Supplementary Information

Optimization of Semisynthetic Approach for Glycosyl Interferon-β-polypeptide by Utilizing Bacterial Protein Expression and Chemical Modification

Yie Kie Chong[a], Chaitra Chandrashekar[a], Donglin Zhao[a], Yuta Maki[a], Ryo Okamoto[a], Yasuhiro Kajihara[a,*]

[a] *Department of Chemistry and Project Research Centre for Fundamental Sciences, Graduate School of Science, Osaka University, 1-1 Machikaneyama, Toyonaka, Osaka 560-0043, Japan*

Table of Contents

Experimental Procedures

General and abbreviation:

1-Hydroxybenzotriazole (HOBt), Boc-Lys(2-Cl-Z)-OH, Boc-Glu(OBzl)-OH, Boc-Val-OH, Boc-Ile-OH, Boc-Leu-OH, Boc-Ala-OH, Boc-Gly-OH, Boc-Met-OH, Boc-Pro-OH, Boc-Phe-OH, Boc-Gln-OH, Boc-Thr(Bzl), Boc-Cys(Acm), Boc-Ser(Bzl), Boc-Asp(OBzl) and Boc-Trp(For) were purchased from Peptide Institude.Inc. Boc-Tyr(Br-Z)-OH, Boc-Asn(Xan)-OH, Boc-Arg(Mtr)-OH, Benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP), N,N,N'N'-Tetramethyl-O-(1H-benzotriazol-1-yl)uranium hexafluorophosphate (HBTU) and 3,4-Dihydro-3-hydroxyl-4-oxo-1,2,3-benzotriazole (HOOBt) were purchased from Watanabe Chemical Industries, LTD. Aminomethyl ChemMatrix® resin was purchased from Sigma Aldrich. Acetonitrile, dimethylformamide, dichloromethane were purchased from Kanto Chemical Co., Inc. Boc-OSu was purchased from Combi-Blocks. N-methylimidazole, Triisopropylsilane (TIPS), Tris(2-carboxyethyl)phosphine Hydrochloride (TCEP), Sodium 2-mercaptoethanesulfonate (MESNA), tert-butyl carbazate, hexamethyleneimine, Trifluoromethanesulfonic acid (TfOH), and 1,2-ethanedithiol (EDT) were purchased from Tokyo Chemical Industry Co., Ltd. Acetic acid, thioanisol, S-trityl-3-mercaptopropionic acid, and piperidine were purchased from Nacalai Tesque. Boc-L-thiozolidine-4-carboxylic acid was purchased from BACHEM. Hydrazine monohydrate, 1-methyl-2-pyrrolidone, N,N'-diisopropylcarbodiimifr (DIC), silver acetate, 2,2'-azobis dihydrochloride (VA-044), dithiothreitol (DTT), dimethyl sulfoxide (DMSO), 2-amino-2-hydroxymethyl-1,3-propanediol (Tris), trifluoroacetic acid (TFA), Guanidine Hydrochloride (Gn-HCl), p-mercaptophenylacetic acid (MPAA) were purchased from Wako Chemical Industries, Ltd. Sodium phosphate buffer was freshly prepared in our laboratory.

Section 1: Synthesis of IFN-β-(79-88) Segment C 10

Synthesis of IFN-β-(81-88)-Segment C: H_2N-[Glu^{81}–Thr^{82}–Ile^{83}– Val^{84}–Glu^{85}–Asn^{86}–Leu^{87}–Leu^{88}]-NHNHBoc 6

IFN-β-(81-88)-Segment C-NHNHBoc **6** was synthesized on Aminomethyl ChemMatrix® resin (100 μmol) by using the improved Boc SPPS.[1] The resin was swelled by washing with DMF,

DCM and DMF. *S*-Trityl-3-mercaptopropionic acid (400 μmol) was used as a linker and was pre-activated with HBTU (380 μmol) and DIEA (800 μmol) in DMF (2.0 ml) for 30 sec, these activated solution was then poured into the reaction column containing the swelled resin and gently shaken for 20 min at room temperature. This protocol was repeated twice toward the resin. The resulting resin was then washed with DMF (2.0 ml) and DCM (2.0 ml) for five times and then treated twice with TFA/TIPS/H_2O (95:2.5:2.5) at room temperature for 1 min each to deprotect the trityl group. It was then followed by coupling of first amino acid, Boc-Leu(400 μmol), which was pre-activated by HBTU (380 μmol) and DIEA (800 μmol) in DMF (2.0 mL) for 1 min and was then poured into the column containing resin. The mixture was gently shaken for 20 min. Next, after washing the resin with DMF (2.0 mL) and DCM (2.0 mL) for several times. The resin was treated twice with TFA (3 mL) for 2 min and after the filtration of TFA, the resin was treated again with TFA (3 mL) for 5 min to completely removal of Boc protecting group. The resin was washed with DCM (2.0 mL) and DMF (2.0 mL) for several times to completely remove the TFA. Subsequently, peptide was elongated in the same protocol as above. After completion of all coupling steps, the side chain protecting groups was removed by treating with TFA:thioanisole:EDT:TfOH (20:2:1:1, 5 mL) in ice bath for 1 hour for twice. The resin was washed with TFA (2.0 ml), DCM (2.0 ml) and DMF (2.0 ml) thoroughly for several times to remove acidic cocktail. Then, the resultant peptide was directly detached from the resin with NHNH-Boc (2.5 mmol), N-methylimidazole (175 μmol), HOOBt (600 μmol), AgNO3 (100 μmol) in DMF (5.0 mL) for 4 hours at room temperature. The resulting peptide-NHNHBoc **6** was then purified by preparative RP-HPLC (Proteonavi, ø 10 mm×250 mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 75:25 to 10:90 over 60 min at 2.5 ml/min). Isolated yield: 31.0 mg, 30%.

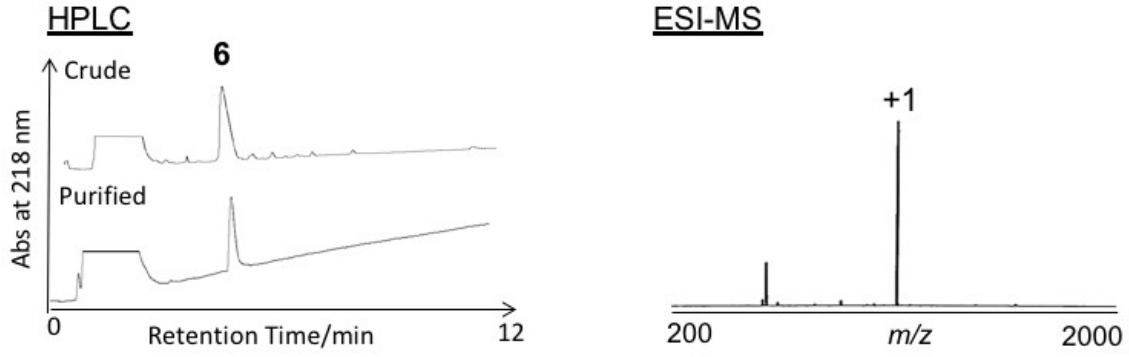


Figure S1: HPLC chromatogram and ESI-MS spectrum of **6**. ESI-MS: m/z calcd for $C_{46}H_{81}N_{11}O_{16}$: $[M+H]^+$ 1044.22, observed for 1044.68.

Synthesis of IFN-β-(80-88) Segment C: Fmoc-[(Asn)^{80}-(Glycan)-Glu^{81}–Thr^{82}–Ile^{83}– Val^{84}–Glu^{85}–Asn^{86}–Leu^{87}–Leu^{88}]-NHNHBoc 7

Fmoc-Asn-(oligosaccharide) **1** was ligated to peptide **6** in solution phase by condensation reaction. First, Fmoc-Asn-(oligosaccharide) **1** (13.4 mg, 4.79 μmol) was pre-activated with PyBOP (3.74 mg, 7.2 μmol) and DIEA (1.25 μL, 7.2 μmol) in 0.67 mL of DMF/DMSO (7:3) at -5 °C for 10 min. Then, the resulting mixture was added to DMF/DMSO (7:3, 0.67 mL) containing peptide **6** (10 mg, 9.6 μmol) and was then gently stirred to react at -5 °C for 1 hour. The product was precipitated in the reaction solution by the addition of precooled diethyl ether (10 times volumes, v/v) and was centrifuged at 4000 rpm for 15 min. This step was repeated twice to remove unreacted reagents. After discarded the diethyl ether from the centrifuge tube carefully, the resultant product **7** was directly used for the next reaction without RP-HPLC purification.

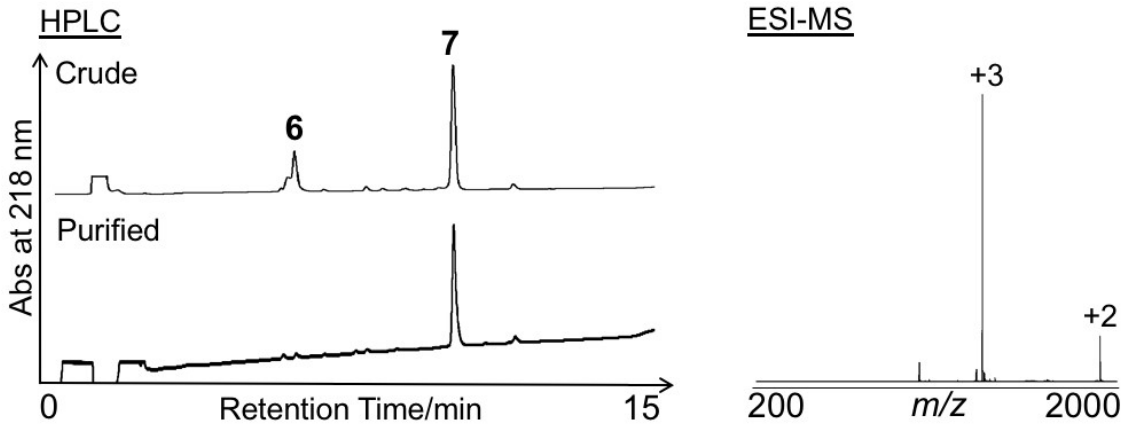


Figure S2: HPLC chromatogram and ESI-MS spectrum of **7**. ESI-MS: m/z calcd for $C_{165}H_{245}N_{19}O_{83}$: $[M+H]^+$ 3822.83, observed (deconvoluted): 3822.60; $[M+2H]^{2+}$ 1912.42, found for 1912.3; $[M+3H]^{3+}$ 1275.28, found for 1275.10.

Synthesis of IFN-β-(80-88) Segment C: H_2N-[Asn^{80}-(Glycan)-Glu^{81}–Thr^{82}–Ile^{83}–Val^{84}–Glu^{85}–Asn^{86}–Leu^{87}–Leu^{88}]-NHNHBoc 8

The Fmoc group at the N-terminal of glycopeptide **7** was deprotected with HOBt (1.3 mg, 9.6 μmol), 1-methylpyrrolidine (50 μL, 480 μmol), hexamethyleneimine (1.4 μL, 12 μmol) in DMF/DMSO solution (7:3, 1.78 mL) at room temperature for 1 hour.[2] After completion of reaction, the product was precipitated with 10 times (v/v) excess of precooled diethyl ether and was centrifuged at 4000 rpm for 15 min, this step was repeated twice to make sure unreacted

reagents was removed. After discarded the diethyl ether from the centrifuge tube carefully, the resulting mixture was then purified by preparative RP-HPLC (Proteonavi, Ø 10 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 65:35 to 30:70 over 60 min at 2.5 ml/min) to give peptide **8**. Isolated yield: 11.21 mg, 65% (over two steps).

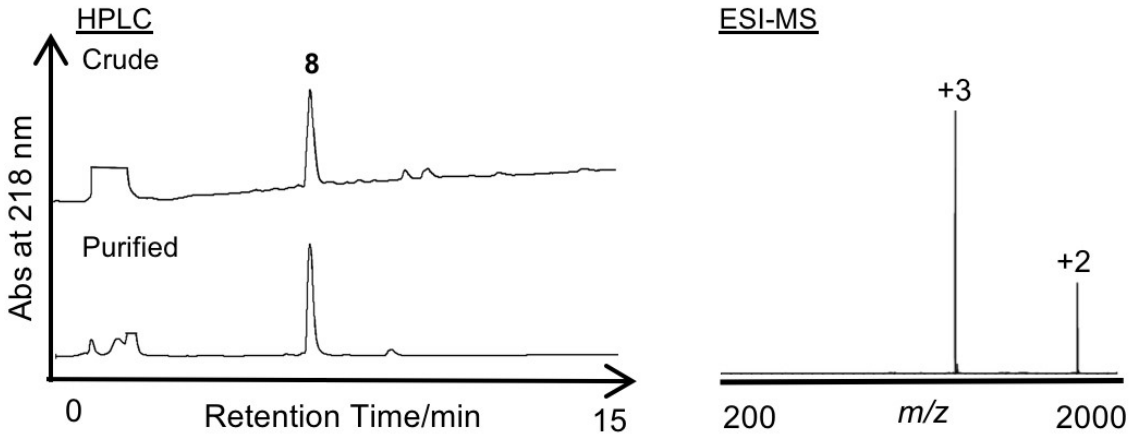


Figure S3: HPLC chromatogram and ESI-MS spectrum of **8**. ESI-MS: m/z calcd for $C_{150}H_{235}N_{19}O_{81}$: $[M+H]^+$ 3600.58, observed (deconvoluted): 3600.40; $[M+2H]^{2+}$ 1801.29, found for 1801.20; $[M+3H]^{3+}$ 1201.19, found for 1201.00.

Synthesis of IFN-β-(79-88) Segment C: H_2N-[Trp^{79}-Asn^{80}-(Glycan)-Glu^{81}–Thr^{82}–Ile^{83}–Val^{84}–Glu^{85}–Asn^{86}–Leu^{87}–Leu^{88}]-NHNHBoc 10

Fmoc-Trp(Boc)-OH (2.63 mg, 5 μmol) was pre-activated with PyBOP (1.95 mg, 3.75 μmol) and DIEA (0.65 μL, 3.75 μmol) in DMF/DMSO (7:3, 1.4 mL) at -5 °C for 5 min. Then, the reaction mixture was added to glycopeptide **8** (9.0 mg, 2.5 μmol) and was gently stirred to react at -5 °C for 1 hour. The product was precipitated by the addition of precooled diethyl ether (10 times volumes, v/v) and was centrifuged at 4000 rpm for 15 min, this step was repeated twice to make sure unreacted reagents was removed.. After discarded the diethyl ether, the resultant product (Fmoc-[Trp^{79}-Asn^{80}-(Glycan)-Glu^{81}–Thr^{82}–Ile^{83}–Val^{84}–Glu^{85}–Asn^{86}–Leu^{87}–Leu^{88}]-NHNHBoc **9**) was air-dried for 10 min and directly used for the next reaction without RP-HPLC purification. ESI-MS: m/z calcd for $C_{181}H_{263}N_{21}O_{86}$: $[M+H]^+$ 4109.16, observed (deconvoluted): 4109.35.

The Fmoc group at the N-terminal of glycopeptide **9** (2.5 μmol) was deprotected with HOBt (0.68 mg, 5 μmol), 1-methylpyrrolidine (26 μL, 250 μmol), hexamethyleneimine (0.74 μL, 6.25 μmol) in DMF/DMSO (7:3, 1.8 mL) at room temperature for 1 hour. The product was

then precipitated with 10 times excess of precooled diethyl ether, this step was repeated twice to make sure unreacted reagents was removed. After discarded the diethyl ether, the resultant product **10** will further purified by preparative RP-HPLC (Proteonavi, Ø 10 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 65:35 to 30:70 over 60 min at 2.5 ml/min). Isolated yield: 6.80 mg, 70% (over two steps).

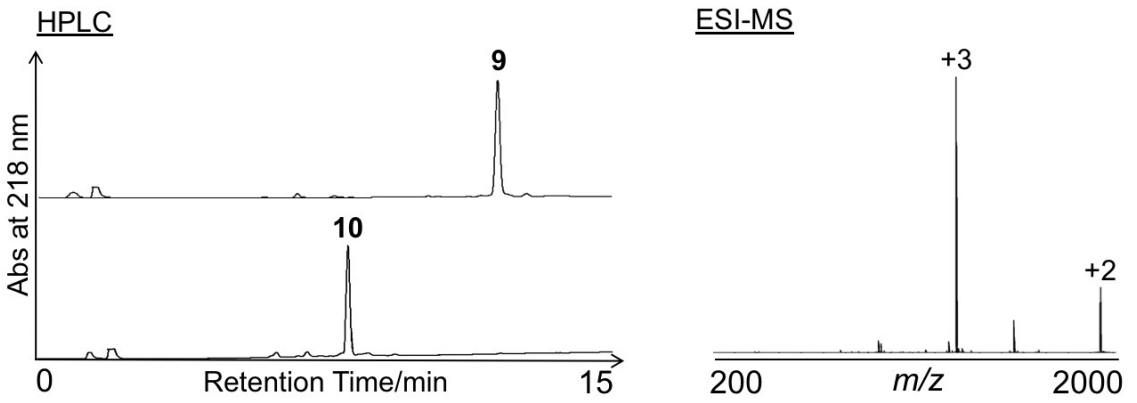


Figure S4: HPLC chromatogram of **9** and **10** and ESI-MS spectrum of **10**. ESI-MS: m/z calcd for $C_{166}H_{253}N_{21}O_{84}$: $[M+H]^+$ 3886.91, observed (deconvoluted): 3886.62; $[M+2H]^{2+}$ 1944.46, found for 1944.30; $[M+3H]^{3+}$ 1296.64, found for 1296.54.

Section 2: Synthesis of IFN-β-(31-78) Segment B 15

Synthesis of IFN-β-(31-78) Segment B 15

Recombinant expression of His-SUMO-Segment B-Cys by jar fermenter

Plasmid was ordered from Thermo Fischer Scientific and was shown in S23 page in detail. His-SUMO-Segment B-Cys production was carried out in a 3L bioreactor (Bioneer-Neo, Marubishi Co., Ltd., Tokyo, Japan) with a working volume of 2 L of production medium. First, the feeding media containing glycerol (300 g), Adekanol LG295S (0.15 g), bacto yeast extract (45 g) and bacto tryptone (84 g) in 255 mL of Milli-Q water was prepared. Then, the starter culture was prepared by adding with 0.1 mL of glycerol stock and 0.1 mL of ampicillin (0.5 g/mL) to 100 mL of LB media and stirred at 160 rpm in incubator, 25°C for 16 hours. This starter culture (100 mL) was then used to incubate LB media (2.0 L) and added with ampicillin (0.5 g/mL, 2.0 mL). The resulting cultivation (pH = 7.0) was carried out under 37°C and agitation rate at 400 rpm. The cultivation solution was allowed to cool to 25°C and then feeding (less than 0.28 g/min) was started when oxygen concentration of the system decrease, the sudden decrease of the oxygen indicate cell death due to lack of nutrient availability. When the optical density (OD) at 600 nm reached 0.9, temperature of culture system was adjusted to 37°C and induced with isopropylthiogalactoside (IPTG,1.0 M solution, 1.5 mL) for 3 hours. *E.coli* were collected by

centrifugation at 8000 rpm for 10 min at 4°C. The cells (pellets) were dissolved in appropriate amount of lysis buffer on ice bath. The composition of lysis buffer are 50 mM of sodium phosphate, pH 7.8, 400 mM NaCl, 100 mM KCl, 10% glycerol, 0.1% triton X-100, 10 mM imidazole, which referred to the protocol provided by Invitrogen. After cell lysis by ultrasonication for 1 min (every 2 sec with 1 sec pauses) on ice bath for 4 times, it was subjected to centrifugation at 10000 rpm for 10 min at 4°C. His-SUMO-Segment B-Cys remained in the supernatant while the pellet containing cell debris was discarded. Purification of His-SUMO-Segment B-Cys was performed using cOmplete™ His-tag resin under native conditions using varying compositions of Buffer A (50 mM NaH_2PO_4, pH 8, 300 mM NaCl) and then Buffer B (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 250 mM imidazole). The eluent containing His-SUMO-Segment B-Cys was confirmed with LC-MS analysis and dialyzed against water for 12 hours. The resulting product was lyophilized and used to the next reaction without RP-HPLC purification. However, in order to confirm the yield of His-SUMO-Segment B-Cys, one of the batch was purified by RP-HPLC. (Proteonavi, Ø 10 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 55:45 to 20:80 over 40 min at 2.5 ml/min). Yield: 135 mg/L (2 steps, dialysis purification and subsequent RP-HPLC purification).

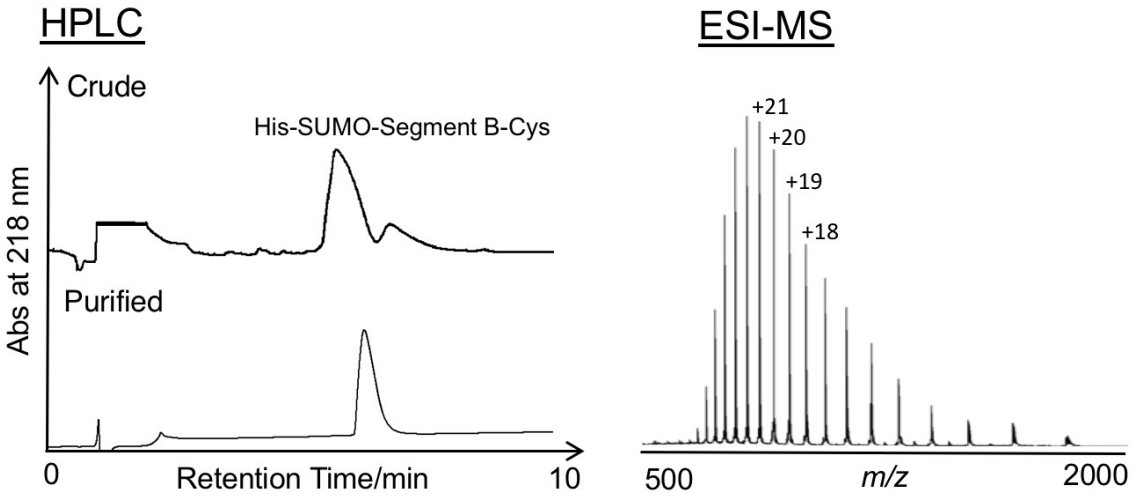


Figure S5: HPLC chromatogram and ESI-MS spectrum of His-SUMO-Segment B-Cys. ESI-MS: m/z calcd for $[M+H]^+$ 18053.04, observed (deconvoluted): 18052.31; $[M+18H]^{18+}$ 1003.95, found for 1003.90; $[M+19H]^{19+}$ 951.16, found for 951.12; $[M+20H]^{20+}$ 903.65, found for 903.61; $[M+21H]^{21+}$ 860.67, found for 860.63.

Synthesis of Thz-Segment B-Cys 11 by SUMO protease cleavage and thiazolidine protection of His-SUMO-Segment B-Cys

Plasmid was ordered from Thermo Fischer Scientific and was shown in S23 page in detail.

SUMO protease cleavage and thiazolidine protection were performed in one pot. The lyophilized His-SUMO-Segment B-Cys (110 mg, 6.09 μmol) was dissolved in Tris buffer (20 mM, 20 mL, pH 8) containing of SUMO protease (0.1 units/μg), formaldehyde (0.9 mM), DTT (15 mM) at 30°C for 4 hours. The reaction was monitored by RP-HPLC and was quenched by adding with GnHCl (11.46 g, 0.12 mol) in the reaction mixture. The resulting product, Thz-Segment B-Cys **11** was dialyzed against water for 12 hours, followed by lyophilized and then used to the next reaction without RP-HPLC purification. Purification was performed with RP-HPLC. (Proteonavi, Ø 10 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 55:45 to 20:80 over 30 min at 2.5 ml/min). This protocol gave product **11** in 22.8 mg, 65% (average isolated yield).

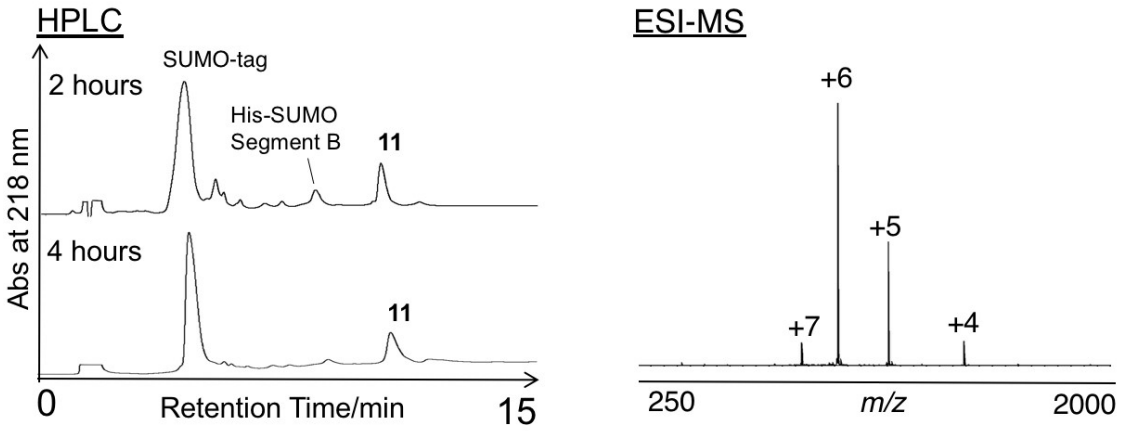


Figure S6: HPLC chromatogram and ESI-MS spectrum of **11**. ESI-MS: m/z cald for $C_{253}H_{396}N_{66}O_{80}S_4$: $[M+H]^+$ 5770.57, observed (deconvoluted): 5769.72; $[M+4H]^{4+}$ 1443.64, found for 1443.34; $[M+5H]^{5+}$ 1155.11, found for 1154.87; $[M+6H]^{6+}$ 962.76, found for 962.56.

Cyanylation of Thz-Segment B-Cys(SCN) 12

Thz-Segment B-Cys **11** (65 mg, 11.3 μmol) was dissolved in Tris buffer solution (20 mM, 57 mL) containing 6 M GnHCl, TCEP (45 mg, 158 μmol) and NTCB (38 mg, 170 μmol) and adjusted the pH to pH 8.3 with triethylamine (TEA). The reaction mixture was stirred for 1 hour at room temperature and was then subjected to purify by RP-HPLC (Proteonavi, Ø 10 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 60:40 to 25:75 over 40 min at 2.5 ml/min) to give peptide **12**. Isolated yield: 43.5 mg, 66 %.

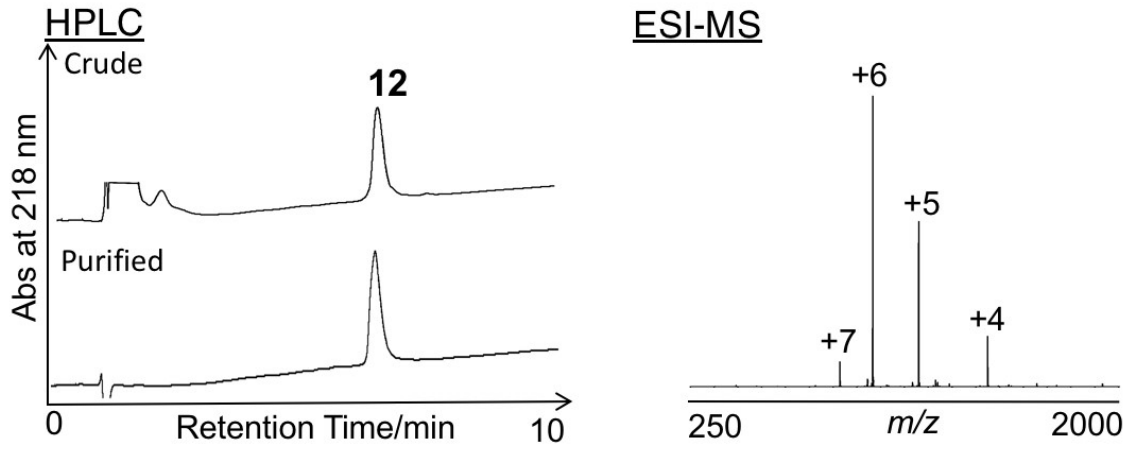


Figure S7: HPLC chromatogram and ESI-MS spectrum of **12**. ESI-MS: m/z calcd for $C_{254}H_{395}N_{67}O_{80}S_4$: $[M+H]^+$ 5795.58, observed (deconvoluted): 5795.08; $[M+4H]^{4+}$ 1449.90, found for 1449.72; $[M+5H]^{5+}$ 1160.12, found for 1159.98; $[M+6H]^{6+}$ 966.85, found for 966.82; $[M+7H]^{7+}$ 828.94, found for 828.84.

Hydrazinolysis: Thz-Segment B-$NHNH_2$ 13

Thz-Segment B-Cys(SCN) **12** (32 mg, 5.53 μmol) was added into pre-prepared Tris buffer solution (20 mM, 48 mL, pH 8.9) containing 6 M GnHCl and 2.5% v/v of hydrazine monohydrate at 30 °C to react for 3 hours. The reaction was completed as judged by LC-MS and was purified by RP-HPLC (Proteonavi, Ø 10 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 70:30 to 20:80 over 30 min at 2.5 ml/min) to give peptide-$NHNH_2$ **13**. Isolated yield: 19.6 mg, 63 %.

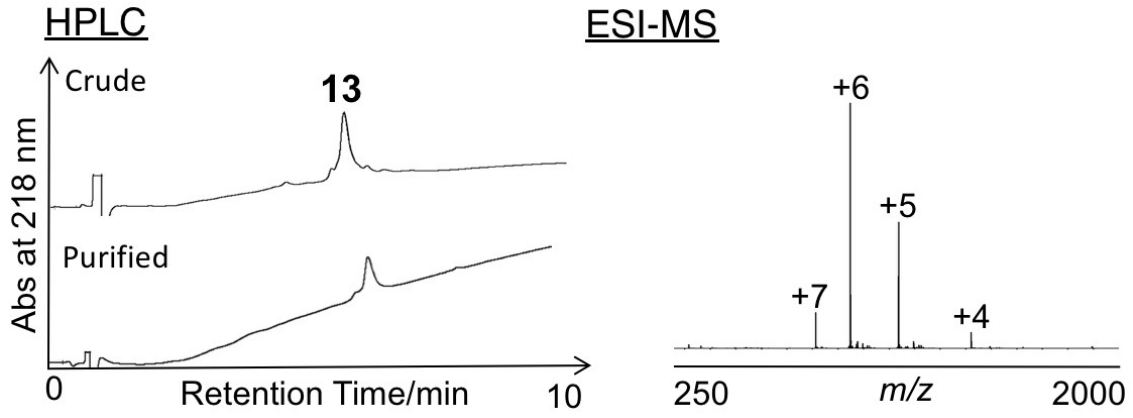


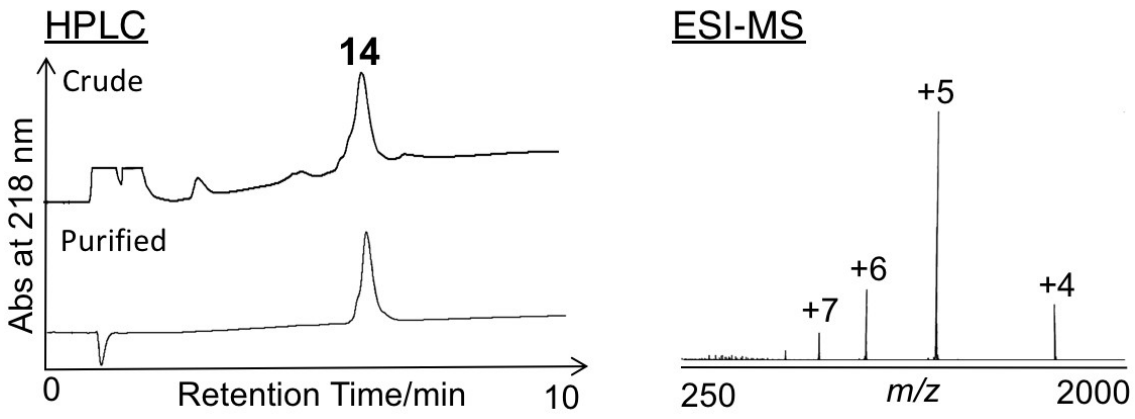
Figure S8: HPLC chromatogram and ESI-MS spectrum of **13**. ESI-MS: m/z calcd for $C_{250}H_{393}N_{67}O_{78}S_3$: $[M+H]^+$ 5681.47, observed (deconvoluted): 5681.10; $[M+4H]^{4+}$ 1421.37, found for 1421.21; $[M+5H]^{5+}$ 1137.30, found for 1137.17; $[M+6H]^{6+}$ 947.91, found for 947.81;

$[M+7H]^{7+}$ 812.64, found for 812.55.

Thioesterification and Boc protection of Thz-Segment B(Boc)-thioester 15

The thioesterification of peptide-hydrazide with Thz group at the N-terminal was performed with reported manner.[3] Thz-Segment B-$NHNH_2$ **13** (10 mg, 1.72 μmol) was dissolved in an acidic cocktail solution containing TFA, TIPS, H_2O, m-cresol, thioanisole (80:2.5:2.5:5:10, 1.0 mL) at -30°C. To this solution, it was added with 10% w/v $NaNO_2$ (7 μL) and the mixture was left for 20 minutes at -30°C. After the reaction completed, the resultant peptide-azide was then precipitated by addition of 10 times excess (10 mL) of precooled diethyl ether. After centrifuged at 4000 rpm for 15 min at 4°C, the diethyl ether was decanted carefully and peptide was air dried for about 10 min. Then, the peptide was dissolved in a sodium phosphate buffer (0.1M, pH 7.3, 1.2 mL) containing 2.5% (w/v) of MESNA and 6M GnHCl at room temperature for 1 hour. The product was purified by preparative RP-HPLC. (Proteonavi, Ø 10 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 55:45 to 25:75 over 40 min at 2.5 ml/min). Isolated yield: 7.27 mg, 73%.

In order to perform thioester ligation, the Lys of IFN-β-(31-78) Segment B **14** was protected with the Boc group to produce segment B **15**. Segment B **14** (10 mg, 1.73 μmol) was dissolved in DMF (0.87 mL) containing Boc-OSu (6.7 mg, 31.1 μmol) and DIEA (16 μL, 93.3 μmol) and the mixture was reacted for 1 hour. The resultant product was precipitated by the addition of the precooled diethyl ether (8.7 mL) to remove the excess Boc-OSu and DIEA and centrifuged at 4000 rpm for 15 min. After discarded diethyl ether from the centrifuge tube carefully, peptide **15** was air-dried for 10 min (7.3 mg, 73%) and proceed to the next reaction.



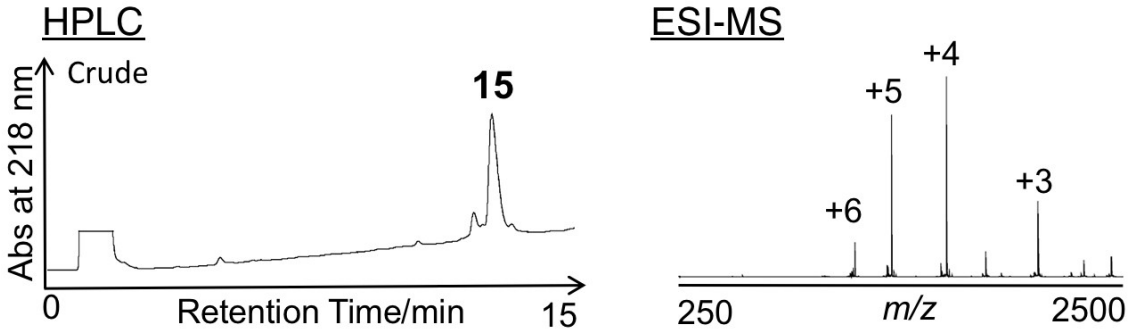


Figure S9: HPLC chromatogram and ESI-MS spectrum of **14** (top) and **15** (bottom) and corresponding mass spectram. ESI-MS of **14** (top) : m/z calcd for $C_{252}H_{394}N_{65}O_{81}S_5$: $[M+H]^+$ 5790.60, observed (deconvoluted): 5790.19. $[M+4H]^{4+}$ 1448.65, found for 1448.55; $[M+5H]^{5+}$ 1159.12, found for 1159.04; $[M+6H]^{6+}$ 966.10, found for 966.03. ESI-MS of **15** (bottom): m/z calcd for $C_{267}H_{418}N_{65}O_{87}S_5$: $[M+H]^+$ 6090.95, observed (deconvoluted): 6090.80; $[M+3H]^{3+}$ 2031.32, found for 2031.60; $[M+4H]^{4+}$ 1523.74, found for 1523.70; $[M+5H]^{5+}$ 1219.19, found for 1219.16; $[M+6H]^{6+}$ 1016.16, found for 1016.13.

Section 3: Synthesis of IFN-β-(1-30)-Segment A 23

Synthesis of IFN-β-(1-16)-Segment A-thioester: H_2N-[Met[1]–Ser–Tyr-Asn-Leu-Leu-Gly-Phe-Leu-Gln-Arg-Ser-Ser-Asn-Phe-Gln[16]]-SR 23(i)

IFN-β-(1-16)-Segment A-thioester **23(i)** was synthesized on Aminomethyl ChemMatrix® resin (100 μmol) by using Boc SPPS conditions which was described in the preparation of IFN-β-(81-88)-Segment C-NHNHBoc **1**. The first amino acid, Boc-Gln (400 μmol) was pre-activated by HBTU (380 μmol) and DIEA (800 μmol) in DMF (2 ml) for 1 min and was then poured into the reaction column containing resin to react for 20 min. After washing the resin with DMF (2 mL) and DCM (2 mL) for five times, the resin was treated twice with TFA (3 mL) for 2 min and after the filtration of TFA, the resin was treated with TFA again for 5 min to completely removal of Boc protecting group. The resin was then washed with DCM (2 mL) and DMF (2 mL) for five times to completely remove the TFA. The subsequent peptide was elongated in the same method as first amino acid. After completion of all coupling steps, the side chain protecting groups was removed by treating with TFA:Thionisole:EDT:TfOH (20:2:1:1, 5 mL) in ice bath for 1 hours (twice). The resin was washed with TFA (2.0 ml), DCM (2.0 ml) and DMF (2.0 ml) thoroughly for several times to remove acidic cocktail. Then, the peptide-thioester was directly detached from the resin by adding a sodium phosphate buffer solution (0.2 M at pH 6.2, 5 mL) containing 5% (w/v) of MESNA, and 6M Gn-HCl for 24 hours. The

target peptide-thioester, H_2N-[Met[1]–Ser–Tyr-Asn-Leu-Leu-Gly-Phe-Leu-Gln-Arg-Ser-Ser-Asn-Phe-Gln[16]]-SR **23(i)** resulting in the mixture was then purified by preparative RP-HPLC (Proteonavi, Ø 10 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 70:30 to 20:80 over 45 min at 2.5 ml/min). Isolated yields: 24.34 mg, 12%.

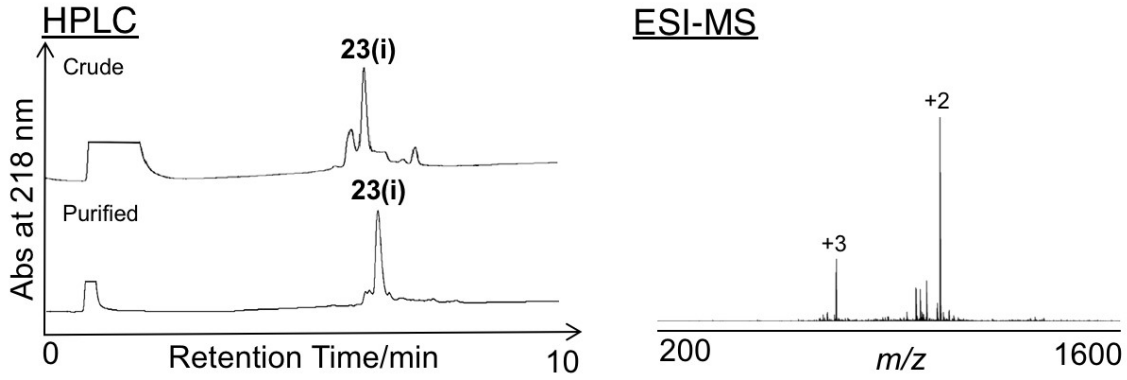


Figure S10: The HPLC profile and ESI-MS spectra of **23(i)**. ESI-MS: m/z calcd for $C_{87}H_{132}N_{23}O_{27}S_3$: $[M+H]^+$ 2028.33, observed (deconvoluted): 2028.71; $[M+2H]^{2+}$ 1015.17, found for 1015.50; $[M+3H]^{3+}$ 677.11, found for 677.25.

Synthesis of IFN-β-(17-30)-Segment A-hydrazide: H_2N-[Cys[17]-Gln-Lys-Leu-Leu-Trp-Gln-Leu-Asn-Gly-Arg-Leu-Glu-Tyr[30]]-$NHNH_2$ 23(ii)

IFN-β-(17-30)-Segment A-hydrazide **23(ii)** was synthesized by the same manner in the preparation of synthesis of IFN-β-(81-88)-Segment C-NHNHBoc **1** and synthesis of IFN-β-(1-16)-Segment A-thioester **23(i)** except for the final hydrazinolysis condition. After the completion of all amino acid (400 μmol) coupling steps, the side chain protecting groups was deprotected by treating with DMS : m-cresol : TFA: TfOH: EDT (3:1:5:1:0.5, 5 mL) and followed by TFA : Thionisole : EDT : TfOH (20:2:1:1, 5 mL) in ice bath, each of the acidic cocktail treatment was carried out for 2 hours. The resulting peptide on resins were washed with DCM (2 mL) and DMF (2 mL) for five time. Then, the peptide was detached from the resin by adding a sodium phosphate buffer solution (0.2M, pH 9.0, 5 mL) containing 6M GnHCl and hydrazine monohydrate (2% v/v) and then the mixture was left at room temperature for 2 hours. After completion of reaction, N-terminal thiazolidine (Thz) of the peptide was converted to Cys by adding with O-methylhydroxylamine hydrochloride (200 mM) and pH was adjusted to pH 4. After 4 hours, the reaction mixture was purified by preparative RP-HPLC (Proteonavi, Ø 10 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 70:30 to 30:70 over 40 min at 2.5 ml/min) to give H_2N-[Cys[17]-Gln-Lys-Leu-Leu-Trp-Gln-Leu-Asn-Gly-Arg-Leu-Glu-Tyr[30]]-$NHNH_2$ **23(ii)**. Isolated yields: 16.70 mg, 9.4%.

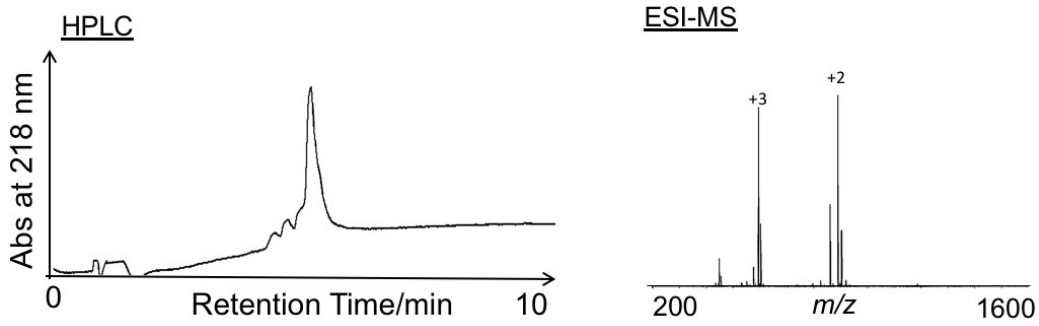


Figure S11: The HPLC profile and ESI-MS spectra of IFN-β-(17-30)-Segment A-hydrazide **23**. ESI-MS: m/z cald for $C_{80}H_{128}N_{24}O_{20}S$: $[M+H]^+$ 1778.11, observed (deconvoluted): 1778.06; $[M+2H]^{2+}$ 890.06, found for 890.03; $[M+3H]^{3+}$ 593.70, found for 594.12.

Native chemical ligation to yield IFN-β-(1-30) Segment A-hydrazide 23

IFN-β-(1-16)-Segment A-thioester **23(i)** (19 mg, 9.4 µmol) and IFN-β-(17-30)-Segment A-hydrazide **23(ii)** (16.7 mg, 9.4 µmol) were dissolved in a freshly prepared sodium phosphate buffer (200 mM, pH 6.5, 4.7 mL) containing 6M Gn-HCl,100 mM MPAA (79 mg, 470 µmol) and 100 mM TCEP (117.6 mg, 470 µmol), which were then stirred at room temperature for 4 hours. After completion of the reaction, the resultant product, Segment A-hydrazide **23,** was purified by preparative RP-HPLC (Proteonavi, Ø 10 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 70:30 to 30:70 over 30 min at 2.5 ml/min). Isolated yields: 8.60 mg, 25%.

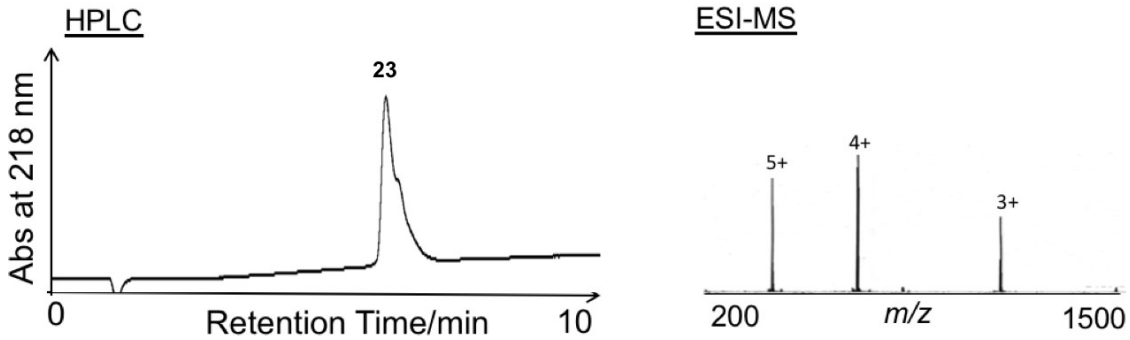


Figure S12: The HPLC profile and ESI-MS spectra of IFN-β-(1-30) Segment A-hydrazide **23**. ESI-MS: m/z calcd for $C_{165}H_{255}N_{47}O_{44}S_2$: $[M+H]^+$ 3665.26, observed (deconvoluted): 3665.03; $[M+3H]^{3+}$ 1222.75, found for 1222.61; $[M+4H]^{4+}$ 917.32, found for 917.21; $[M+5H]^{5+}$ 734.05, found for 733.97.

Section 4: Synthesis of IFN-β-(89-166) Segment D 19

Recombinant expression of His-SUMO-Segment D 16

His-SUMO-Segment D **16** was overexpressed in *E.coli* BL21(DE3) strain in LB media at 37°C for 12 hours and the enhancement of expression was evaluated after the induction with isopropylthiogalactoside (IPTG) (50 mM) at 37°C for 6 hours. The cells were collected by centrifugation at 8000 rpm for 10 min at 4°C and were resuspended in lysis buffer (compositions as mentioned in the expression of His-SUMO-Segment B) and lysed by ultrasonicator for 1 min (every 2 sec with 1 sec pauses) on ice bath for 4 times. The lysed cells were subjected to centrifugation at 10000 rpm for 10 min at 4°C. His-SUMO-Segment D **16** remained in the supernatant was purified by cOmplete^{TM} His-tag resin under native conditions using varying compositions of Buffer A (50 mM NaH_2PO_4, pH 8, 300 mM NaCl) and then Buffer B (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, 500 mM imidazole). The eluent containing His-SUMO-Segment D **16** was confirmed with LC-MS analysis and was directly purify by HPLC as the peptide prone to aggregation when dialysed against water (Proteonavi, Ø 10 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 70:30 to 20:80 over 40 min at 2.5 ml/min). Average isolated yield: 13mg/L.

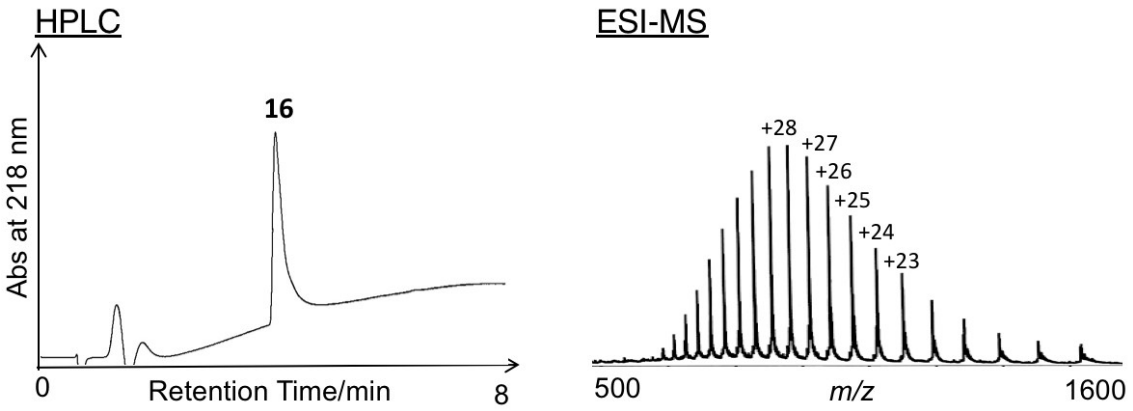


Figure S13: HPLC chromatogram and ESI-MS spectrum of **16**. ESI-MS: m/z cald for $C_{972}H_{1528}N_{280}O_{281}S_8$: $[M+H]^+$ 21875.16, observed (deconvoluted): 21872.55; $[M+27H]^{27+}$ 811.19, found for 811.09.

Synthesis of Thz-Segment D 17 by SUMO protease cleavage and thiazolidine protection

His-SUMO-Segment D **16** (15 mg, 0.69 µmol) was dissolved in distilled water (13.5 mL) and adding with the Arg solution (42.66 mg, 15 mM) to assist the solubility of the hydrophobic His-SUMO-Segment D **16**. The pH was adjusted to pH 5.7-6.0 and SUMO protease was added (70 µL, 0.1 units/µg) followed by addition of 0.1 M formaldehyde (10 µL) which was incubated at 30°C for 18h. The desired product formed precipitated in the reaction mixture. To quench the reaction and also dissolve the peptides, GnHCl (7.74 g, 0.081 mol) was added followed by

purification using RP-HPLC (Proteonavi, Ø 10 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 65:35 to 20:80 over 40 min at 2.5 ml/min) to give peptide **17**. Isolated yield: 2.84 mg, 43%.

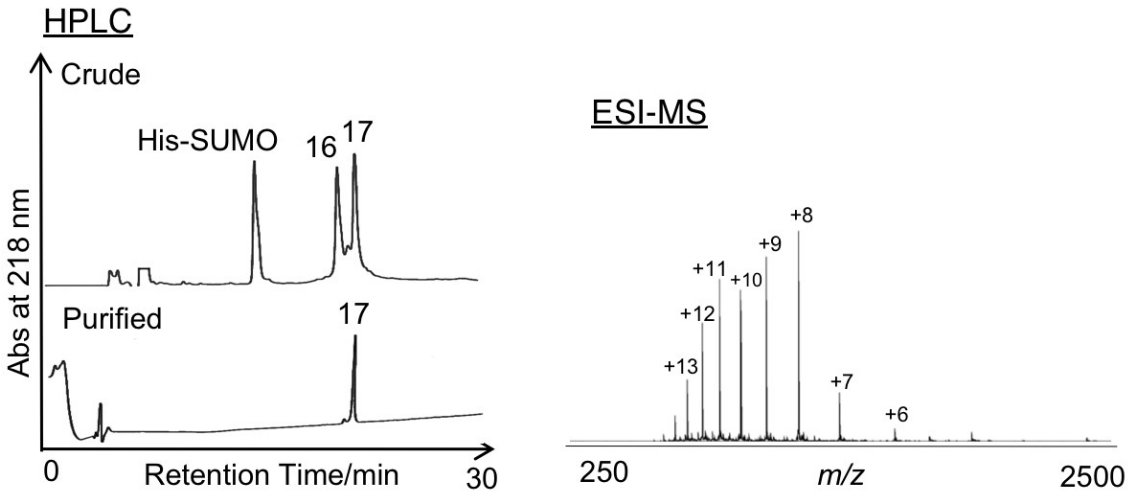


Figure S14: HPLC chromatogram and ESI-MS spectrum of **17**. ESI-MS: m/z calcd for $C_{439}H_{683}N_{123}O_{113}S_3$: $[M+H]^+$ 9588.22, observed (deconvoluted): 9588.08; $[M+6H]^{6+}$ 1599.04, found for 1598.83; $[M+7H]^{7+}$ 1370.75, found for 1370.71; $[M+8H]^{8+}$ 1199.53, found for 1199.37; $[M+9H]^{9+}$ 1066.36, found for 1066.22; $[M+11H]^{11+}$ 872.66, found for 872.54.

Synthesis of Segment D 18 and 19 by phenacyl bromide protection and ring opening of Segment D 17

Thz-Segment D **17** (1.9 mg, 0.20 μmol) was dissolved in 100 μL of distilled water and added with 0.2 M of Pac-Br (1.6 μL, 0.32 μmol) to react for 30 minutes. The pH of solution in this reaction was to be around neutral. After the formation of Pac-protected peptide **18**, 40% acetonitrile (100 μL) containing with 0.1% (v/v) TFA and dipyridyldisulfide (DPDS, 2 mg, 9.08 μmol) was added to the reaction mixture. The mixture was then gently stirred at 37 °C for 2-3 hours to open the Thz. After completion of reaction, the resultant peptide was purified by preparative RP- HPLC (Proteonavi, Ø 4.6 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 65:35 to 15:85 over 30 min at 1 ml/min) to give peptide **19**. Isolated yield: 1.10 mg, 56%.

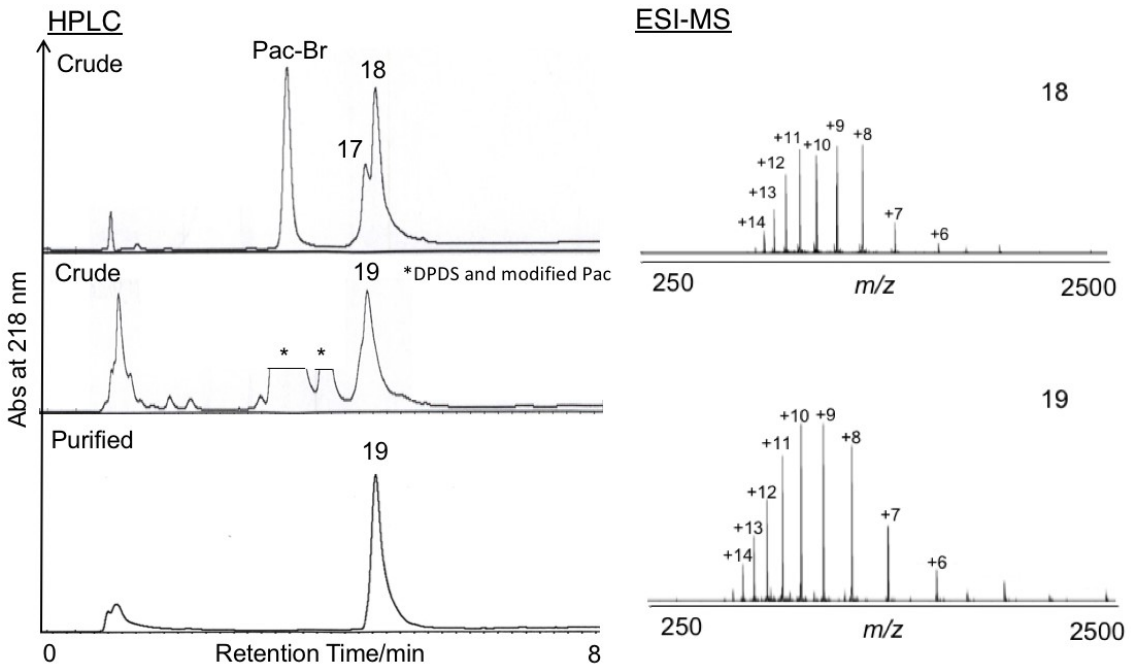


Figure S15: HPLC chromatogram and ESI-MS spectrum of Pac-protected Thz-Segement-D **18** and Cys-Segment-D **19**. ESI-MS: m/z calcd for **18**; $C_{447}H_{689}N_{123}O_{114}S_3$: $[M+H]^+$ 9706.36, observed (deconvoluted): 9706.07; $[M+6H]^{6+}$ 1618.73, found for 1618.51; $[M+8H]^{8+}$ 1214.30, found for 1214.14; $[M+11H]^{11+}$ 883.40, found for 883.37.

ESI-MS: m/z calcd for **19;** $C_{451}H_{692}N_{124}O_{114}S_4$: $[M+H]^+$ 9803.49, observed (deconvoluted): 9803.00; $[M+6H]^{6+}$ 1634.92, found for 1634.82; $[M+7H]^{7+}$ 1401.50, found for 1370.71; $[M+8H]^{8+}$ 1226.44, found for 1199.37; $[M+9H]^{9+}$ 1090.28, found for 1090.21; $[M+10H]^{10+}$ 981.35, found for 981.29.

Section 5: Peptide ligation to yield full length glycosyl-interferon β polypeptide

First ligation: Thioester ligation between IFN-β-(31-78)-Segment B 15 and IFN-β-(79-88)-glycopeptide Segment C 10

For thioester ligation, IFN-β-(79-88)-glycopeptide Segment C **10** (6.2 mg, 1.64 µmol) was dissolved in DMF (0.82 mL) containing $AgNO_3$ (0.84 mg, 4.92 µmol), N-methylimidazole (1.35 µL, 16.4 µmol) and HOOBt (6.42 mg, 39.4 µmol).[3] This reactant solution was added to the Boc protected Segment B **15**. The resultant mixture was stirred at room temperature for 2 hours. After completion of the reaction, the product Segment-BC **20** was precipitated with precooled diethyl ether (8.2 mL) and was centrifuged at 4000 rpm for 15 min. After the supernatant was discarded, the resultant pellet was added with a solution [0.1% TFA in water : 0.1% TFA in 90% MeCN (3:7), 2 ml] containing dipyridyldisulfide (DPDS, 20 mM) and the

mixture was gently stirred at 37 °C for 2 hours to convert the N-terminal Thz to Cys residue.[4] After completion of reaction, the resultant peptide Segment-BC **21** was purified by preparative RP- HPLC (Proteonavi, Ø 4.6 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 60:40 to 0:100 over 30 min at 1 ml/min). Isolated yield: 4.07 mg, 25%.

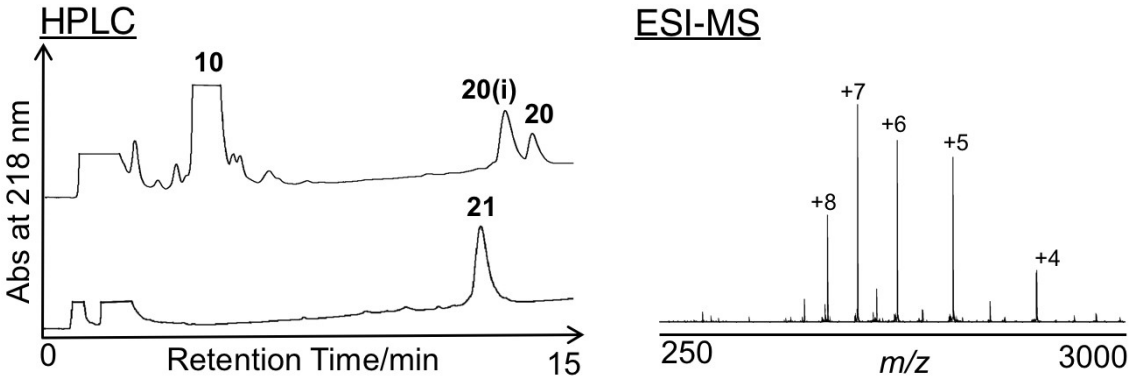


Figure S16: HPLC chromatogram of Thz opening reaction of Segment BC **20.** Thz-Segment BC **20** and PyS-Cys-Segment-BC **21** and ESI-MS spectrum of **21**. Compound **20(i)** corresponds to the Cys-Segment-BC which was derived from **20** by partially Thz open reaction. ESI-MS: m/z calcd for $C_{435}H_{669}N_{87}O_{168}S_4$: $[M+H]^+$ 9933.82, observed (deconvoluted): 9933.00; $[M+4H]^{4+}$ 2484.46, found for 2484.25; $[M+5H]^{5+}$ 1987.76, found for 1987.60; $[M+6H]^{6+}$ 1656.64, found for 1656.49; $[M+7H]^{7+}$ 1420.12, found for 1419.85; $[M+8H]^{8+}$ 1242.73, found for 1242.61.

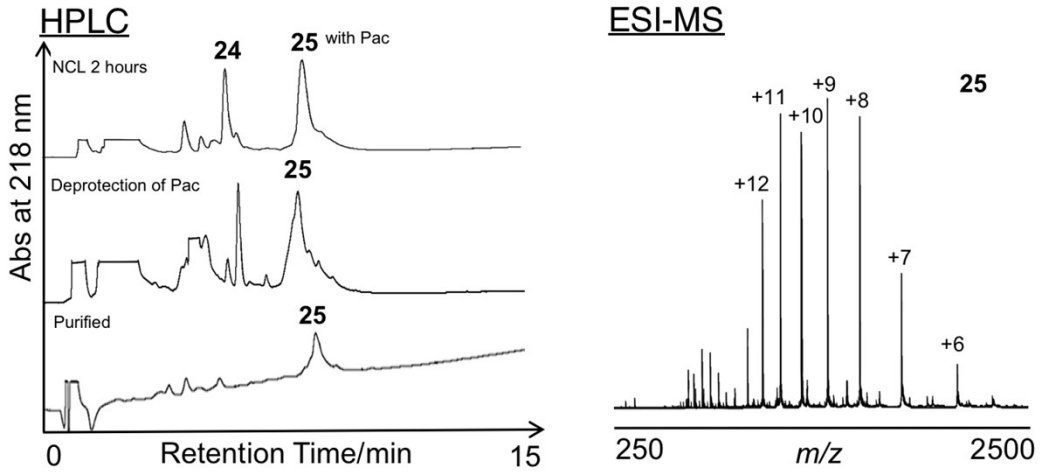
Second ligation: Native chemical ligation between IFN-β-(1-30)-Segment A 24 and IFN-β-(31-88)-glycopeptide Segment BC 22

Before native chemical ligation with IFN-β-(1-30)-Segment A **23**, Boc deprotection of IFN-β-(31-88)-segment BC **21** (2.0 mg, 0.204 μmol) was carried out by treatment with TFA:TIPS: H_2O (95 : 2.5 : 2.5, 0.2 mL). The peptide was precipitated with the precooled diethyl ether (2.0 mL) and then the solution was centrifuged at 4000 rpm for 15 min. After discarded of diethyl ether from the centrifuge tube carefully, the resultant segment BC **22** was air-dried for 10 min and proceed to the next reaction without RP-HPLC purification.

To use IFN-β-(1-30)-Segment A-α-hydrazide **23** for NCL, the hydrazide form was converted into a thioester. For this reaction, IFN-β-(1-30)-Segment A-α-hydrazide **23** (1.5 mg, 0.410 μmol) was dissolved in the degassed sodium phosphate (200 mM, pH 3.5, 50 μL) solution containing 6M GnHCl and was allowed to stand in an ice bath. To this reaction mixture was added $NaNO_2$ (300 mM, 5 μL) and the mixture was left in an ice bath for 30 min. After completion of reaction, to this reaction mixture was quickly added another degassed sodium

phosphate (200 mM, pH 3.5, 5.0 µL) solution containing 6M GnHCl, MPAA (300 mM), TCEP (300 mM) and then pH of the solution was adjusted to pH 6.5. The resulting solution was stirred at room temperature for 30 min to yield IFN-β-(1-30)-Segment A-S-C_6H_5-CH_2-COOH **24**. This thioester was used to the next NCL without purification.

For NCL, to the resultant Boc deprotected Segment BC **22** was added with the above sodium phosphate buffer solution (200 mM, pH 6.5, 5.0 µL) containing 6M GnHCl, MPAA (300 mM), TCEP (300 mM) and IFN-β-(1-30)-Segment A-S-C6H5-CH_2-COOH **24**. The mixture was incubated at room temperature for 5 hours. After completion of NCL, the solution was adjusted to pH 9.5 by the addition of NaOH (5.0 M) solution dropwisely and left the solution for 2 hours to fully deprotect the phenacyl group on sialic acid. This resultant mixture was treated with 100 mM of TCEP and then the product **25** was purified by preparative RP-HPLC (Proteonavi, Ø 4.6 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 60:40 to 0:100 over 30 min at 1 ml/min). Isolated yield: 1.5 mg, 58%.



F**igure S17:** HPLC chromatogram and ESI-MS spectrum of **25**. HRMS (ESI): m/z calcd for $C_{554}H_{865}N_{131}O_{200}S_5$: $[M+H]^+$, 12721.0310, observed (deconvoluted): 12721.0048. $[M+6H]^{6+}$ 2121.1718 found for 2121.1708; $[M+7H]^{7+}$ 1818.2901, found for 1818.2934; $[M+8H]^{8+}$ 1591.1289, found for 1591.1265; $[M+9H]^{9+}$ 1414.4479, found for 1414.4444; $[M+11H]^{11+}$ 1157.4574, found for 1157.4529.

Synthesis of IFN-β-(1-88)-Segment ABC 27 by one pot Cys-thiol protection and thioesterification

IFN-β-(1-88)-Segment ABC-hydrazide **25** (1.5 mg, 0.118 µmol) was dissolved in the degassed sodium phosphate buffer solution (200 mM, pH 6.7, 150 µL) containing 6M GnHCl. To this buffer solution was added 0.1 M phenacyl-bromide in DMF (1.8 µL) and the resulting solution was stirred for 1 hour at room temperature.[5] After the completion of phenacyl protection of

cysteines as monitored by LC-MS (ESI-MS: m/z calcd for $C_{570}H_{877}N_{131}O_{202}S_5$: $[M+H]^+$ 12957.301, observed (deconvoluted): 12956.84), the resulting solution (150 µL) was added to the freshly prepared degassed sodium phosphate buffer solution (200 mM, pH 3.5, 600 µL) containing 6 M GnHCl. This solution was allowed to pre-cool under -20 °C for 10 min and added with $NaNO_2$ solution (600 mM, 75 µL). The reaction was carried out for 30 min at -20 °C. Next, thioesterification was carried out by adding with the freshly prepared degassed sodium phosphate buffer solution (200 mM, pH 3.5, 750 µL) containing 6 M GnHCl and MPAA (300 mM) at pH 6.5 for 30 minutes at room temperature. The resultant solution was incubated for 30 minutes at room temperature. After the completion of reaction as judged by LC-MS, the resultant mixture was treated with TCEP (200 mM, 1.5 mL). Purified by preparative RP- HPLC (Proteonavi, Ø 4.6 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 55:45 to 15:85 over 30 min at 1 ml/min) to give **27**. Isolated yield: 0.82 mg, 53% (two steps).

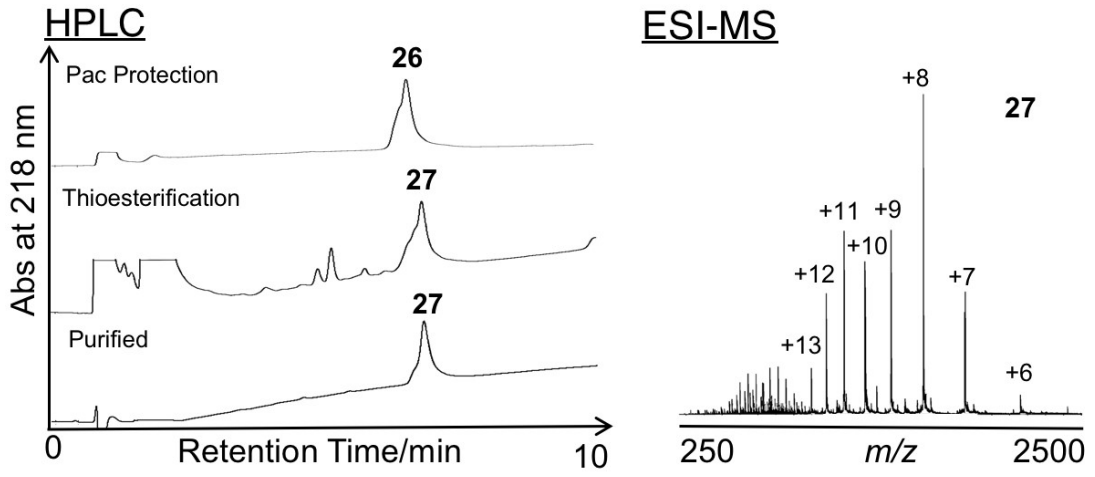


Figure S18: HPLC chromatogram of **26** and **27** as well as the ESI-MS spectrum of **27**. HRMS (ESI): m/z calcd for $C_{578}H_{881}N_{129}O_{204}S_6$: $[M+H]^+$ 13093.4650, observed (deconvoluted): 13092.5612. $[M+6H]^{6+}$ 2183.2442, found for 2183.2353; $[M+7H]^{7+}$ 1871.4950, found for 1871.3700; $[M+8H]^{8+}$ 1637.6831, found for 1637.5570; $[M+9H]^{9+}$ 1455.8294, found for 1455.7178; $[M+10H]^{10+}$ 1310.3465, found for 1310.2426; $[M+11H]^{11+}$ 1191.3150, found for 1191.1395.

Third ligation: Native chemical ligation between IFN-β-(1-88)-Segment ABC 27 and IFN-β-(89-166)-Segment D 19

IFN-β-(1-88)-Segment ABC-thioester **27** (0.8 mg, 0.061 µmol) and IFN-β-(89-166)-Segment D **19** (1.2 mg, 0.122 µmol) were dissolved in the freshly prepared degassed sodium phosphate buffer solution (200 mM, pH 6.5, 35 µL) containing 8 M GnHCl and MPAA (40 mM), MPAA

(40 mM) and TCEP (40 mM). This solution was incubated at room temperature for 5 hours. The reaction was monitored by LC-MS. After the completion of reaction, the solution was treated with TCEP (200 mM, 35 µL) before purification by preparative RP- HPLC (Proteonavi, Ø 4.6 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 55:45 to 25:75 over 30 min at 1 ml/min). Glycopeptide **28** was obtained in 0.8 mg, 58% (Isolated yield).

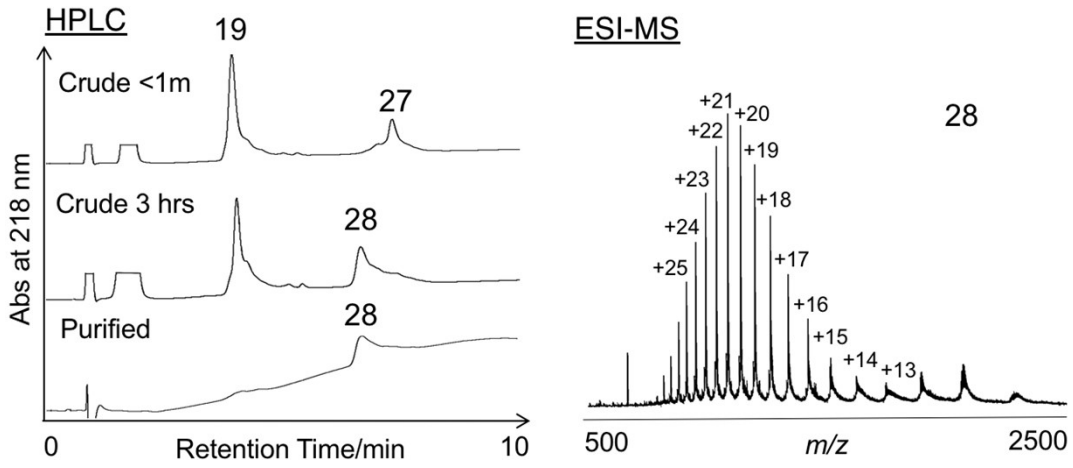


Figure S19: HPLC chromatogram and ESI-MS spectrum of **28**. HRMS (ESI): m/z cald for $C_{1016}H_{1562}N_{252}O_{316}S_8$: $[M+H]^+$ 22619.6000, observed (deconvoluted): 22619.1012. $[M+16H]^{16+}$ 1414.7250, found for 1414.7005; $[M+17H]^{17+}$ 1331.5647, found for 1331.5129; $[M+18H]^{18+}$ 1257.6444, found for 1257.6579; $[M+19H]^{19+}$ 1191.5053, found for 1191.4156; $[M+20H]^{20+}$ 1131.9800, found for 1131.9420; $[M+21H]^{21+}$ 1078.1238, found for 1078.1065; $[M+22H]^{22+}$ 1029.1636, found for 1029.1517.

Synthesis of IFN-β-(1-166)-Segment ABCD 29 by desulfurization

IFN-β-(1-166)-Segment ABCD-OH **28** (0.8 mg, 0.061 µmol) was added to the degassed sodium phosphate buffer (200 mM, pH 7.5, 160 µL) containing GnHCl (8 M), TCEP (250 mM), 14.4 uL of 2-methyl-2-propanethiol (0.8 M, 14 µL) and 0.1 M VA-044 (18 mM, 29 µL).[6] The reaction mixture was incubated for 24 hours at 37 °C. Purification by RP- HPLC purification (Proteonavi, Ø 4.6 mm × 250 mm 0.1% TFA: 0.1% TFA in 90% MeCN = 55:45 to 25:75 over 30 min at 1 ml/min) to give **29**. Isolated yield: 0.1 mg, 7.3%.

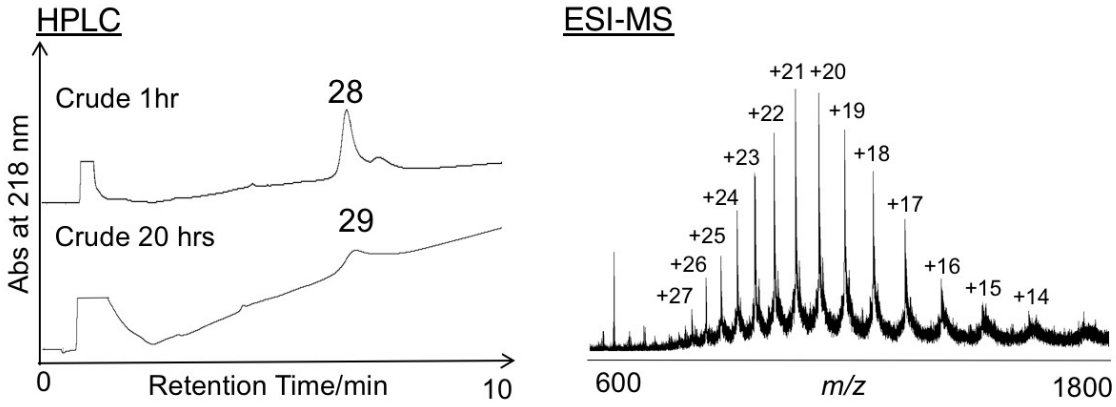


Figure S20: HPLC chromatogram and ESI-MS spectrum of **29**. HRMS (ESI): m/z cald for $C_{1016}H_{1562}N_{252}O_{316}S_7$: $[M+H]^+$ 22587.5400, observed (deconvoluted): 22587.5351 $[M+18H]^{18+}$ 1255.8633, found for 1255.8014; $[M+19H]^{19+}$ 1189.8179, found for 1189.7637; $[M+20H]^{20+}$ 1130.3770, found for 1130.3625; $[M+21H]^{21+}$ 1076.5971, found for 1076.5556; $[M+22H]^{22+}$ 1027.7064, found for 1027.6357.

Synthesis of IFN-β-(1-166)-Segment ABCD 30 by deprotection of phenacyl groups

The desulfurized IFN-β-(1-166)-Segment ABCD-OH **29** (0.1 mg, 0.008 μmol) was dissolved in a solution containing sodium phosphate (200 mM, 0.1 mL), 47% (v/v) of acetic acid, 8M GnHCl and zinc powder (18 mg). The resulting mixture was incubated at 25 °C for 1 hour. Based on our extensive studies, lyophilization after HPLC purification can accelerate the aggregation due to the highly hydrophobic character of the glycosyl polypeptide. Therefore, this glycosyl polypeptide was isolated by dialysis by centrifugal filter units (Amicon Ultra-0.5, MWCO 10K). The reaction solution was poured into the centrifugal filter and remove the reaction solution by centrifugation and then the resultant peptide **30** was washed with a solution containing acetic acid (47%, v/v), 8M GnHCl and 200 mM of sodium phosphate. The resultant solution should be kept without lyophilization. HPLC and Mass spectrum indicated that deprotection of Pac group was to be good yield.

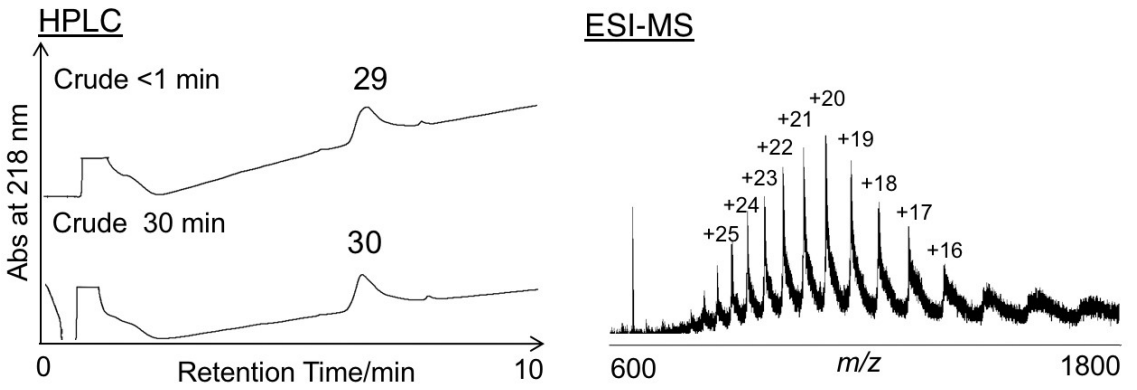
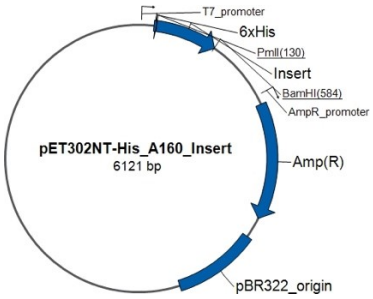


Figure S21: HPLC chromatogram and ESI-MS spectrum of **30**. HRMS (ESI): m/z calcd for $C_{992}H_{1544}N_{252}O_{313}S_7$: $[M+H]^+$ 22233.1350, observed: 22233.7281. $[M+18H]^{18+}$ 1236.1742, found for 1236.1626; $[M+19H]^{19+}$ 1171.1650, found for 1171.1214; $[M+20H]^{20+}$ 1112.6568, found for 1112.6975; $[M+21H]^{21+}$ 1059.7207, found for 1059.7156; $[M+22H]^{22+}$ 1011.5971, found for 1011.5603.

Plasmid preparation

IFN peptide (31-78).



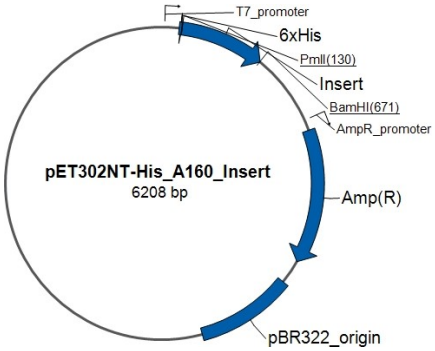
Plasmid structure ordered to Thermo Fisher Scientific. The following sequence was inserted between PmlI (CACGTG) and BamHI sequence (GGATCC).

```
                M  S  D  S  E  V  N  Q  E  A  K  P  E  V  K  P  E  V  K  P  E
1.  CACGTGATGAGCGATAGCGAAGTTAATCAAGAAGCAAAACCGGAAGTTAAGCCGGAAGTGAAACCTGAA
    T  H  I  N  L  K  V  S  D  G  S  S  E  I  F  F  K  I  K  K  T  T  P
70. ACACATATTAACCTGAAAGTGAGTGATGGCAGCAGCGAAATCTTCTTCAAAATCAAAAAAACCACACCG
    L  R  R  L  M  E  A  F  A  K  R  Q  G  K  E  M  D  S  L  R  F  L  Y
139. CTGCGTCGTCTGATGGAAGCATTTGCAAAACGTCAGGGTAAAGAAATGGATAGCCTGCGTTTTCTGTAT
    D  G  I  R  I  Q  A  D  Q  T  P  E  D  L  D  M  E  D  N  D  I  I  E
208. GATGGTATTCGTATTCAGGCAGATCAGACACCGGAAGATCTGGATATGGAAGATAACGATATTATCGAA
    A  H  R  E  Q  I  G  G  C  L  K  D  R  M  N  F  D  I  P  E  E  I  K
277. GCACATCGTGAGCAGATTGGTGGTTGTCTGAAAGATCGTATGAACTTTGATATCCCGGAAGAGATTAAA
    Q  L  Q  Q  F  Q  K  E  D  A  A  L  T  I  Y  E  M  L  Q  N  I  F  A
346. CAGCTGCAGCAGTTCCAGAAAGAAGATGCAGCACTGACCATTTATGAAATGCTGCAGAACATCTTTGCC
    I  F  R  Q  D  S  S  S  T  G  C
415. ATCTTTCGTCAGGATAGCAGCAGCACCGGTTGTTAATAGTGAGGATCC
```

After SUMO sequence (MSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSE

IFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIG G), we add IFN peptide (31-78) with cysteine at the C-terminal.

IFN peptide (89-166).



Plasmid structure ordered to Thermo Fisher Scientific. The following sequence was inserted between PmlI (CACGTG) and BamHI sequence (GGATCC).

```
          M  S  D  S  E  V  N  Q  E  A  K  P  E  V  K  P  E  V  K  P  E
1.  CACGTGATGAGCGATAGCGAAGTTAATCAAGAAGCAAAACCGGAAGTTAAGCCGGAAGTGAAACCTGAA
    T  H  I  N  L  K  V  S  D  G  S  S  E  I  F  F  K  I  K  K  T  T  P
70. ACACATATTAACCTGAAAGTGAGTGATGGCAGCAGCGAAATCTTCTTCAAAATCAAAAAAACCACACCG
    L  R  R  L  M  E  A  F  A  K  R  Q  G  K  E  M  D  S  L  R  F  L  Y
139. CTGCGTCGTCTGATGGAAGCATTTGCAAAACGTCAGGGTAAAGAAATGGATAGCCTGCGTTTTCTGTAT
    D  G  I  R  I  Q  A  D  Q  T  P  E  D  L  D  M  E  D  N  D  I  I  E
208. GATGGTATTCGTATTCAGGCAGATCAGACACCGGAAGATCTGGATATGGAAGATAACGATATTATCGAA
    A  H  R  E  Q  I  G  G  C  N  V  Y  H  Q  I  N  H  L  K  T  V  L  E
277. GCACATCGTGAGCAGATTGGTGGTTGTAATGTTTATCATCAGATCAACCATCTGAAAACCGTGCTGGAA
    E  K  L  E  K  E  D  F  T  R  G  K  L  M  S  S  L  H  L  K  R  Y  Y
346. GAAAAACTGGAAAAAGAGGATTTTACACGCGGTAAACTGATGAGCAGCCTGCATCTGAAACGTTATTAT
    G  R  I  L  H  Y  L  K  A  K  E  Y  S  H  C  A  W  T  I  V  R  V  E
415. GGTCGTATTCTGCACTACCTGAAAGCCAAAGAATATAGCCATTGTGCATGGACCATTGTGCGTGTTGAA
    I  L  R  N  F  Y  F  I  N  R  L  T  G  Y  L  R  N
484. ATTCTGCGTAACTTCTATTTCATTAATCGCCTGACCGGTTATCTGCGCAATTAATAGTGAGGATCC
```

After SUMO sequence (MSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSE IFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIG G), we add IFN peptide (89-166), but Ala was mutated with cysteine.

References

1 M. Murakami, R. Okamoto, M. Izumi and Y. Kajihara, *Angew. Chemie - Int. Ed.*, 2012, **51**, 3567–3572.

2 X. Li, T. Kawakami and S. Aimoto, *Tetrahedron Lett.*, 1998, **39**, 8669–8672.

3 K. Sato, S. Tanaka, K. Yamamoto, Y. Tashiro, T. Narumi and N. Mase, *Chem. Commun.*, 2018, **54**, 9127–9130.

4 H. Katayama and S. Morisue, *Tetrahedron*, 2017, **73**, 3541–3547.

5 M. Matveenko, S. Hackl and C. F. W. Becker, *ChemistryOpen*, 2018, **7**, 106–110.

6 T. Kiuchi, M. Izumi, Y. Mukogawa, A. Shimada, R. Okamoto, A. Seko, M. Sakono, Y. Takeda, Y. Ito and Y. Kajihara, *J. Am. Chem. Soc.*, 2018, **140**, 17499–17507.