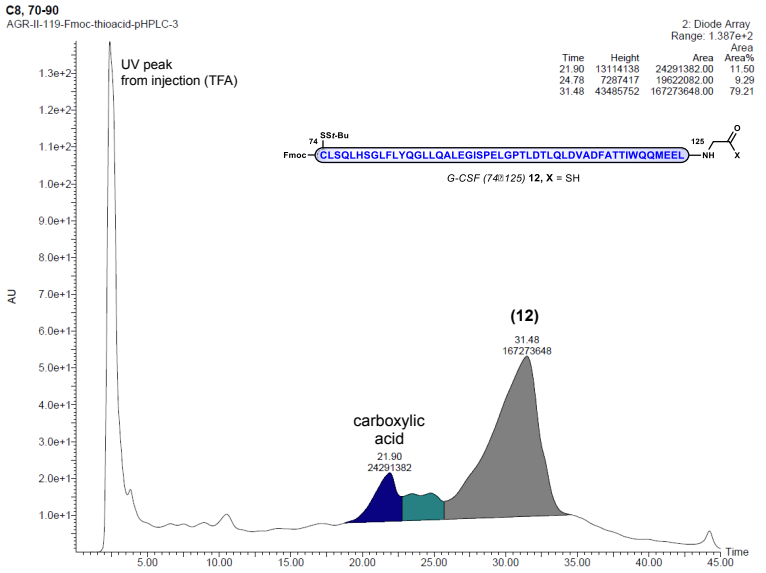


Figure 1.10A Top: UV Trace from LC-MS analysis of HPLC purified peptide **12**; gradient of 70% to 90% solvent B over 30 min at a flow rate of 0.2 mL/min, X-Bridge C8 analytical column.

Compound (**12**). (LC-MS analysis)

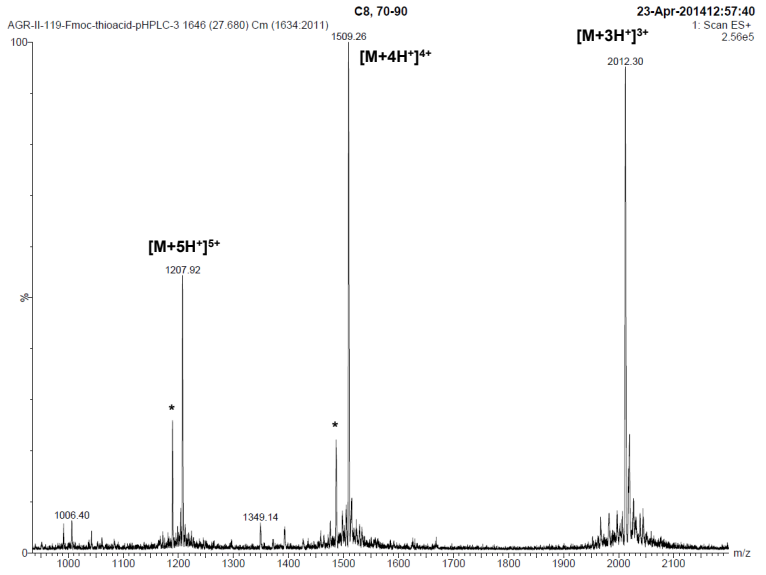


Figure 1.10B ESI-MS (+) of peptide **12**. M = $C_{276}H_{419}N_{61}O_{82}S_4$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 2011.67, found: 2012.30; $[M+4H^+]^{4+}$ *m/z*: 1509.00, found: 1509.26; $[M+5H^+]^{5+}$ *m/z*: 1207.40, found: 1207.92. * indicates presence of corresponding carboxylic acid in spectrum.

Compound (**12**) (UPLC-MS analysis)

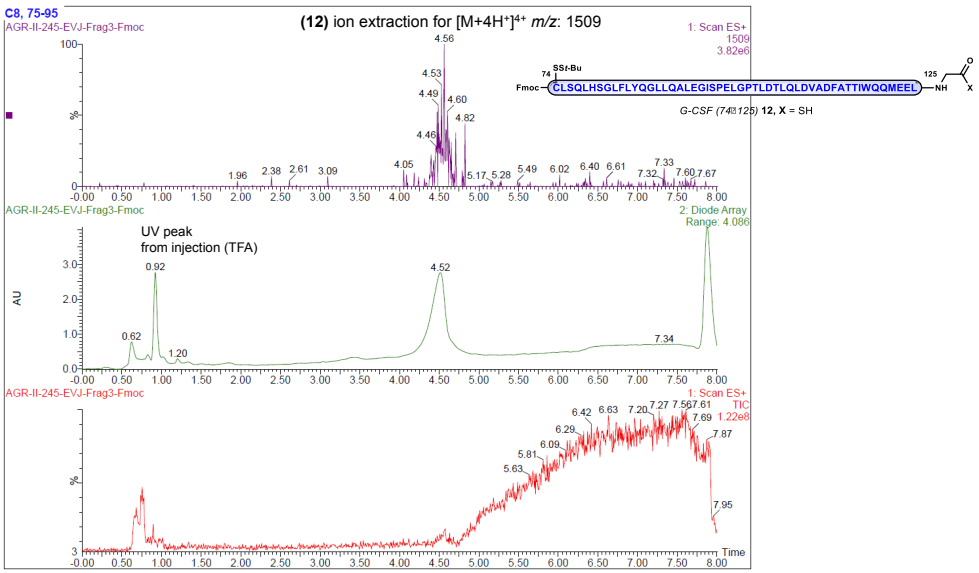


Figure 1.10C Top: ESI (+), ion extraction for $[M+4H^+]^{4+}$ *m/z*: 1509. **Middle:** UV Trace from UPLC-MS analysis of HPLC purified peptide **12**; gradient of 75% to 95% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**12**) (UPLC-MS analysis)

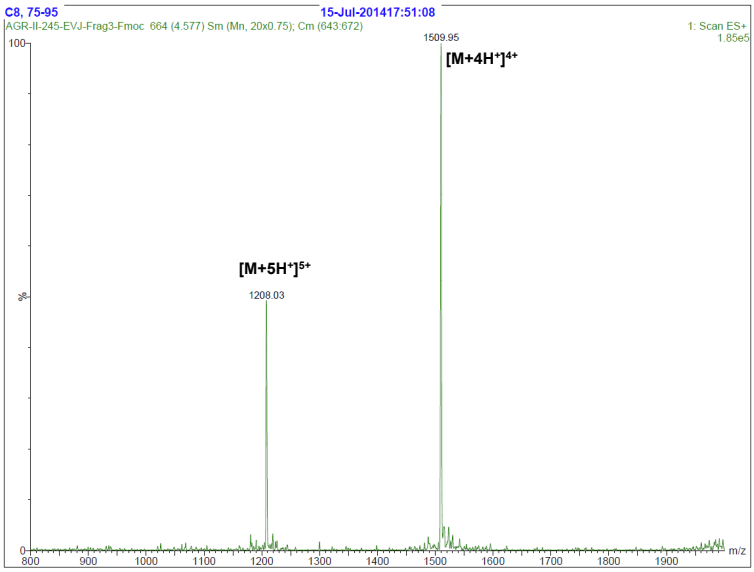


Figure 1.10D ESI-MS (+) of peptide **12**. M = $C_{276}H_{419}N_{61}O_{82}S_4$. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1509.00, found: 1509.95; $[M+5H^+]^{5+}$ *m/z*: 1207.40, found: 1208.03.

Compound (**12**).

G-CSF (74‖125) **12, X** = SH

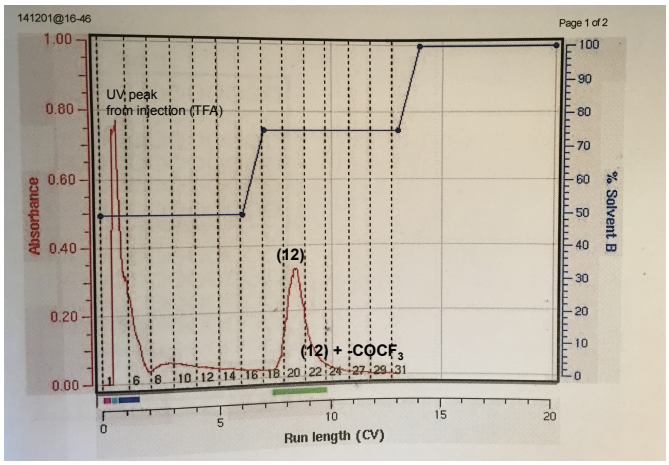


Figure 1.11A UV-trace from flash chromatography purification of **12** with C_2 silica gel. Flow rate of 12 mL/min, λ = 220 nm, 8 mL fractions collected.

Compound (**12**)

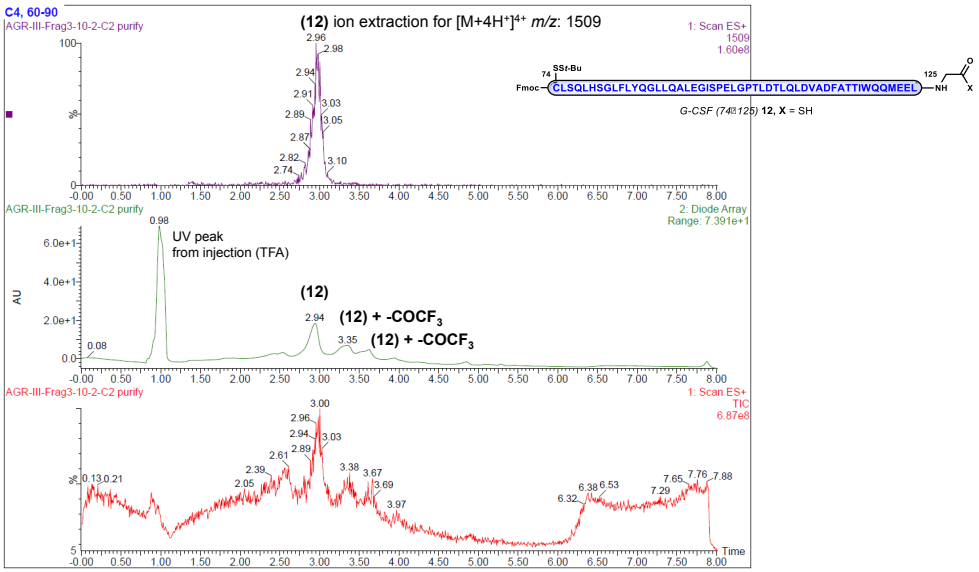


Figure 1.11B Top: ESI (+), ion extraction for $[M+4H^+]^{4+}$ *m/z*: 1509. **Middle:** UV Trace from UPLC-MS analysis of C_2 silica gel purified peptide **12**; gradient of 60% to 90% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**12**)

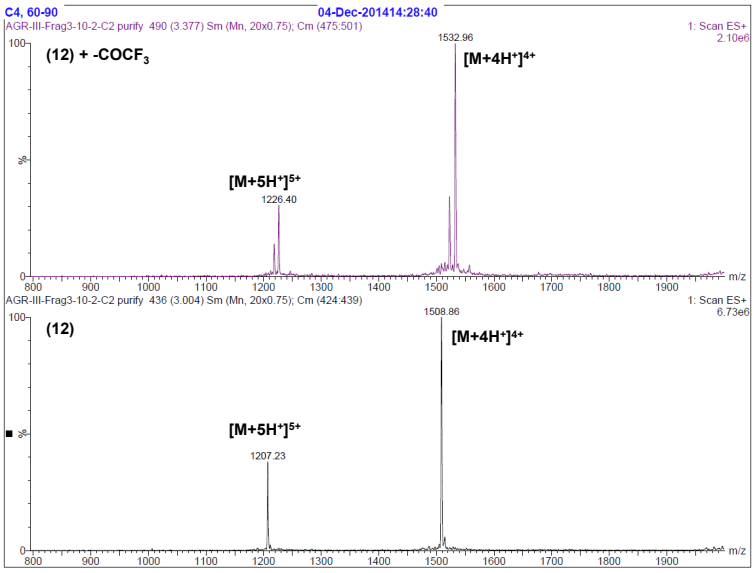


Figure 1.11C Top: ESI-MS (+) of peptide **12 + -COCF$_3$**. M = $C_{278}H_{418}F_3N_{61}O_{83}S_4$. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1533.00, found: 1532.96; $[M+5H^+]^{5+}$ *m/z*: 1226.60, found: 1226.40. **Bottom:** ESI-MS (+) of peptide **12**. M = $C_{276}H_{419}N_{61}O_{82}S_4$. ESI calculated for $[M+4H^+]^{4+}$ *m/z*: 1509.00, found: 1508.86; $[M+5H^+]^{5+}$ *m/z*: 1207.40, found: 1207.23.

Compound (**13**)

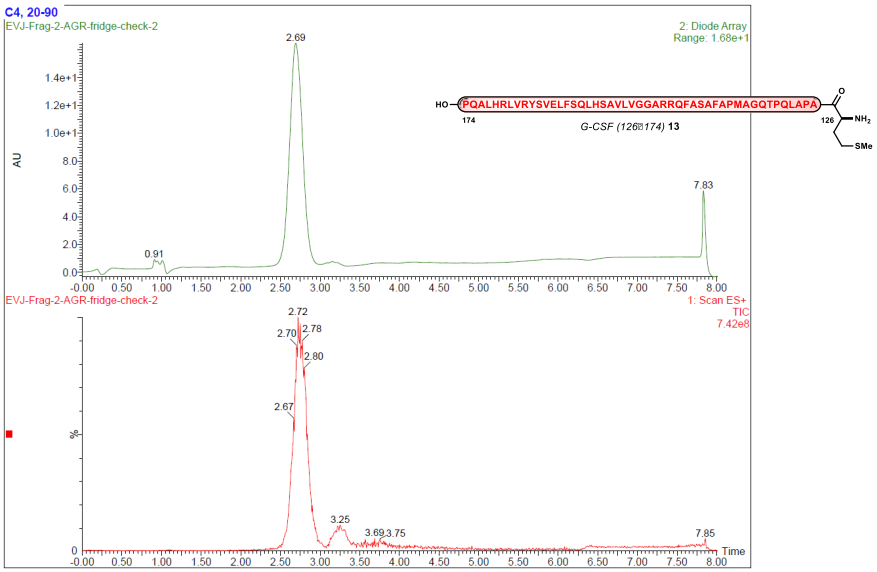


Figure 1.12A Top: UV Trace from UPLC-MS analysis of purified peptide **13**; gradient of 20% to 90% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**13**).

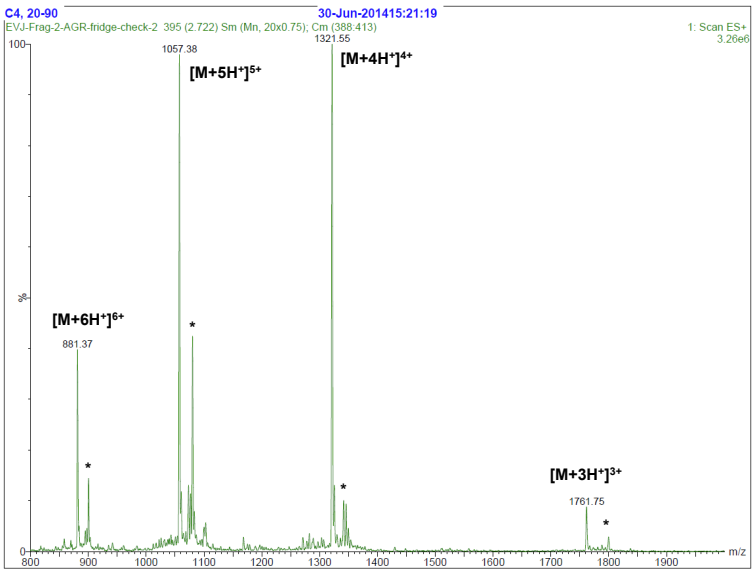


Figure 1.12B ESI-MS (+) of peptide **13**. M = $C_{237}H_{376}N_{70}O_{63}S_2$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1760.39, found: 1761.75; $[M+4H^+]^{4+}$ *m/z*: 1320.55, found: 1321.55; $[M+5H^+]^{5+}$ *m/z*: 1056.64, found: 1057.38; $[M+6H^+]^{6+}$ *m/z*: 880.70, found: 881.37. * indicates presence of TFA adduct in spectrum.

Compound (**14**). (crude reaction)

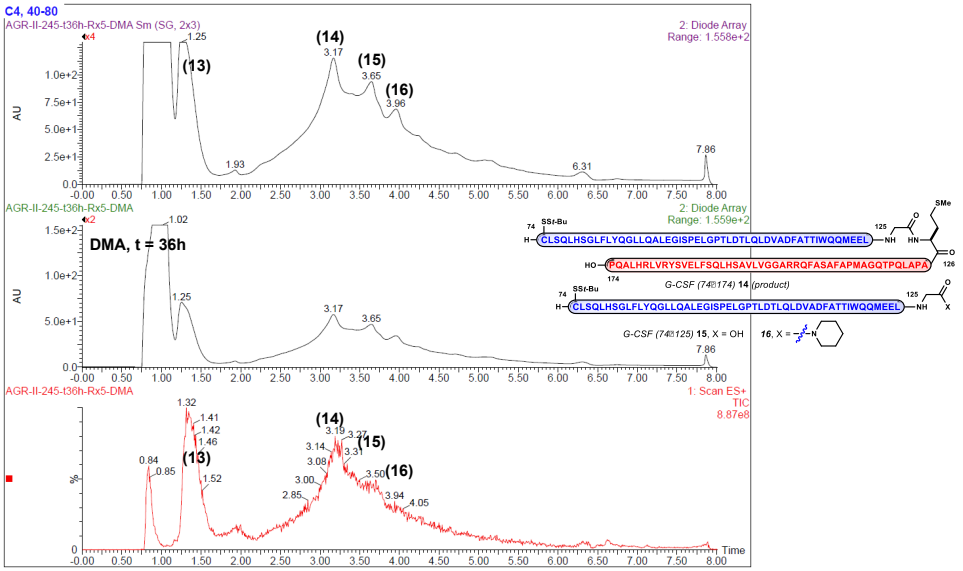


Figure 1.13A Top and Middle: Enlarged (4×) and (2×) UV Trace from UPLC-MS analysis of crude reaction in DMA to prepare peptide **14**. Following reaction aliquot, treatment with piperidine, precipitation with Et_2O and dissolution. Gradient of 40% to 80% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**14**)
(crude reaction)

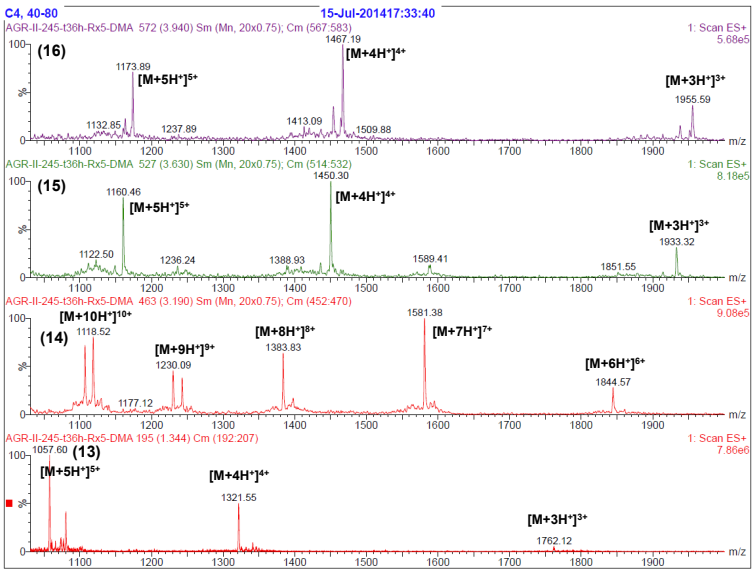


Figure 1.13B Top: Product identified as piperidide (**16**) at 3.94 min, found $[M+3H^+]^{3+}$, $[M+4H^+]^{5+}$, $[M+5H^+]^{5+}$. **Middle (1):** Product identified as carboxylic acid (**15**) at 3.63 min, found $[M+3H^+]^{3+}$, $[M+4H^+]^{5+}$, $[M+5H^+]^{5+}$. **Middle (2):** Product identified as desired peptide (**14**) at 3.19 min, found $[M+6H^+]^{6+}$, $[M+7H^+]^{7+}$, $[M+8H^+]^{8+}$ $[M+9H^+]^{9+}$, $[M+10H^+]^{10+}$. **Bottom:** Remaining (**13**) at 1.34 min, found $[M+3H^+]^{3+}$, $[M+4H^+]^{5+}$, $[M+5H^+]^{5+}$.

Compound (**14**).
(crude reactions)

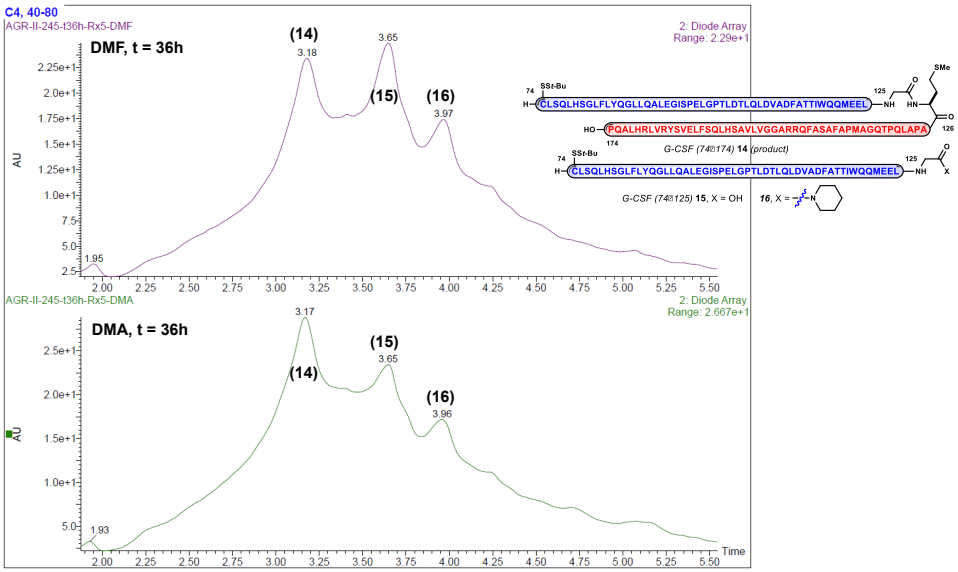


Figure 1.13C Top: UV Trace from UPLC-MS analysis of crude reaction in DMF to prepare peptide **14**. **Bottom:** UV Trace from UPLC-MS analysis of crude reaction in DMA to prepare peptide **14**. Following reaction aliquot, treatment with piperidine, precipitation with Et_2O and dissolution. Gradient of 40% to 80% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column.

Compound (**14**).
(crude reactions)
t = 20h, 1d

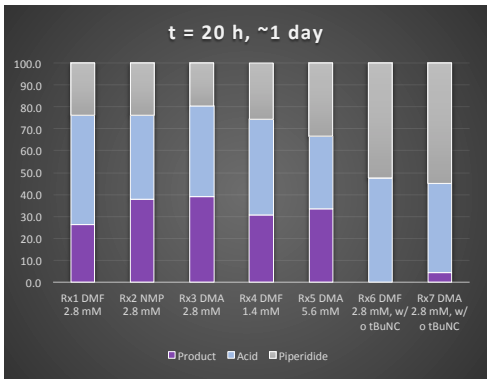


Figure 1.14A Graphical representation of relative ratio of product **14**, acid **15** and piperidide **16** at t = 20h.

Compound (**14**).
(crude reactions)
t = 44h, 2d

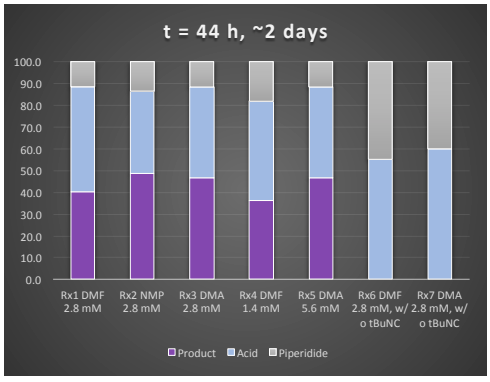


Figure 1.14B Graphical representation of relative ratio of product **14**, acid **15** and piperidide **16** at t = 44h.

Compound (**14**).
(crude reactions)
t = 72h, 3d

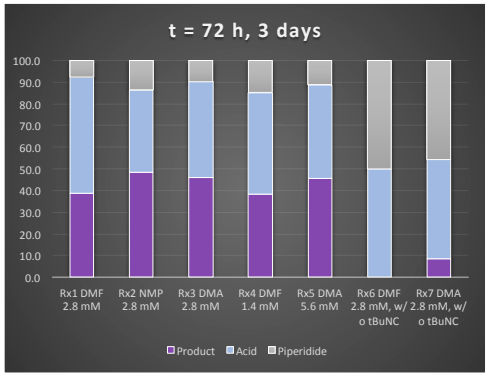


Figure 1.14C Graphical representation of relative ratio of product **14**, acid **15** and piperidide **16** at t = 72h.

Compound (**14**).
(purification)

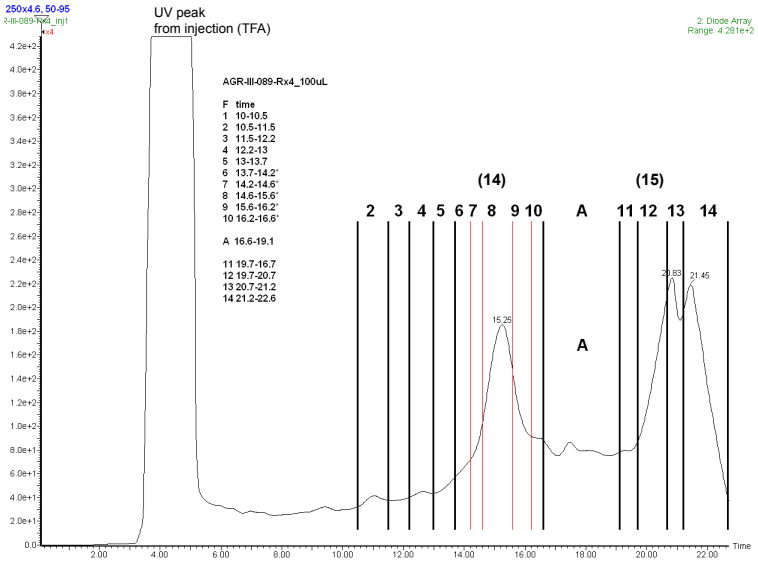


Figure 1.15 UV trace from HPLC purification of peptide **14** (range 0 to 22 min); gradient of 55% to 95% solvent B over 30 min at a flow rate of 0.8 mL/min, Higgins C4 semi-preparative column. Fractions 6 to 10 contain product **14** (t = 15.3 min) and were pooled for isolated yield. Fractions 11 and 12 contain carboxylic acid **15** (t = 20.8 min). Fraction A contains impure desired **14** and trifluoroacetate esters of **14**.

Compound (**14**)

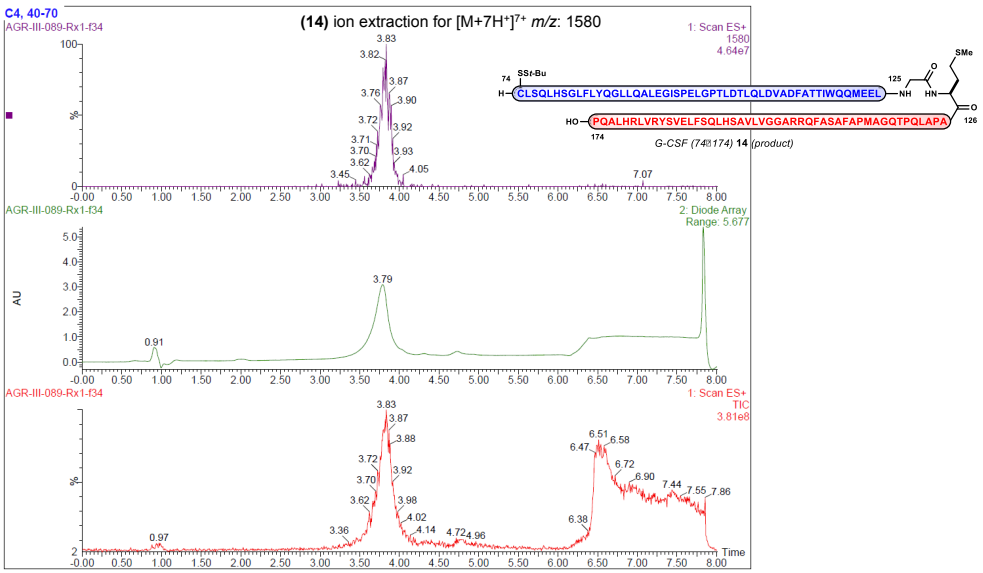


Figure 1.16A Top: ESI (+), ion extraction for $[M+7H^+]^{7+}$ *m/z*: 1580. **Middle:** UV Trace from UPLC-MS analysis of purified peptide **14**; gradient of 40% to 70% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**14**).

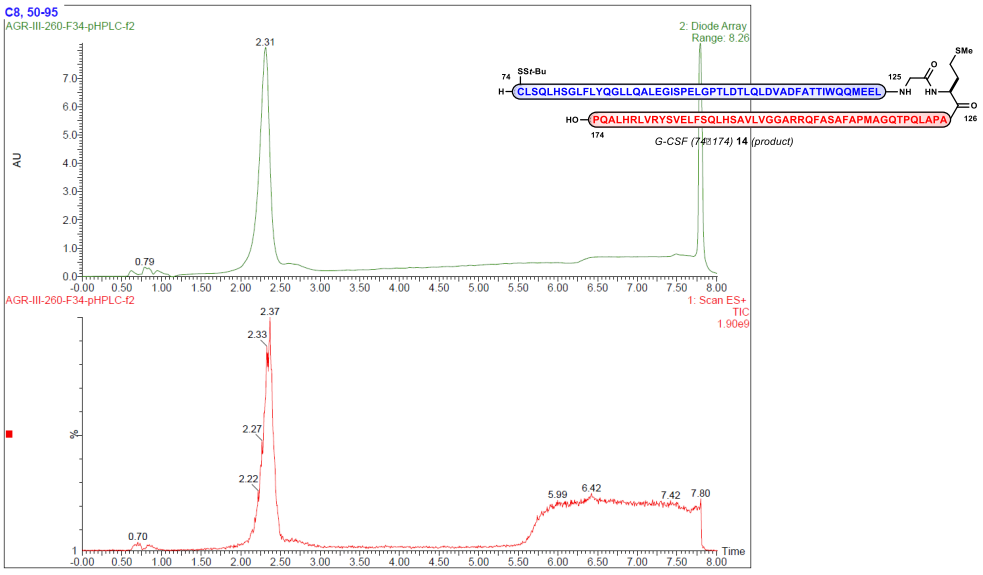


Figure 1.16B Top: UV Trace from UPLC-MS analysis of purified peptide **14**; gradient of 50% to 95% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**14**).

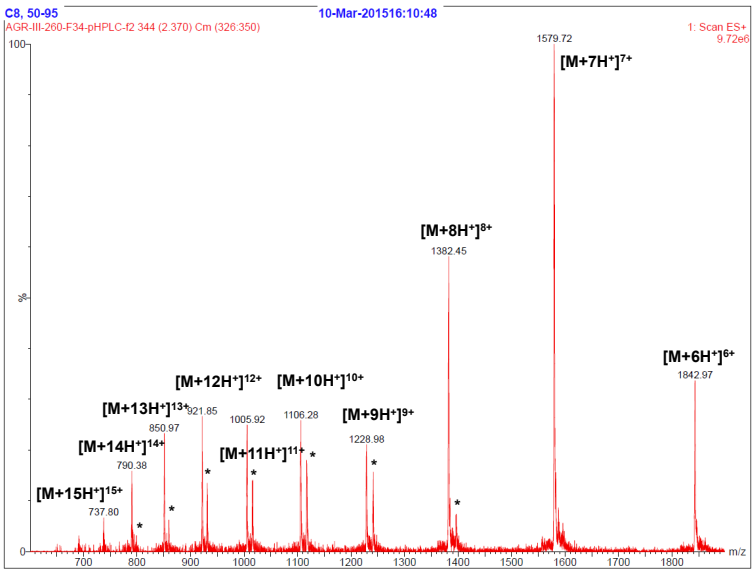


Figure 1.16C ESI-MS (+) of peptide **14**. M = $C_{498}H_{783}N_{131}O_{143}S_5$. ESI calculated for $[M+6H^+]^{6+}$ *m/z*: 1843.31, found: 1842.97; $[M+7H^+]^{7+}$ *m/z*: 1580.12, found: 1579.72; $[M+8H^+]^{8+}$ *m/z*: 1382.74, found: 1382.45; $[M+9H^+]^{9+}$ *m/z*: 1229.21, found: 1228.98; $[M+10H^+]^{10+}$ *m/z*: 1106.39, found: 1106.28; $[M+11H^+]^{11+}$ *m/z*: 1005.90, found: 1005.92; $[M+12H^+]^{12+}$ *m/z*: 922.16, found: 921.85; $[M+13H^+]^{13+}$ *m/z*: 851.30, found: 850.97; $[M+14H^+]^{14+}$ *m/z*: 790.57, found: 790.38; $[M+15H^+]^{15+}$ *m/z*: 737.93, found: 737.80. * indicates presence of a TFA adduct.

Compound (**15**).

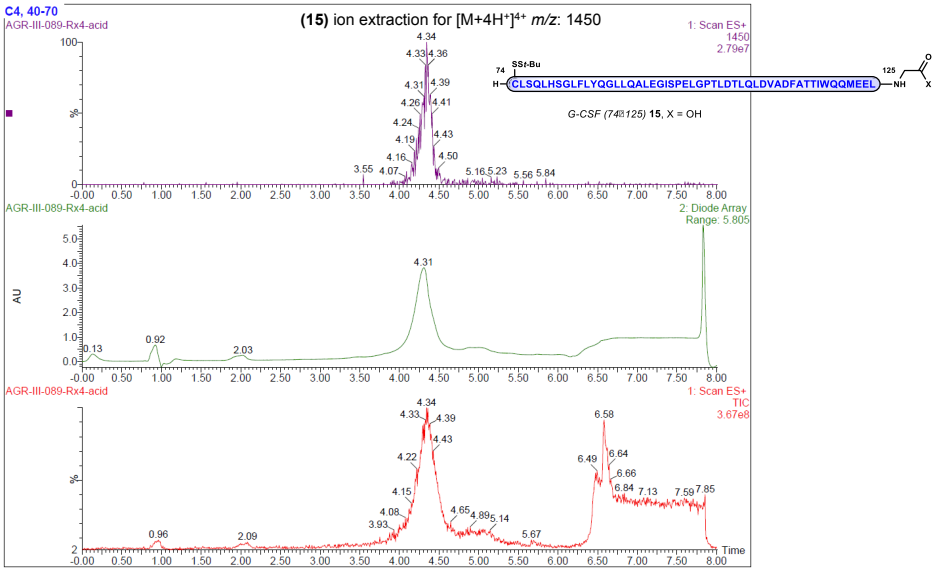


Figure 1.17A Top: ESI (+), ion extraction for $[M+4H^+]^{4+}$ *m/z*: 1450. **Middle:** UV Trace from UPLC-MS analysis of purified peptide **15**; gradient of 40% to 70% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**15**).

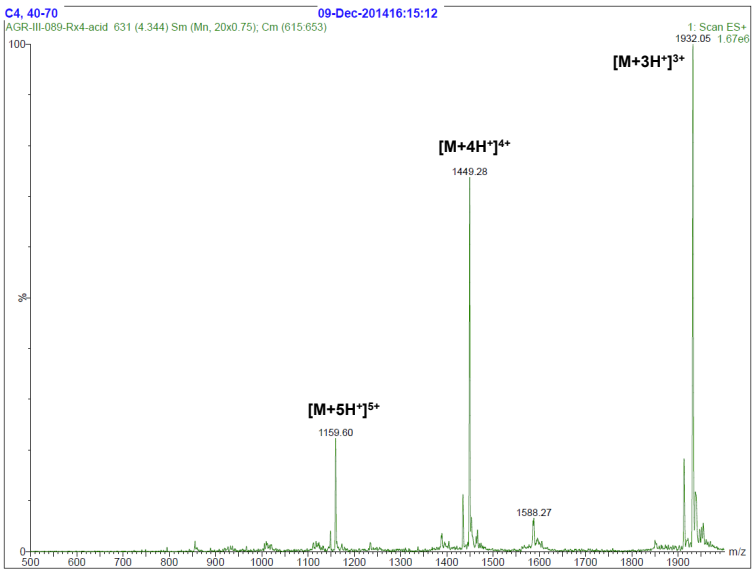


Figure 1.17B ESI-MS (+) of peptide **15**. M = $C_{261}H_{409}N_{61}O_{81}S_3$. ESI calculated for $[M+3H^+]^{5+}$ *m/z*: 1932.23, found: 1923.05; $[M+4H^+]^{5+}$ *m/z*: 1449.43, found: 1449.28; $[M+5H^+]^{5+}$ *m/z*: 1159.74, found: 1159.60.

Compound (**22**).

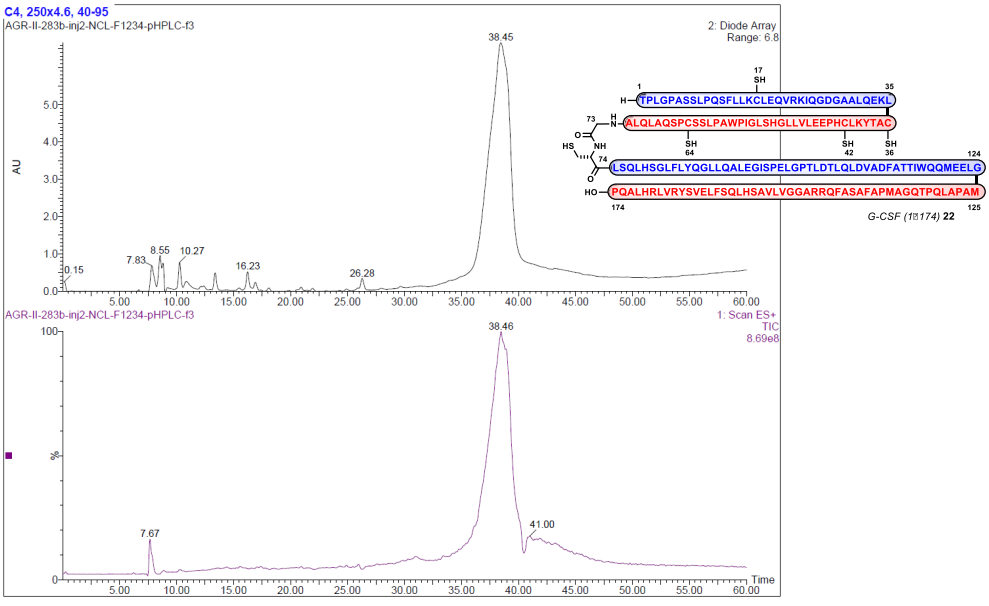


Figure 1.18A Top: UV Trace from LC-MS analysis of linear G-CSF **22**; gradient of 40% to 95% solvent B over 60 min at a flow rate of 0.2 mL/min, Higgins C4 semi-preparative column. **Bottom:** ESI-MS (+) TIC.

Compound (**22**).

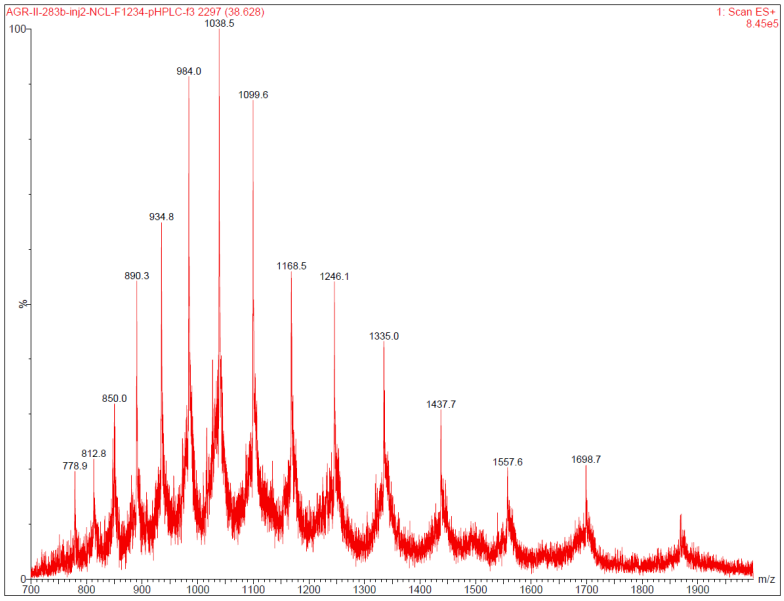


Figure 1.18B ESI-MS (+) of linear G-CSF **22**. M = $C_{840}H_{1334}N_{222}O_{242}S_8$. ESI calculated for $[M+11H^+]^{11+}$ *m/z*: 1698.4, found: 1698.7; $[M+12H^+]^{12+}$ *m/z*: 1557.0, found: 1557.6; $[M+13H^+]^{13+}$ *m/z*: 1437.3, found: 1437.7; $[M+14H^+]^{14+}$ *m/z*: 1334.7, found: 1335.0; $[M+15H^+]^{15+}$ *m/z*: 1245.8, found: 1246.1; $[M+16H^+]^{16+}$ *m/z*: 1168.0, found: 1168.5; $[M+17H^+]^{17+}$ *m/z*: 1099.3, found: 1099.6; $[M+18H^+]^{18+}$ *m/z*: 1038.3, found: 1038.5; $[M+19H^+]^{19+}$ *m/z*: 983.7, found: 984.0; $[M+20H^+]^{20+}$ *m/z*: 934.6, found: 934.8 ; $[M+21H^+]^{21+}$ *m/z*: 890.1, found: 890.3; $[M+22H^+]^{22+}$ *m/z*: 849.7, found: 850.0; $[M+23H^+]^{23+}$ *m/z*: 812.8, found: 812.8; $[M+24H^+]^{24+}$ *m/z*: 779.0, found: 778.9.

Compound (**2a**). (crude in folding buffer) UPLC-MS

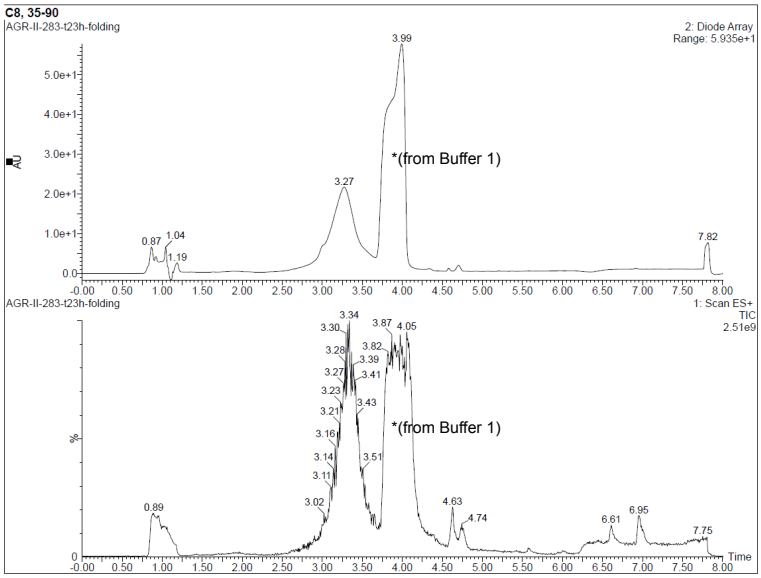


Figure 1.19A Top: UV Trace from UPLC-MS analysis of crude G-CSF **2a** in folding buffer at t = 23h; gradient of 35% to 90% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C8 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**2a**) (crude in folding buffer) UPLC-MS

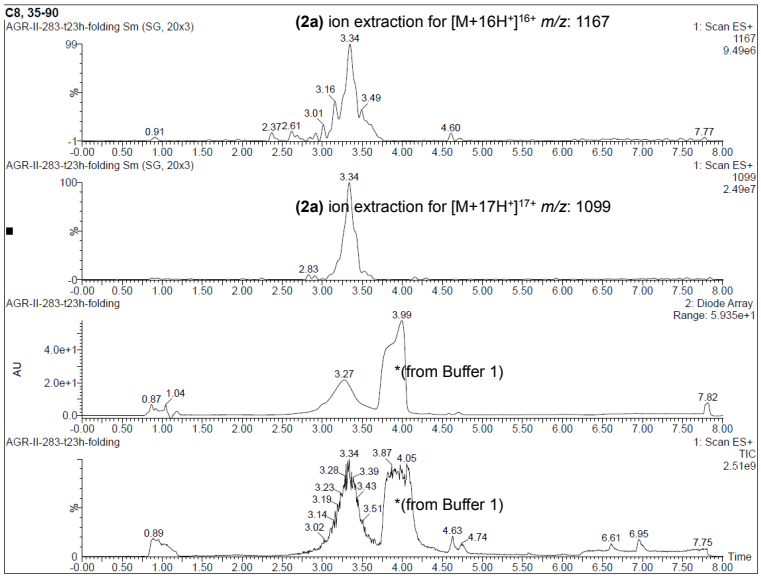


Figure 1.19B Top: ESI-MS (+) TIC, ion extraction for $[M+16H^+]^{16+}$ *m/z*: 1167. **Middle (1):** ESI-MS (+) TIC, ion extraction for $[M+17H^+]^{17+}$ *m/z*: 1099. **Middle (2):** UV Trace from UPLC-MS analysis of crude G-CSF **2a** in folding buffer at t = 23h; gradient of 35% to 90% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C8 analytical column. **Bottom:** ESI-MS (+) TIC.

Comparison (**22**) and crude (**2a**). UPLC-MS

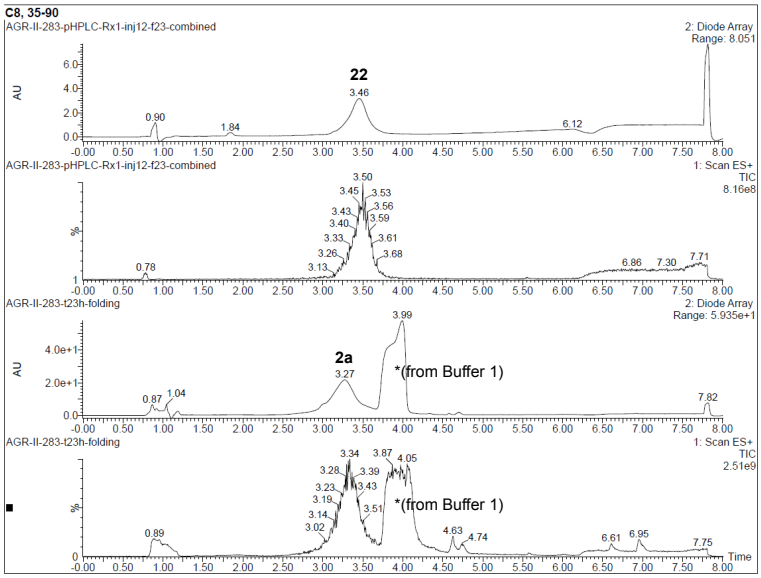


Figure 1.19C Top: UV Trace from UPLC-MS analysis of pure G-CSF **22**; gradient of 35% to 90% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C8 analytical column. **Middle (1)**: ESI-MS (+) TIC. **Middle (2):** UV Trace from UPLC-MS analysis of crude G-CSF **2a** in folding buffer at t = 23h; gradient of 35% to 90% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C8 analytical column. **Bottom:** ESI-MS (+) TIC.

Comparison (**22**) and crude (**2a**). LR-MS

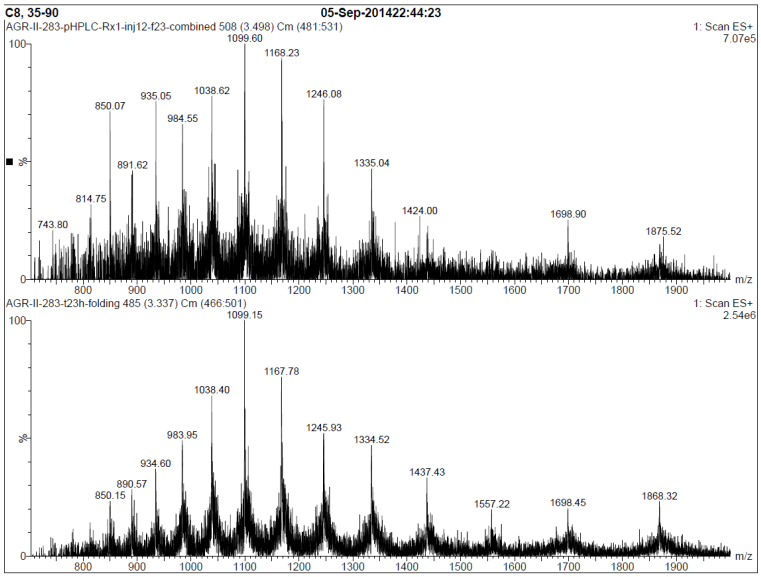


Figure 1.20A Top: ESI-MS (+) TIC of pure G-CSF **22** (non-folded). **Bottom:** ESI-MS (+) TIC of G-CSF **2a** (folded) in crude folding buffer.

Comparison (**22**) and crude (**2a**). LR-MS (smoothened)

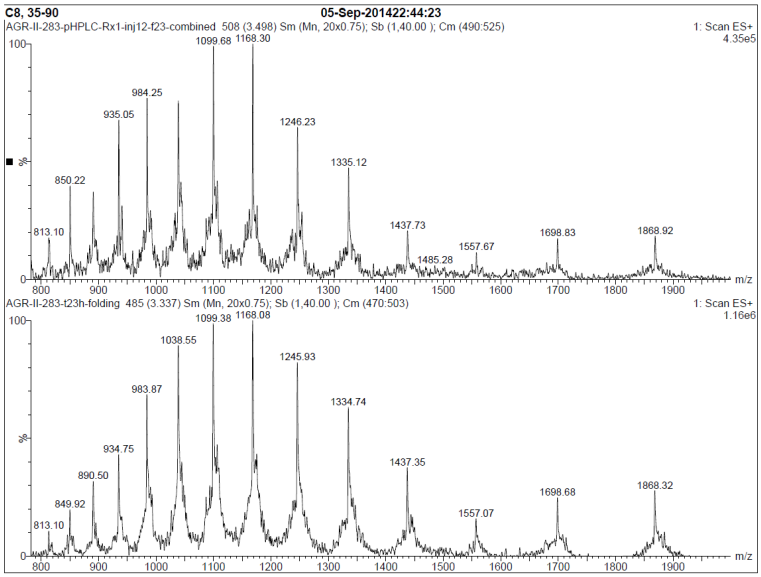


Figure 1.20B Top: ESI-MS (+) TIC of pure G-CSF **22** (non-folded). **Bottom:** ESI-MS (+) TIC of G-CSF **2a** (folded) in crude folding buffer. (data smoothened low resolution data supports decrease in mass, see also HRMS)

Compound (**2b**) (pure) recombinant G-CSF in folding buffer

LC-MS

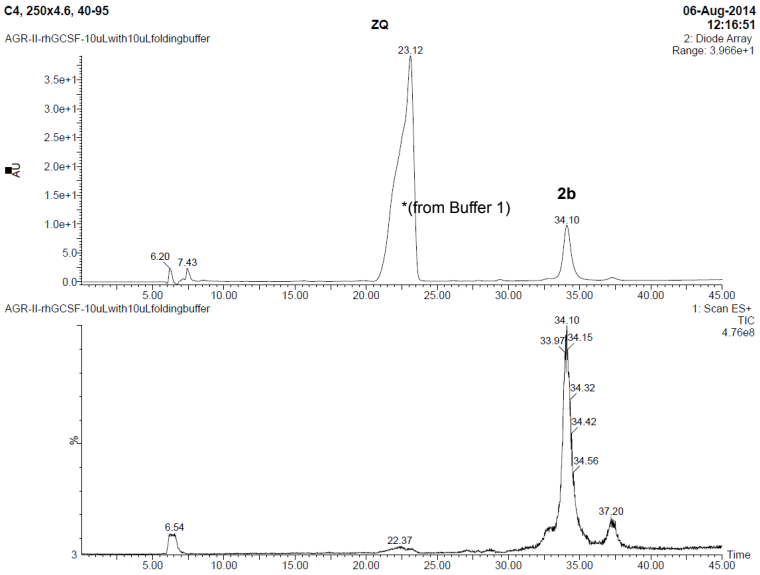


Figure 1.21A Top: UV Trace from LC-MS analysis of recombinant G-CSF **2b**; gradient of 40% to 95% solvent B over 60 min at a flow rate of 0.5 mL/min, Higgins C4 semi-preparative column. **Bottom:** ESI-MS (+) TIC.

Comparison (**2b**) (pure) (**2a**) (crude) (**2a**) (purify)

LC-MS

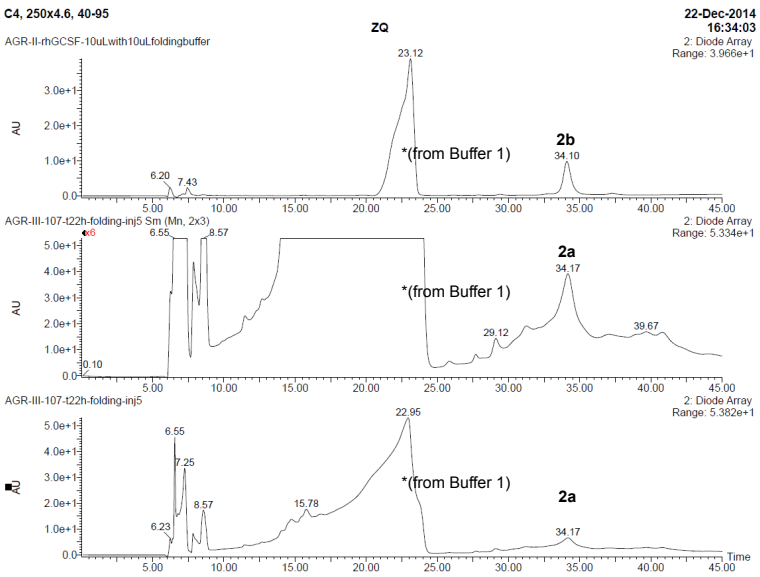


Figure 1.21B Top: UV Trace from LC-MS analysis of recombinant G-CSF **2b** (t = 34.10 min); gradient of 40% to 95% solvent B over 60 min at a flow rate of 0.5 mL/min, Higgins C4 semi-preparative column. **Middle:** UV Trace from LC-MS analysis of crude G-CSF **2a** (t = 34.17 min) after folding; same system as above. **Bottom:** Sample injection for purification.

Comparison
(**2b**) (pure)
(**2a**) (crude)
(**2a**) (pure)

LC-MS

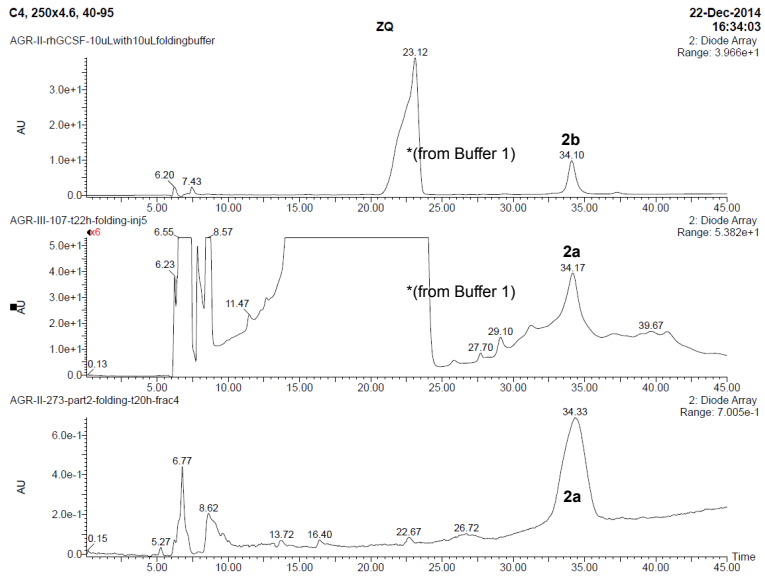


Figure 1.21C Top: UV Trace from LC-MS analysis of recombinant G-CSF **2b** (t = 34.10 min); gradient of 40% to 95% solvent B over 60 min at a flow rate of 0.5 mL/min, Higgins C4 semi-preparative column. **Middle:** UV Trace from LC-MS analysis of crude G-CSF **2a** (t = 34.17 min) after folding; same system as above. **Bottom:** Reinjection of pure fractions after collection, pure G-CSF **2a** (t =34.33 min); same system as above.

Compound (**2a**).
(purification)

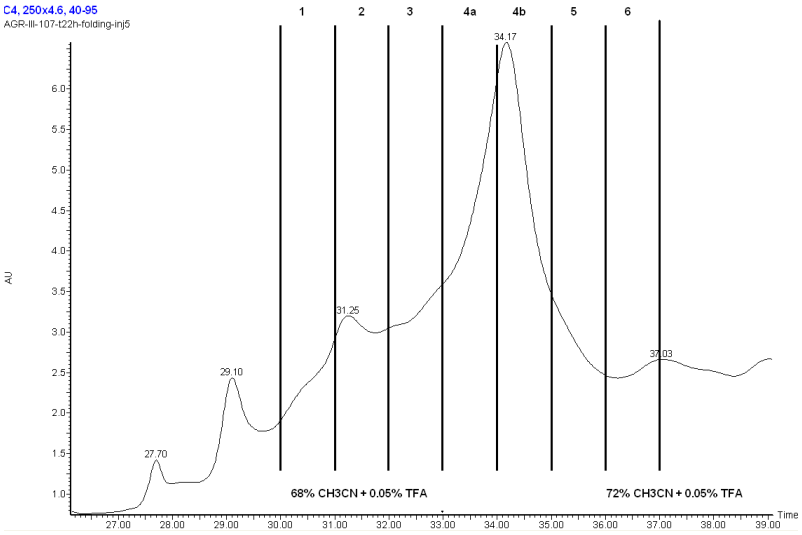


Figure 1.22 UV Trace from HPLC purification of G-CSF **2a** (range 26 to 39 min); gradient of 40% to 95% solvent B over 60 min at a flow rate of 0.5 mL/min, Higgins C4 semi-preparative column. Seven fractions (1, 2, 3, 4a, 4b, 5 and 6) were collected (each approx. 0.5 mL of 70:30 CH_3CN / H_2O with 0.05% TFA neutralized with 0.1 mL with 1.0 M HEPES). Fractions 4a and 4b exhibited the highest activity relative to the others in the NFS-60 cellular assay Isolated yield is based upon combined recovery from collection of fractions 4a and 4b.

Compound (**2a**). (purification)

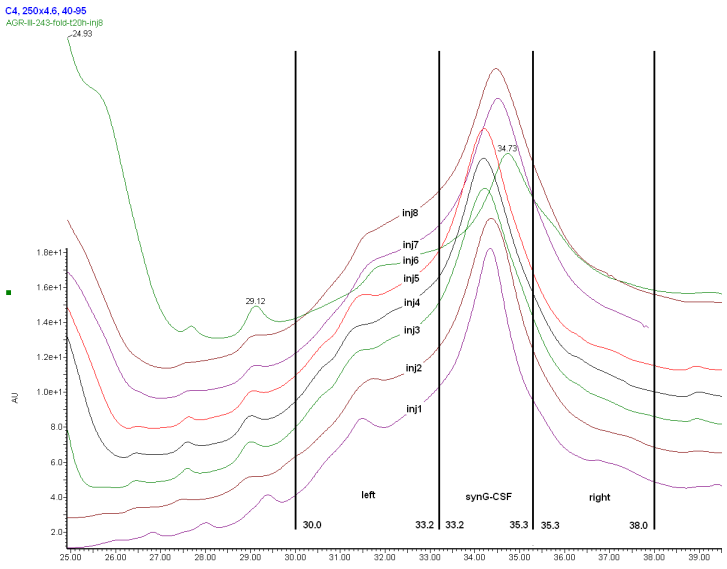


Figure 1.23 UV Traces from HPLC purification (8 injections) of G-CSF **2a** (range 25 to 39 min); gradient of 40% to 95% solvent B over 60 min at a flow rate of 0.5 mL/min, Higgins C4 semi-preparative column.

Compound (**2a**).

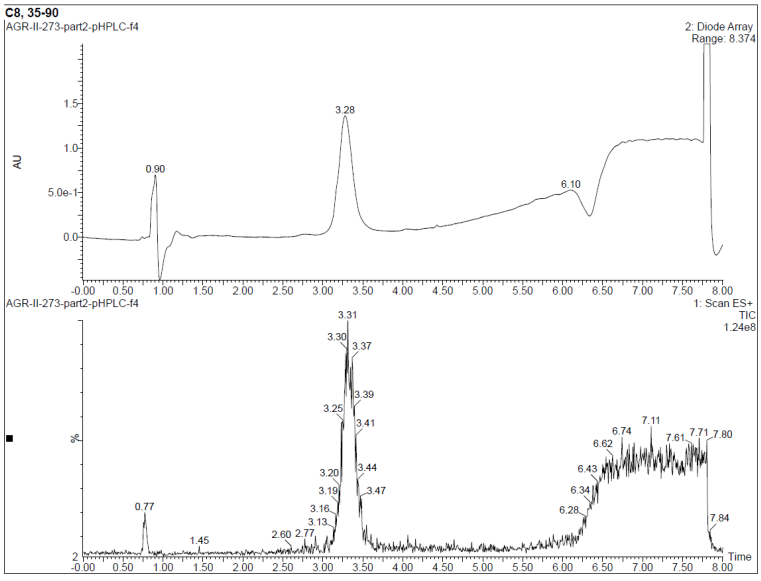


Figure 1.24A Top: UV Trace from UPLC-MS analysis of purified G-CSF **2a**; gradient of 35% to 90% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C8 analytical column. **Bottom:** ESI-MS (+) TIC.

Comparison
(**22**) (pure)
(**2a**) (pure)

UPLC-MS

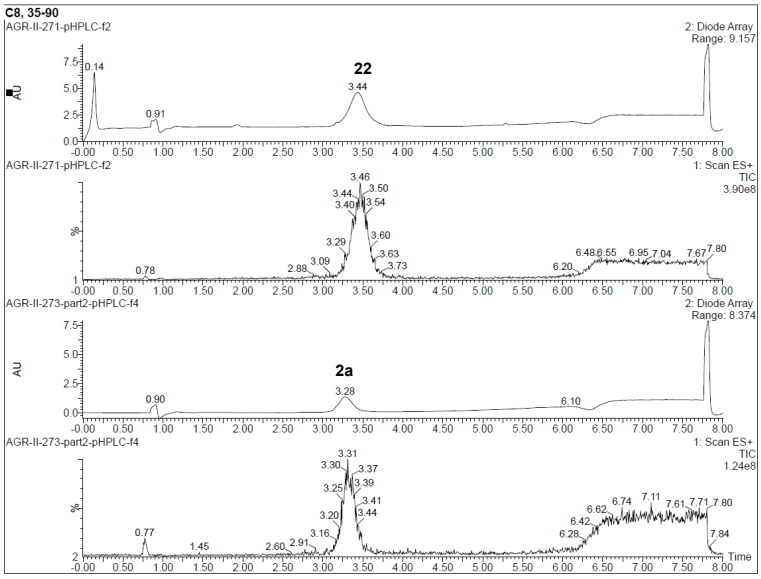


Figure 1.24B Top: UV Trace from UPLC-MS analysis of purified G-CSF **22** (t = 3.44 min); gradient of 35% to 90% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C8 analytical column. **Middle (1)**: ESI-MS (+) TIC. **Middle (2):** UV Trace from UPLC-MS analysis of purified G-CSF **2a** (t = 3.28 min); same system as above. **Bottom:** ESI-MS (+) TIC.

Comparison
(**22**) (pure)
(**2a**) (pure)

LC-MS

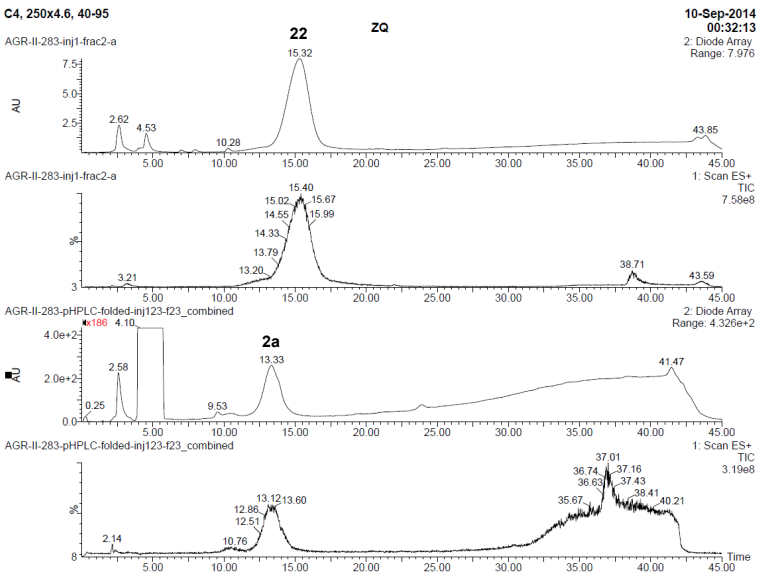


Figure 1.25 Top: UV Trace from LC-MS analysis of purified G-CSF **22** (t = 15.32 min); gradient of 40% to 95% solvent B over 60 min at a flow rate of 0.2 mL/min, X-Bridge C8 analytical column. **Middle (1**): ESI-MS (+) TIC. **Middle (2):** UV Trace from LC-MS analysis of purified G-CSF **2a** (t = 13.33 min); same system as above. **Bottom:** ESI-MS (+) TIC.

Compound (**22**). HRMS

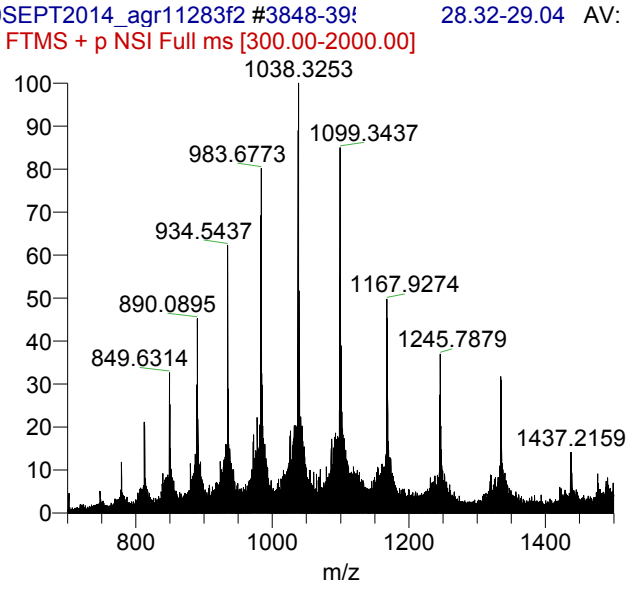


Figure 1.26 High resolution ESI-MS (+) of linear G-CSF **22**. M = $C_{840}H_{1334}N_{222}O_{242}S_8$. See Figure 1.28.

Compound (**2a**). HRMS

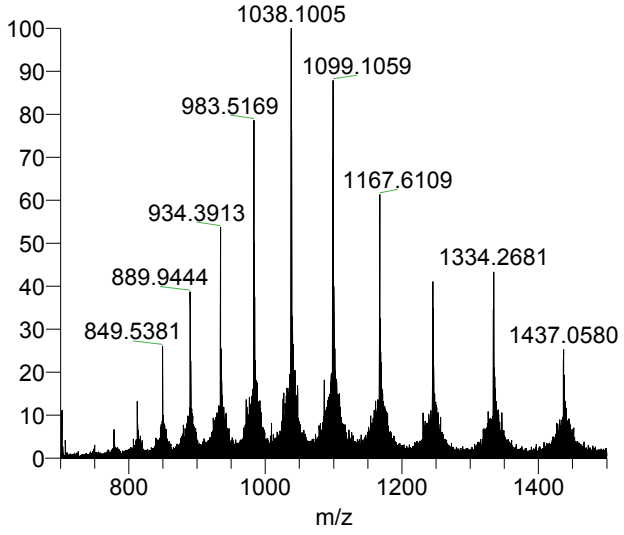


Figure 1.27 High resolution ESI-MS (+) of linear G-CSF **2a**. M = $C_{840}H_{1330}N_{222}O_{242}S_8$. See Figure 1.28.

Comparison
HRMS (**22**) and (**2a**).

G-CSF (**22**)	AGR-II-283-comb.		
Found	**Charge**	**# of H+**	**(M+1H+)+1**
1437.21478	13	13.1033	18671.6968
1334.69973	14	14.1113	18672.6929
1245.72040	15	15.1192	18671.6947
1167.92581	16	16.1272	18671.6937
1099.40185	17	17.1351	18673.7043
1038.26919	18	18.1430	18671.7103
983.72884	19	19.1510	18672.7049
934.59343	20	20.1589	18672.7176
890.13666	21	21.1669	18672.7109
849.67613	22	22.1748	18671.7080
812.86453	23	23.1828	18673.7093
778.95388	24	24.1907	18671.7103
		average	**18672.4312**

G-CSF (**2a**)	AGR-II-283-folded-f23-combined		
Found	**Charge**	**# of H+**	**(M+1H+)+1**
1437.05813	13	13.1033	18669.6603
1334.26812	14	14.1113	18666.6504
1245.51803	15	15.1192	18668.6592
1167.61089	16	16.1272	18666.6550
1098.98785	17	17.1351	18666.6663
1038.10047	18	18.1430	18668.6734
983.51685	19	19.1510	18668.6771
934.39127	20	20.1589	18668.6744
889.94443	21	21.1669	18668.6741
849.53810	22	22.1748	18668.6713
812.73301	23	23.1828	18670.6844
778.74310	24	24.1907	18666.6516
		average	**18668.3951**

Figure 1.28 Left: HRMS data from purified G-CSF **22**, (18,672.4312 Da) **Right:** HRMS data from purified G-CSF **2a**, (18,668.3951 Da). Δ = 4.04 (accounts for disulfide formations, 4 H)

Compound (**17**).

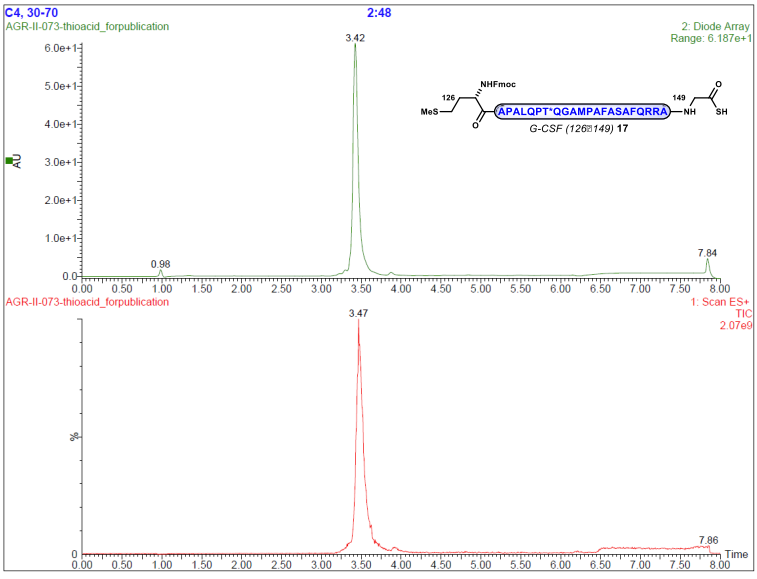


Figure 1.29A Top: UV Trace from UPLC-MS analysis of purified peptide **17**; gradient of 30% to 70% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**17**).

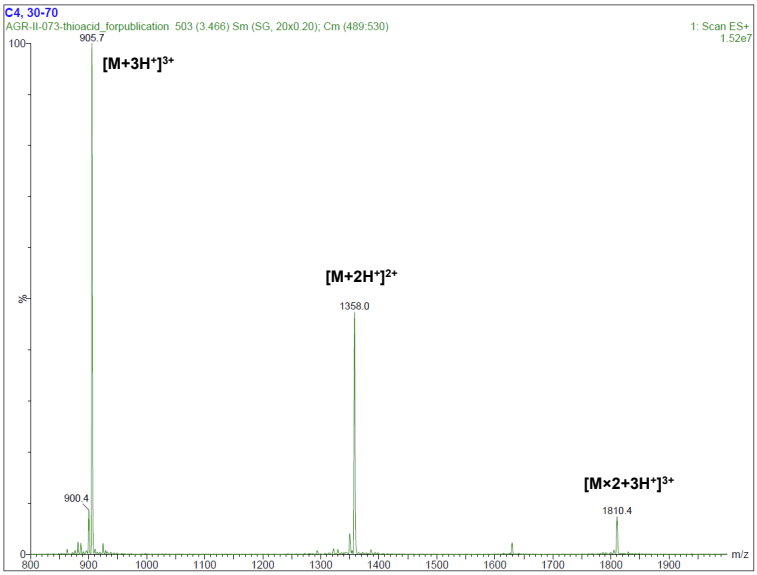


Figure 1.29B ESI-MS (+) of peptide **17**. M = $C_{123}H_{181}N_{33}O_{31}S_3$. ESI calculated for $[M\times2+3H^+]^{3+}$ *m/z*: 1810.5, found: 1810.4; $[M+2H^+]^{2+}$ *m/z*: 1358.1, found: 1358.0; $[M+3H^+]^{3+}$ *m/z*: 905.7, found: 905.7.

Compound (**18**).

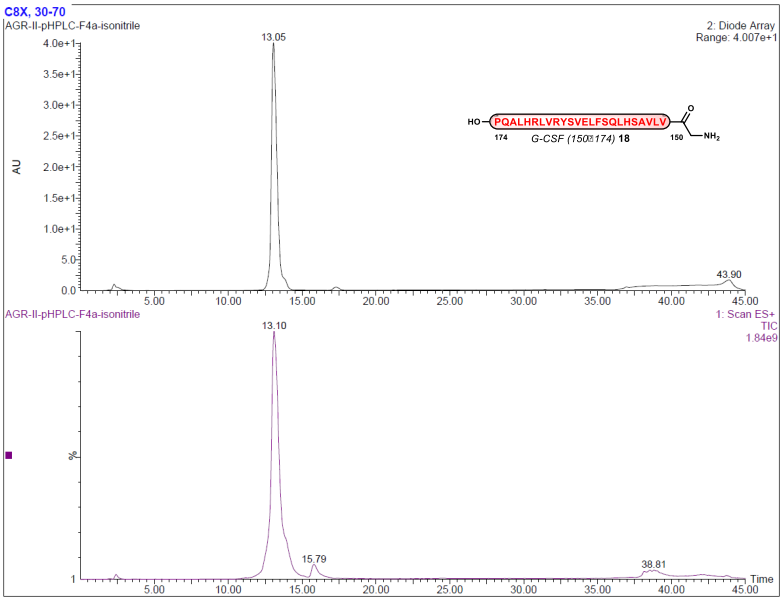


Figure 1.30A Top: UV Trace from LC-MS analysis of HPLC purified peptide **18**; gradient of 30% to 70% solvent B over 30 min at a flow rate of 0.2 mL/min, X-Bridge C8 analytical column.

Compound (**18**).

Figure 1.30B ESI-MS (+) of peptide **18**. M = $C_{129}H_{207}N_{37}O_{34}$. ESI calculated for $[M\times2+3H^+]^{3+}$ *m/z*: 1881.20, found: 1881.82; $[M+2H^+]^{2+}$ *m/z*: 1411.16, found: 1411.41; $[M+3H^+]^{3+}$ *m/z*: 941.11, found: 941.21; $[M+4H^+]^{4+}$ *m/z*: 706.08, found: 706.08.

Compound (**13**). (from coupling)

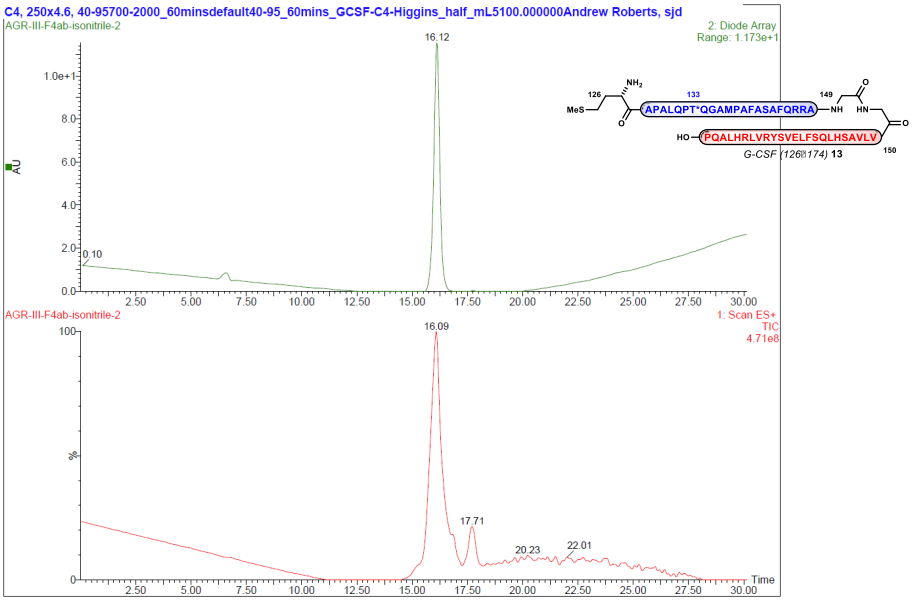


Figure 1.31A Top: UV Trace from LC-MS analysis of HPLC purified peptide **13**; gradient of 30% to 70% solvent B over 30 min at a flow rate of 0.2 mL/min, X-Bridge C8 analytical column.

Compound (**13**). (from coupling)

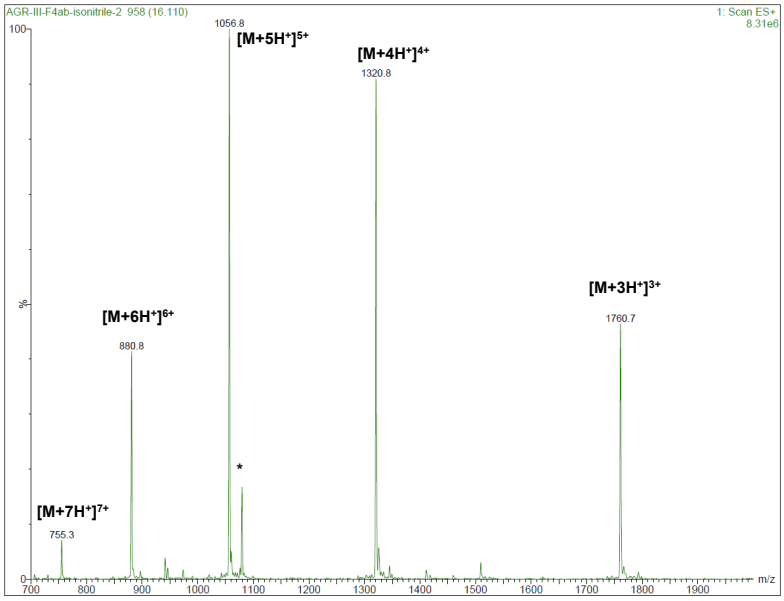


Figure 1.31B ESI-MS (+) of peptide **13**. M = $C_{237}H_{376}N_{70}O_{63}S_2$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1760.4, found: 1760.7; $[M+4H^+]^{4+}$ *m/z*: 1320.6, found: 1320.8; $[M+5H^+]^{5+}$ *m/z*: 1056.6, found: 1056.8; $[M+6H^+]^{6+}$ *m/z*: 880.7, found: 880.8; $[M+7H^+]^{7+}$ *m/z*: 750.0, found: 750.3. * indicates presence of TFA adduct in spectrum.

Compound (**19**)

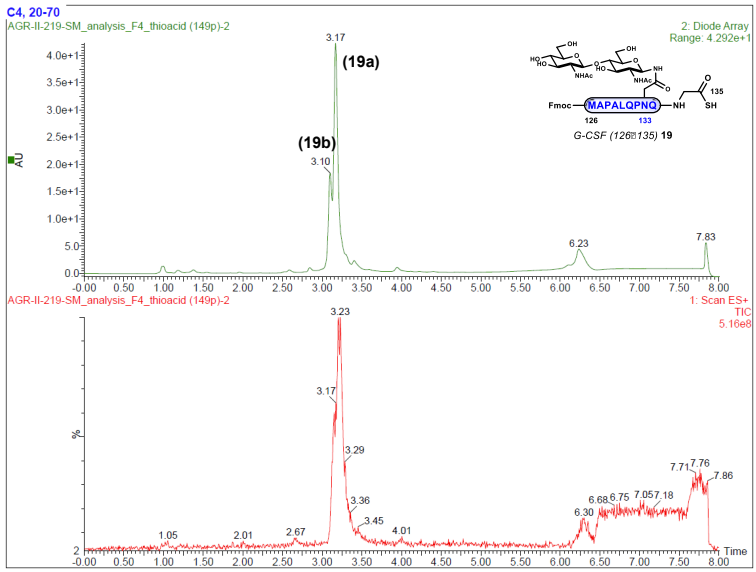


Figure 1.32A Top: UV Trace from UPLC-MS analysis of purified peptide **19a** containing **19b**; gradient of 20% to 70% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**19**)

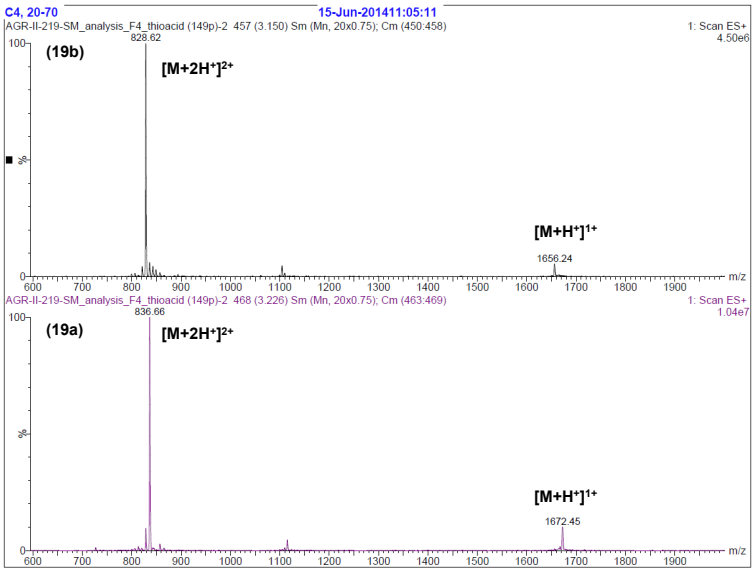


Figure 1.32B Top: ESI-MS (+) of peptide **19b**. M = $C_{74}H_{107}N_{15}O_{26}S_1$. ESI calculated for $[M+H^+]^{1+}$ *m/z*: 1655.82, found: 1656.24; $[M+2H^+]^{2+}$ *m/z*: 828.41, found: 828.62. **Bottom:** ESI-MS (+) of peptide **19a**. M = $C_{74}H_{107}N_{15}O_{25}S_2$. ESI calculated for $[M+H^+]^{1+}$ *m/z*: 1671.88, found: 1672.45; $[M+2H^+]^{2+}$ *m/z*: 836.44, found: 836.66.

Compound (**19**) (purity analysis)

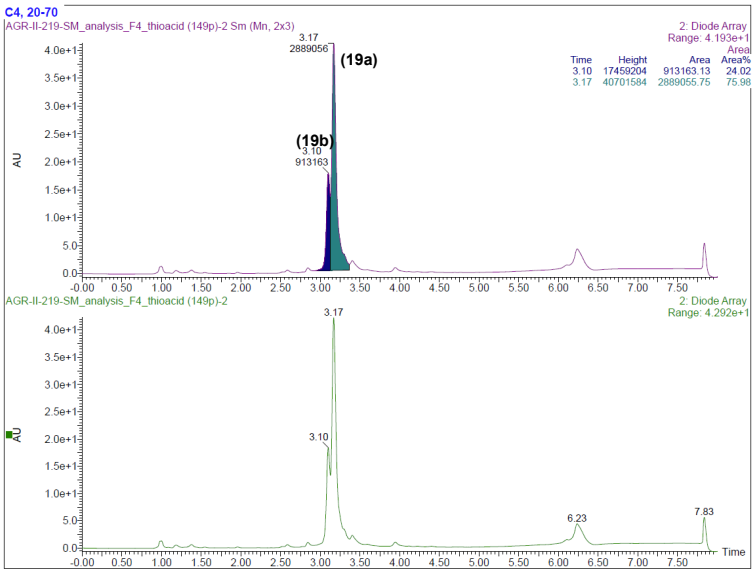


Figure 1.32C Top: Integrated UV Trace from UPLC-MS analysis of purified peptide **19a** (75% AUC) containing **19b** (25% AUC). **Bottom:** UV Trace from UPLC-MS analysis of purified peptide **19a** containing **19b**; gradient of 20% to 70% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column.

Compound (**19**) (purity analysis)

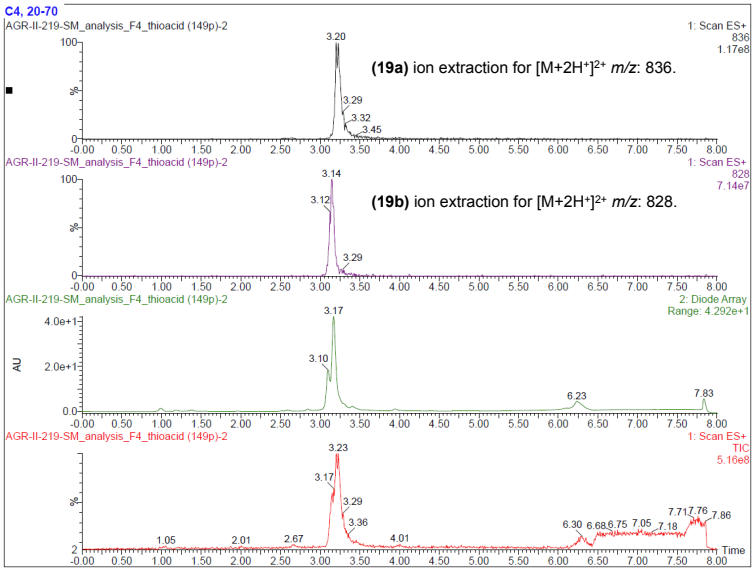


Figure 1.32D Top: ESI (+), ion extraction for $[M+2H^+]^{2+}$ *m/z*: 836 (**19a**). **Middle (1):** ESI (+), ion extraction for $[M+2H^+]^{2+}$ *m/z*: 828 (**19b**). **Middle (2):** UV Trace from UPLC-MS analysis of purified peptide **19a** containing **19b**; gradient of 20% to 70% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**20**).

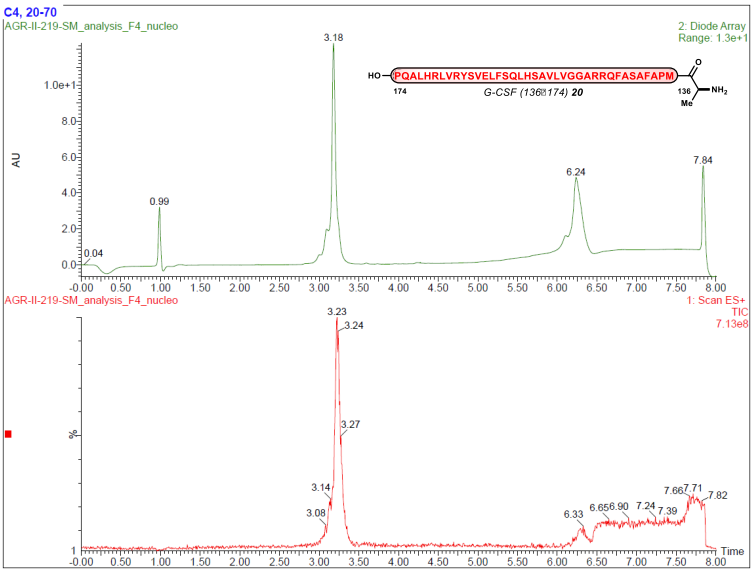


Figure 1.33A Top: UV Trace from UPLC-MS analysis of purified peptide **20**; gradient of 20% to 70% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**20**).

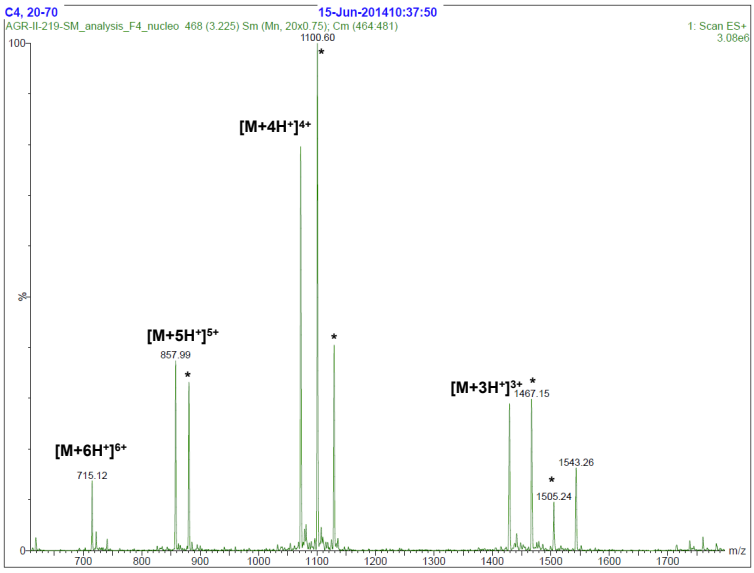


Figure 1.33B ESI-MS (+) of peptide **20**. M = $C_{194}H_{306}N_{58}O_{50}S$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1428.67, found: 1428.21; $[M+4H^+]^{4+}$ *m/z*: 1071.76, found: 1072.16; $[M+5H^+]^{5+}$ *m/z*: 857.61, found: 857.99; $[M+6H^+]^{6+}$ *m/z*: 714.84, found: 715.12. * indicates presence of TFA adduct(s) in spectrum.

Compound (**21**). (crude reaction)

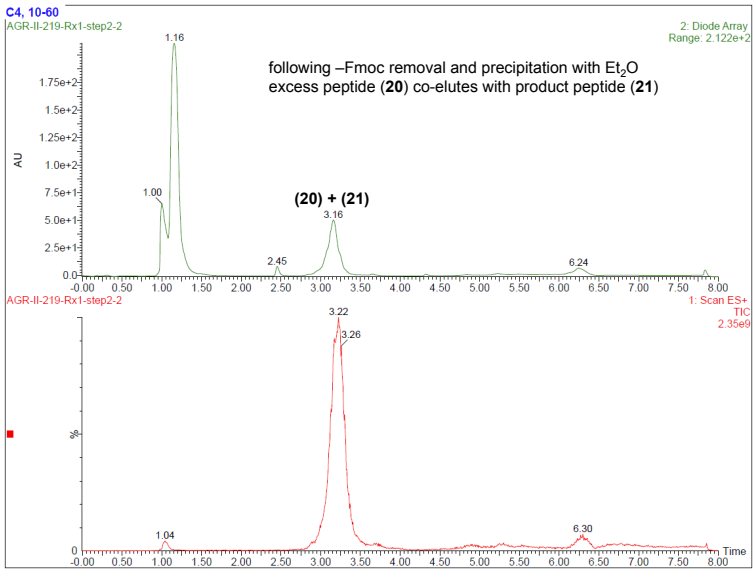


Figure 1.34A Top: UV Trace from UPLC-MS analysis of crude reaction to prepare peptide **21**; gradient of 10% to 60% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**21**). (crude reaction)

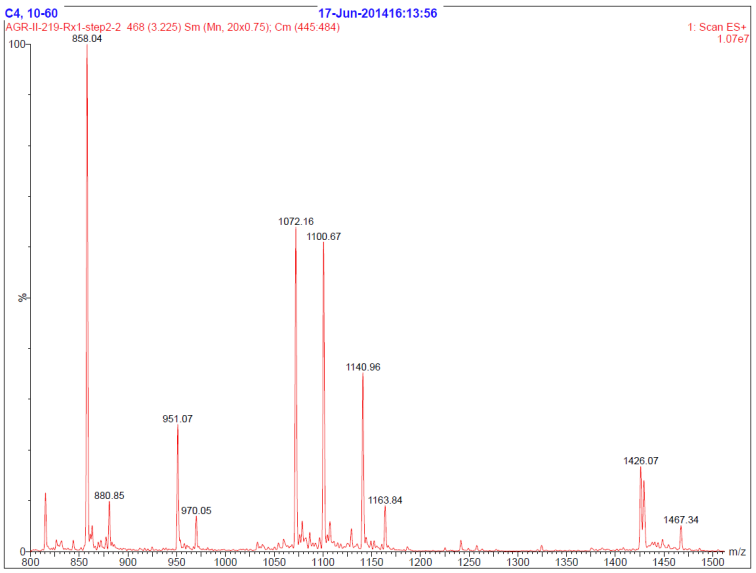


Figure 1.34B ESI-MS (+) of crude reaction mixture to prepare peptide **21**.

Compound (**21**).

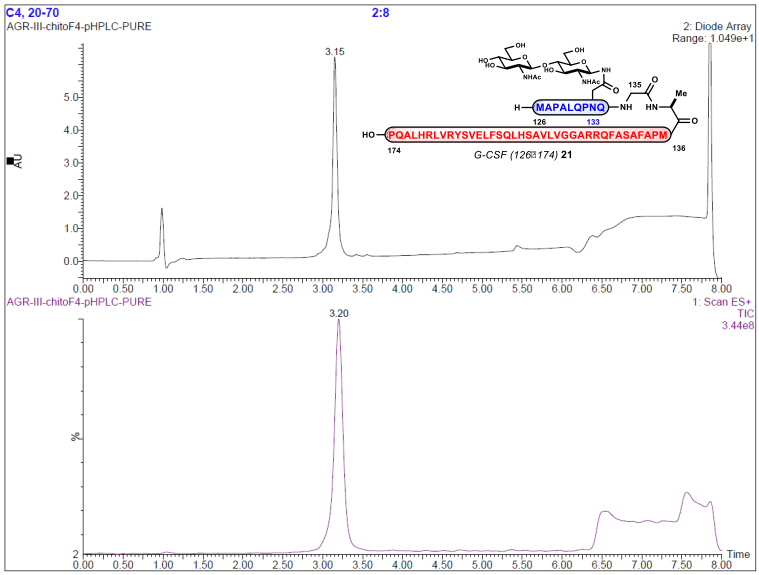


Figure 1.35A Top: UV Trace from UPLC-MS analysis of purified peptide **21**; gradient of 20% to 70% solvent B over 6 min at a flow rate of 0.3 mL/min, Acquity UPLC BEH C4 analytical column. **Bottom:** ESI-MS (+) TIC.

Compound (**21**).

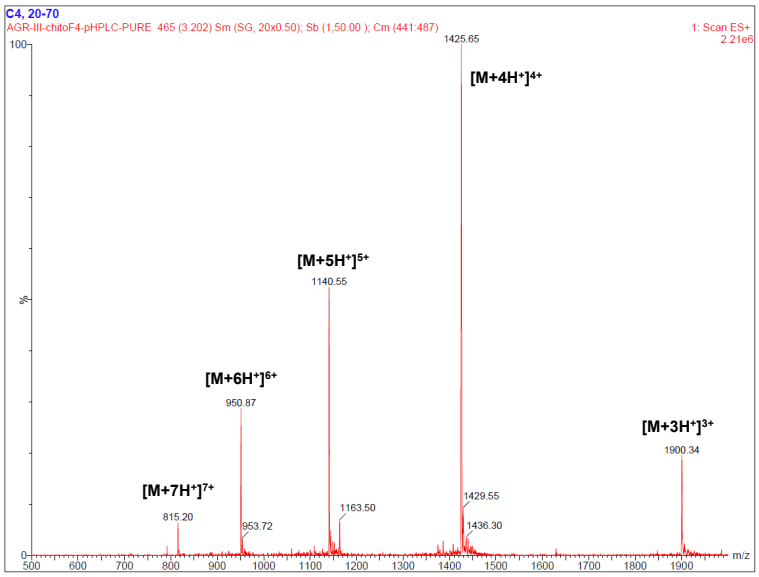


Figure 1.35B ESI-MS (+) of peptide **21**. M = $C_{253}H_{401}N_{73}O_{73}S_2$. ESI calculated for $[M+3H^+]^{3+}$ *m/z*: 1900.19, found: 1900.34; $[M+4H^+]^{4+}$ *m/z*: 1425.40, found: 1425.65; $[M+5H^+]^{5+}$ *m/z*: 1140.52, found: 1140.55; $[M+6H^+]^{6+}$ *m/z*: 950.60, found: 950.87; $[M+7H^+]^{7+}$ *m/z*: 814.94, found: 815.20.

Compound (**11**).

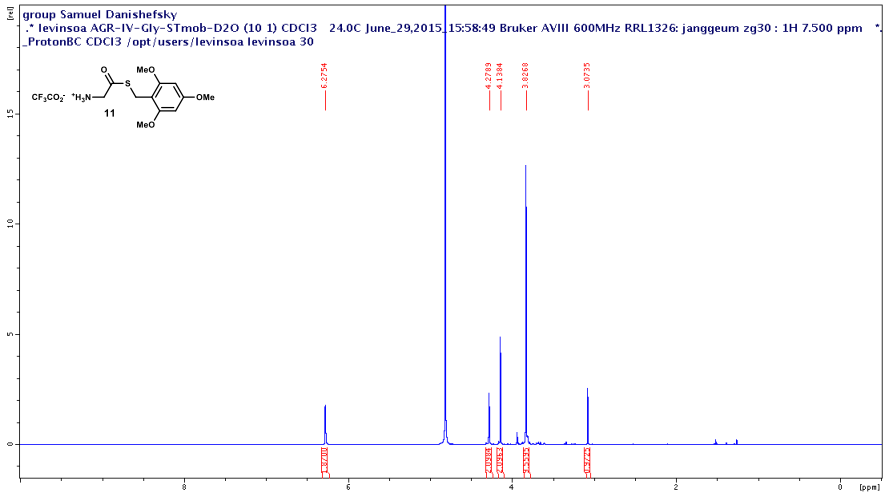


Figure 1.36A ^1H NMR spectrum of compound **11** in D_2O.

Compound (**11**).

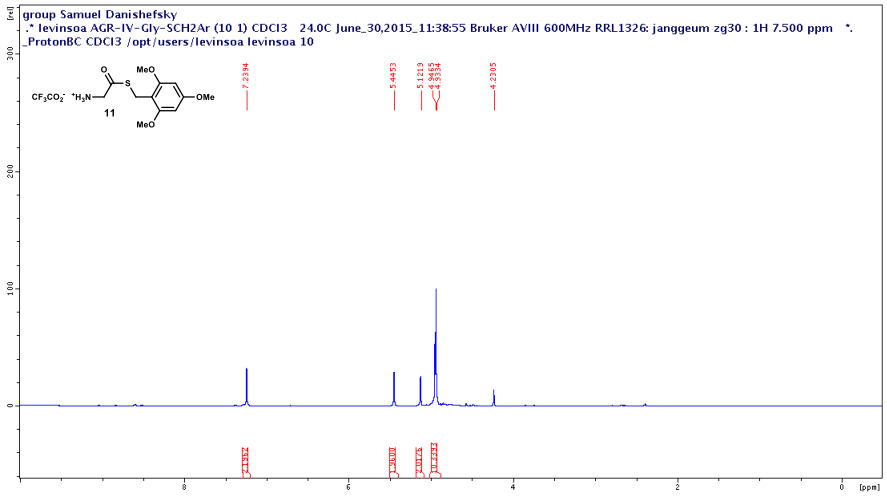


Figure 1.36B ^{1}H NMR spectrum of compound **11** $CDCl_3$ with DMSO-d_6.

Compound (**11**).

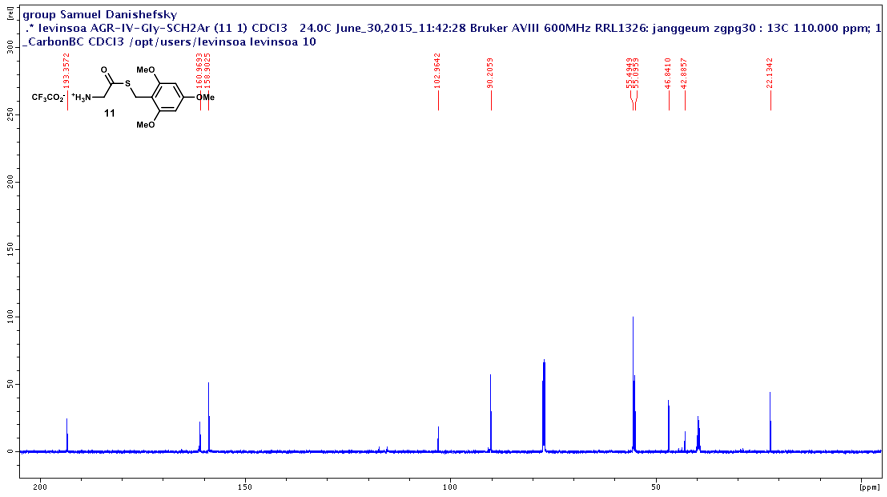


Figure 1.36C ^{13}C NMR spectrum of compound **11** in $CDCl_3$ with DMSO-d_6.

Compound (**S-1**).

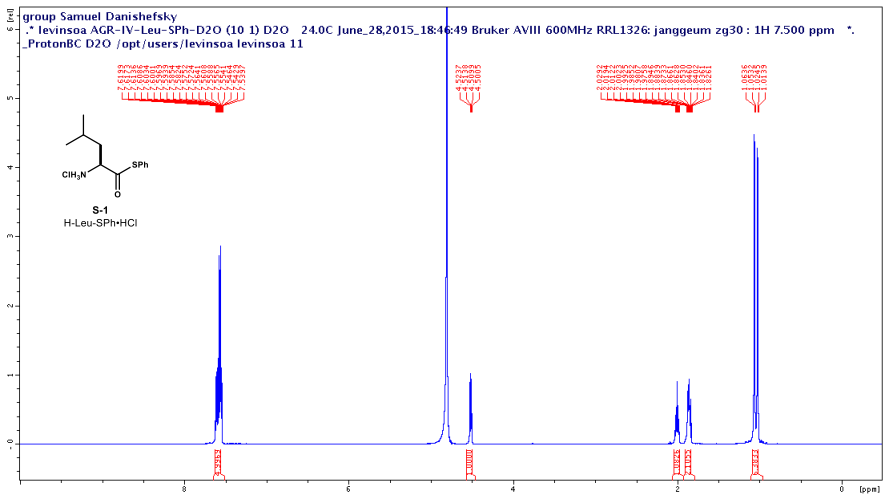


Figure 1.37A ^{1}H NMR spectrum of compound **S-1** in D_2O.

Compound (**S-1**).

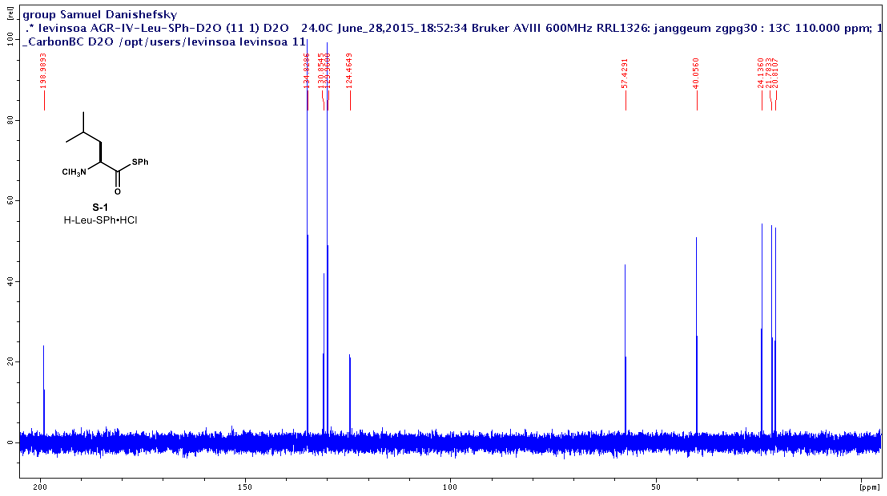


Figure 1.37B ^{13}C NMR spectrum of compound **S-1** in D_2O.

Compound (**S-2**).

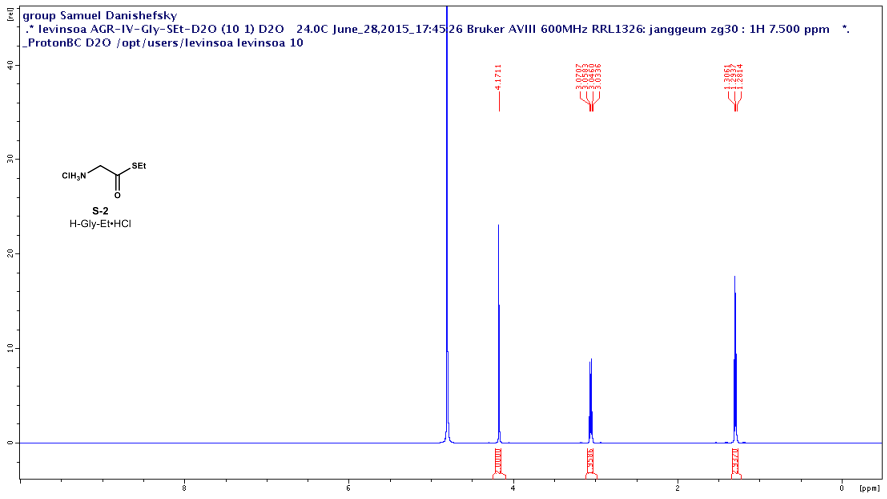


Figure 1.38A ^{1}H NMR spectrum of compound **S-2** in D_2O.

Compound (**S-2**).

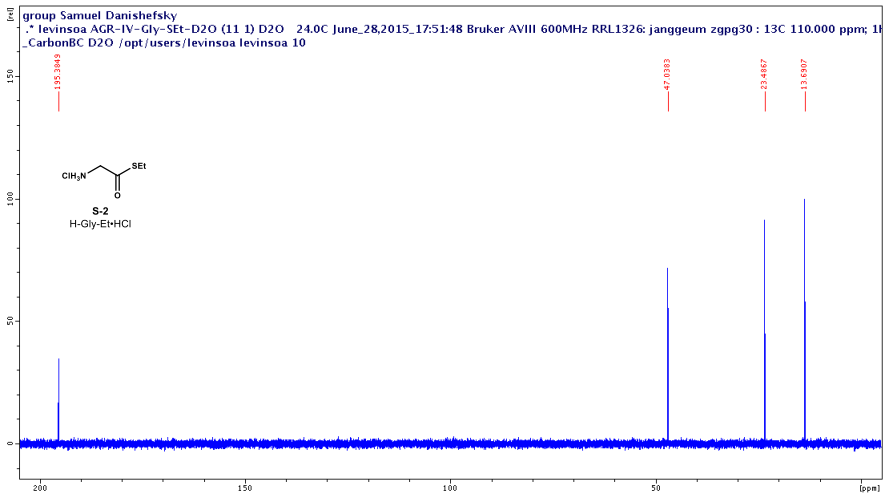


Figure 1.38B ^{13}C NMR spectrum of compound **S-2** in D_2O.

Compound (**S-3**).

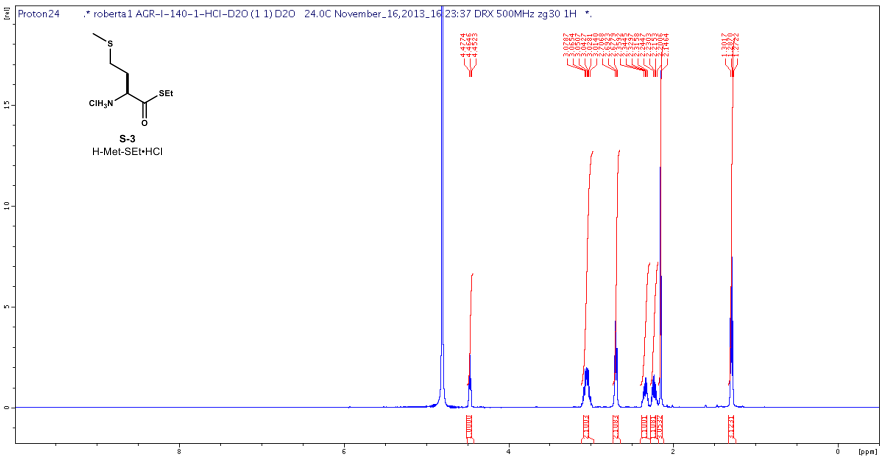


Figure 1.39A ^{1}H NMR spectrum of compound **S-3** in D_2O.

Compound (**S-3**).

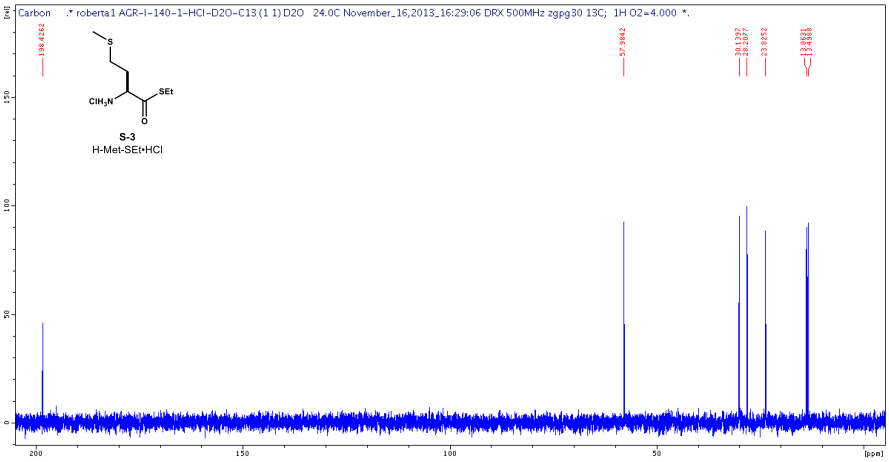


Figure 1.39B ^{13}C NMR spectrum of compound **S-3** in D_2O.

Compound (**S-4**).

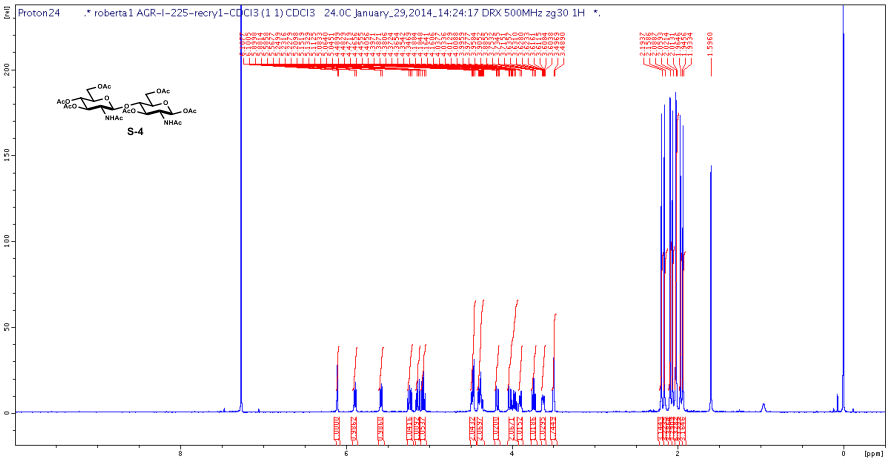


Figure 1.40A ^{1}H NMR spectrum of compound **S-4**.

Compound (**S-4**).

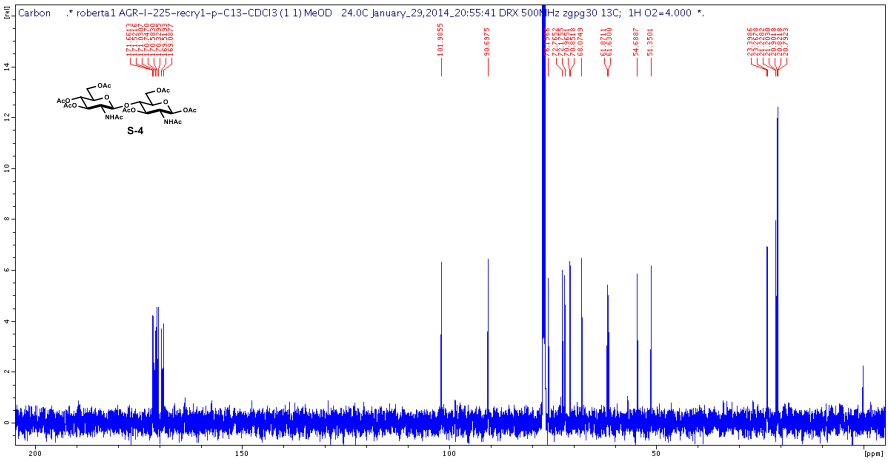


Figure 1.40B ^{13}C NMR spectrum of compound **S-4**.

Compound (**S-5**).

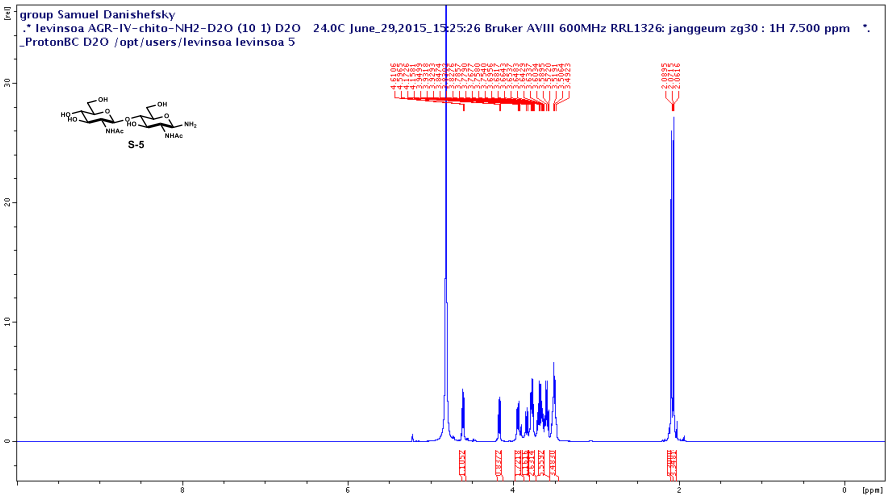


Figure 1.41A ^{1}H NMR spectrum of compound **S-5** in D_2O.

Compound (**S-5**).

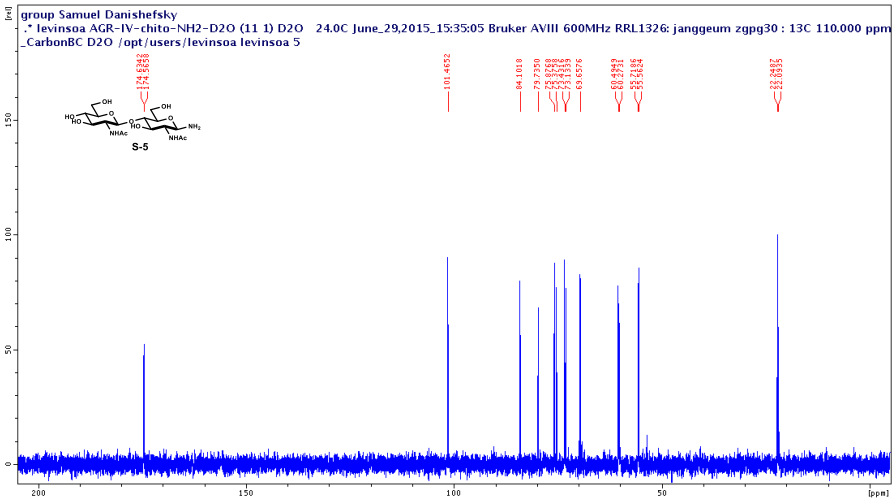


Figure 1.42B ^{13}C NMR spectrum of compound **S-5** in D_2O.

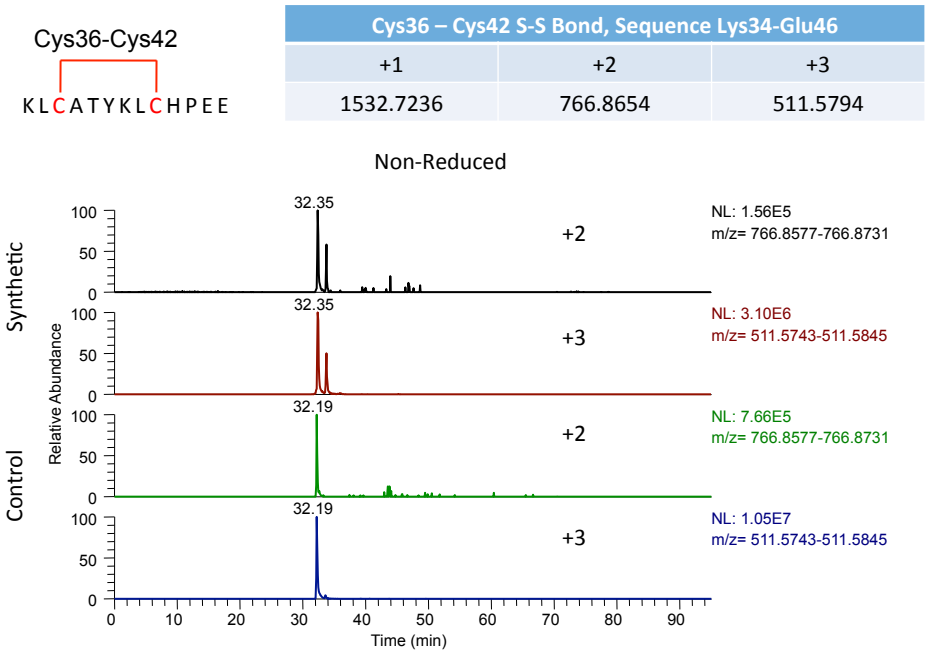


Figure S-3. Data for non-reduced analysis of synthetic G-CSF (**2a**) and recombinant G-CSF (**2b**), Lys34-Glu46.

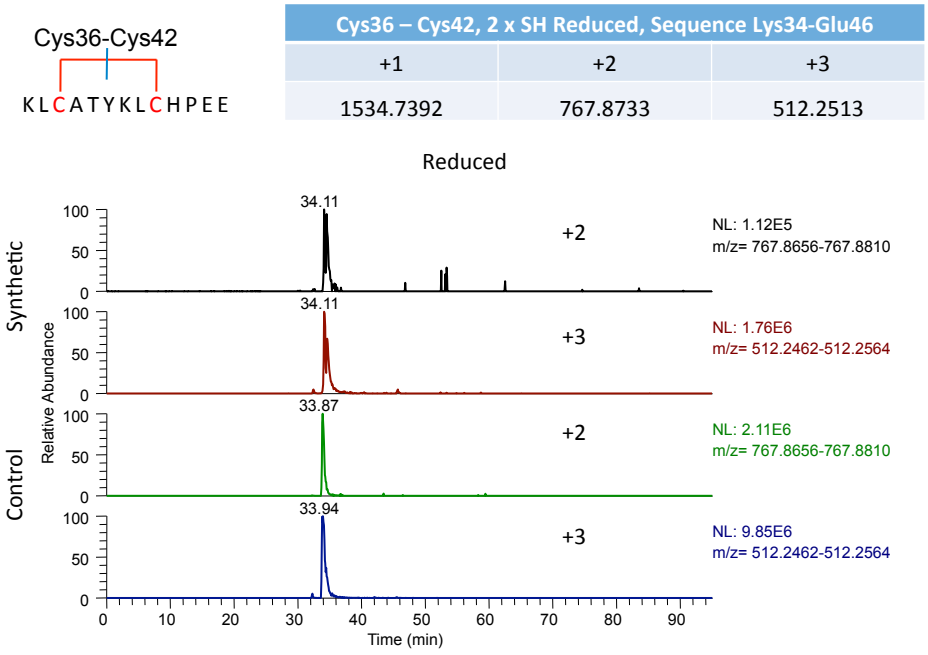


Figure S-4. Data for reduced analysis of synthetic G-CSF (**2a**) and recombinant G-CSF (**2b**), Lys34-Glu46.

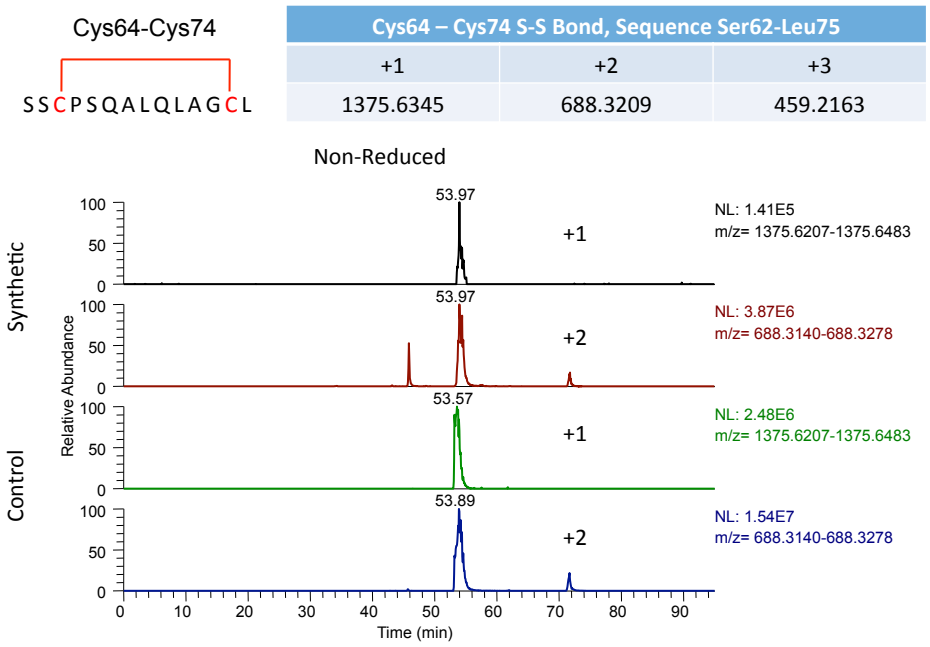


Figure S-5. Data for reduced analysis of synthetic G-CSF (**2a**) and recombinant G-CSF (**2b**), Ser62-Leu75.

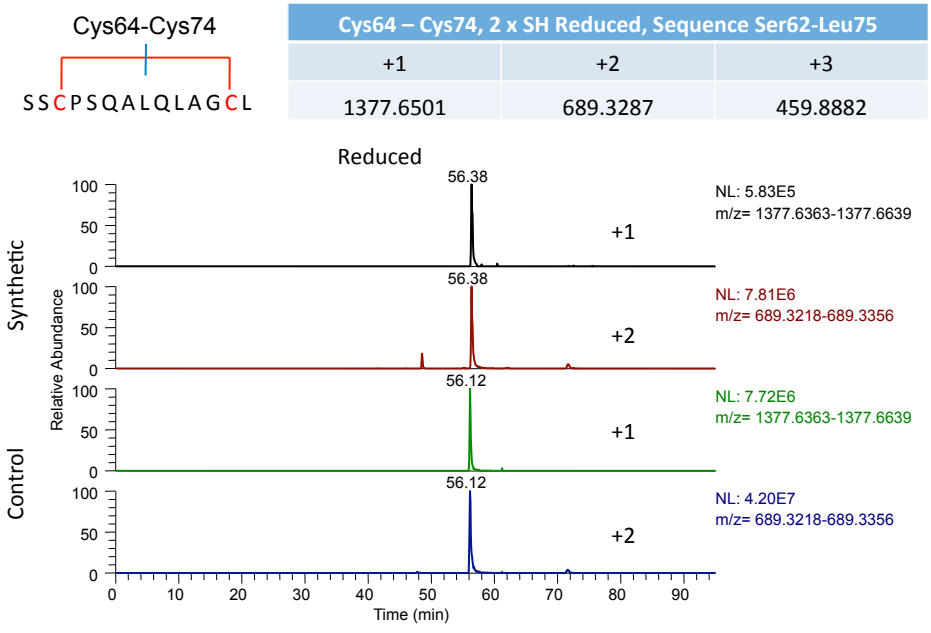


Figure S-6. Data for reduced analysis of synthetic G-CSF (**2a**) and recombinant G-CSF (**2b**), Ser62-Leu75.