

Supporting Information

Monitoring of glycoprotein quality control system with a series of chemically synthesized homogeneous native and misfolded glycoproteins.

Tatsuto Kiuchi,[†] Masayuki Izumi,^{†#} Yuki Mukogawa,[†] Arisa Shimada,[†] Ryo Okamoto,[†] Akira Seko,^{‡#} Masafumi Sakono,^{‡#} Yoichi Takeda,^{‡#} Yukishige Ito^{‡§} and Yasuhiro Kajihara^{††*}

[†]Department of Chemistry, Graduate School of Science, Osaka University, 1-1 Machikaneyama, Toyonaka, Osaka 560-0043, Japan. [‡]ERATO Ito glycotriology project, Japan Science and Technology Agency (JST), 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. [§]Synthetic Cellular Chemistry Laboratory, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan.

Biochemistry section

Fig. S 1. ANS assay of synthetic M9-EPO analogs and synthetic M9-IFN β

Fig. S 2. Structure of proteins binding to 1-anilinonaphthalene-8-sulfonic acid (ANS).

Fig. S3. Hydrophobic amino acids exposed on the surface of M9-IFN β and EPO bearing three glycans.

Fig. S4. The formation of G1M9-glycan in the presence or absence of an inhibitor on the ER assays.

Fig. S5. Monitoring glucosylation and deglucosylation in the ER lysate by means of UDP-[1-¹³C]-glucose.

Fig. S6. Simulation of high-resolution mass spectra of G1M9-glycoprotein and [¹³C]-G1M9-glycoprotein.

Fig. S7. Monitoring of glucosylation amount during refolding process of misfolded M9-IL8 in the ER lysate.

Fig. S8. Monitoring of glucosylation amount of native M9-IL8 in the ER lysate.

Fig. S9. Monitoring of glucosylation amount of refolding process of dimeric misfolded M9-IL8 in the ER lysate.

Fig. S10. Relationship between glucosylation/de-glucosylation and refolding ability of misfolded dimeric G1M9-IL8 in the ER lysate.

Synthesis section

Fig. S11. Amino acid sequence of erythropoietin.

Fig. S12. Synthetic strategy of erythropoietin glycoforms.

Fig. S13. Isolation of homogeneous M9-high-mannose type oligosaccharide from egg yolk and modification by Boc group.

Fig. S14. Hydrophobicity of EPO segment (50-166 peptide) and their mutant N83K.

Peptide fragment synthesis

Fig. S15. HPLC profiles and ESI-MS spectrum of purified peptide segment A.

Fig. S16. HPLC profiles and ESI-MS spectrum of purified peptide segment B.

Fig. S17. HPLC profiles and ESI-MS spectrum of purified peptide segment C.

Fig. S18. HPLC profiles and ESI-MS spectrum of purified peptide segment D.

Fig. S19. HPLC profiles and ESI-MS spectrum of purified peptide segment E.

Fig. S20. HPLC profiles and ESI-MS spectrum of purified glycopeptide segment B.

Fig. S21. HPLC profiles and ESI-MS spectrum of purified glycopeptide segment D.

Fig. S22. HPLC profiles and ESI-MS spectrum of purified glycopeptide segment A.

Fig. S23. HPLC profiles and ESI-MS spectrum of purified peptide segment F.

Synthesis of EPO having three M9-glycans.

Fig. S24. 1st-NCL for the construction of EPO bearing three M9-glycans.

Fig. S25. 2nd-NCL for the construction of EPO bearing three M9-glycans.

Fig. S26. 3rd-NCL for the construction of EPO bearing three M9- glycans.

Fig. S27. 4th-NCL for the construction of EPO bearing three M9- glycans.

Fig. S28. Desulfurization of glycopeptide segment BCDEF for the construction of EPO bearing three M9-glycans.

Fig. S29. Deprotection of Acn groups of glycopeptide segment BCDEF for the construction of EPO bearing three M9-glycans.

Fig. S30. Final NCL for the construction of EPO bearing three M9-glycans.

Fig. S31. Folding of segment ABCDEF for the construction of EPO bearing three M9-glycans.

Synthesis of EPO having a M9-glycans at 24.

Fig. S32. 2nd-NCL for the construction of EPO bearing an M9-glycans at 24.

Fig. S33. 3rd-NCL for the construction of EPO bearing an M9-glycans at 24.

Fig. S34. 4th-NCL for the construction of EPO bearing an M9- glycans at 24.

Fig. S35. Desulfurization of glycopeptide segment BCDEF for the construction of EPO bearing an M9-glycans at 24.

Fig. S36. Deprotection of Acn groups of glycopeptide segment BCDEF for the construction of EPO bearing an M9- glycans at 24.

Fig. S37. Final NCL for the construction of EPO bearing an M9-glycans at 24.

Fig. S38. Folding of segment ABCDEF for the construction of EPO bearing an M9-glycans at 24.

Synthesis of EPO having a M9-glycans at 38.

Fig. S39. 4th-NCL for the construction of EPO bearing an M9-glycans at 38.

Fig. S40. Desulfurization of glycopeptide segment BCDEF for the construction of EPO bearing an M9-glycans at 38.

Fig. S41. Deprotection of Acn groups of glycopeptide segment BCDEF for the construction of EPO bearing an M9-glycans at 38.

Fig. S42. Final NCL for the construction of EPO bearing an M9-glycans at 38.

Fig. S43. Folding of segment ABCDEF for the construction of EPO bearing an M9-glycans at 38.

Synthesis of EPO having a M9-glycans at 83.

Fig. S44. 4th-NCL for the construction of EPO bearing an M9-glycans at 83.

Fig. S45. Desulfurization of glycopeptide segment BCDEF for the construction of EPO bearing an M9-glycans at 83.

Fig. S46. Deprotection of Acn groups of glycopeptide segment BCDEF for the construction of EPO bearing an M9-glycans at 83.

Fig. S47. Final NCL for the construction of EPO bearing an M9-glycans at 83.

Fig. S48. Folding of segment ABCDEF for the construction of EPO bearing an M9-glycans at 83.

Analysis of disulfide bond positions of native and misfolded EPOs.

Fig. S49. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from native folded EPO bearing three glycans (compound 1).

Fig. S50. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from misfolded EPO bearing three glycans (compound 5).

Fig. S51. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from native folded EPO bearing a glycan at 24 (compound 2).

Fig. S52. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from misfolded EPO bearing a glycan at 24 (compound 6).

Fig. S53. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from native folded EPO bearing a glycan at 38 (compound 3).

Fig. S54. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from misfolded EPO bearing a glycan at 38 (compound 7).

Fig. S55. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from native EPO bearing a glycan at 83 (compound 4).

Fig. S56. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from misfolded EPO bearing a glycan at 83 (compound 8).

Circular dichroism of EPOs.

Fig. S57. Circular dichroism (CD) spectrum of native folded EPO bearing one or three glycans

Fig. S58. Circular dichroism (CD) spectrum of misfolded EPO bearing one or three glycans

Characterization of synthetic EPOs.

Fig. S59. Biological activity of commercially available EPO and synthetic native folded EPO bearing three glycans (compound 1).

Fig. S60. Biological activity of commercially available EPO and synthetic EPO bearing a glycan at 38 (compound 3).

Fig. S61. Denaturing of misfolded EPO bearing a glycan at 38 (compound 7).

Fig. S62. Glycoproteins for UGGT substrates.

Synthesis of IFN- β

Fig. S63. Amino acids sequence of M9-IFN β

Fig. S64. Synthetic strategy of M9-IFN β .

Fig. S65. HPLC profile and ESI-MS spectrum of purified glycopeptide segment B.

Fig. S66. 1st-NCL for the construction of M9-IFN β .

Fig. S67. 2nd-NCL for the construction of M9-IFN β .

Fig. S68. Desulfurization of glycopeptide segment ABC for the construction of M9-IFN β .

Fig. S69. Deprotection of Ac group of glycopeptide segment ABC for the construction of M9-IFN β .

Fig. S70. Folding of segment ABC for the construction of M9-IFN β .

Fig. S71. Circular dichroism (CD) spectrum of synthetic M9-IFN β

References

Biochemistry

RP-HPLC analyses (analytical or semipreparative) for chemistry and biological assays were performed using Vydac C18, C8 (GRACE Co.), Cadenza CD-C18 (Imtakt), Proteonavi or Capcell pak (Shiseido). Proteonavi (Shiseido) were used in LC-MS analysis. LC-MS and ESI-MS spectra were recorded on a Bruker Daltonics amaZon-mass spectrometer system and Bruker Daltonics Esquire 3000 mass spectrometer, respectively. High-resolution MS spectra were recorded on a FTICR-MS (Bruker solariX XR)

Biological Activity of correctly folded EPO38 (compound 3) and EPO-24, 38, 83 (ref 1) (compound 1)

The biological activities of correctly folded EPO-38 (compound 3) and EPO-24, 38, 83 (compound 1) were evaluated. A cell proliferation assay using TF-1 cells, which were established from marrow cell, was examined. Commercially available EPO and only the cell culture medium were used as a positive control and negative control, respectively. The plates containing TF-1 cells and EPO were incubated under 5% CO₂ atmosphere at 37°C. After 72 h, cell growth was estimated by using of 2-(2-methoxy-4-nitrophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfohenyl)-2H tetrazolium, monosodium salt (WST-8). WST-8 was added to each well of the plate and incubated for 2 h. Absorbance of WST-8 in each well was measured at 450 nm and 630 nm.

Recombinant UGGT assay for correctly folded EPO analogs and M9-INF- β (ref 2)

UGGT assay were carried out in 50 mM Tris-HCl at pH 7.5 containing 5 mM CaCl₂, 0.5 mM UDP-Glc, 5 μ M of folded EPO or M9-INF- β and 3.9 ng/ μ L recombinant UGGT. Recombinant human UGGT was expressed by yeast. The mixture was incubated at 37°C. This reaction was monitored by LC-MS.

Recombinant UGGT assay for misfolded EPO analogs

UGGT assay with misfolded EPO analogs are difficult to evaluate due to the disordered MS spectra of misfolded EPO analogs. Therefore, trypsin digestion was applied to analyze whether the misfolded EPO analogs were glucosylated or not. After the reaction with misfolded EPO analogs and UGGT as following above, the solutions were heated up to 100°C to deactivate UGGT. Then, trypsin was added to this solution for the digestion of misfolded EPO to peptide and glycopeptide fragments. Finally, this solution was analyzed by LC-MS. Observing glycopeptide fragment reveals the result of UGGT assay.

ANS assay

The hydrophobicity can be estimated by assays with 1-anilino-8-naphthalenesulfonate, which acted as

a gauge of the protein surface hydrophobicity (ref 3). Binding of ANS was measured with FP-6500 (JASCO). Correctly folded EPO analogs (compound **1**, **2**, **3**, **4**: 10 μ M) or misfolded EPO analogs (compound **5**, **6**, **7**, **8**, 10 μ M) or correctly folded M9-INF- β (10 μ M) were incubated with ANS (50 μ M) at 37°C for 10 min. The excitation wavelength was set at 369 nm. The emission was scanned from 400 nm to 600 nm in the suitable spectrum mode. All measurements were carried out in 50 mM Tris-HCl at pH 7.5.

Isolation of rough ER (ref 4)

Rough ER lysate was isolated from rat liver by sucrose density gradient centrifugation. Rat liver 33 g was cut into small pieces and the fibrous were filtered with grind on ice. The extract was homogenized by a homogenizer. The homogenate was centrifuged at 10000 g for 10 min at 4°C. The supernatant was recovered and centrifuged at 10500 g for 60 min at 4°C again. The precipitate was suspended by 2.1 M sucrose containing Tris-HCl, MgCl₂ and KCl. This suspension was filtered and applied to sucrose density gradient (0.3 M, 0.8 M and 1.15 M). This gradient was centrifuged at 150000 g for 90 min at 4°C. The bottom-most layer was recovered and applied to next sucrose density gradient (0.8 M, rough ER containing layer (1.35 M) and 2.1M). This gradient was centrifuged at 300000 g for 60 min at 4°C. The layer between 2.1 M and 1.35 M was recovered. The amount of proteins was estimated by the Bradford protein assay.

ER assay

ER assay were carried out in 50 mM HEPES at pH 7.5 containing 5 mM CaCl₂, 2 mM UDP-Glc, 5 μ M of correctly folded EPO or INF- β and 290 μ g/ml ER lysate. 1-deoxynojirimycin (5 mM) and kifunensine (1 mM) were used as inhibitors of glucosidases and mannosidases, respectively. The mixture was incubated at 37°C. This reaction was monitored by LC-MS. Because ER fraction contains ER-mannosidases, this assay also gave a tiny amount of M8-glycoprotein from M9-glycoprotein substrate by ER-mannosidases. In order to monitor the reactions clearly, we needed to inhibit mannosidase reaction with kifunensine.

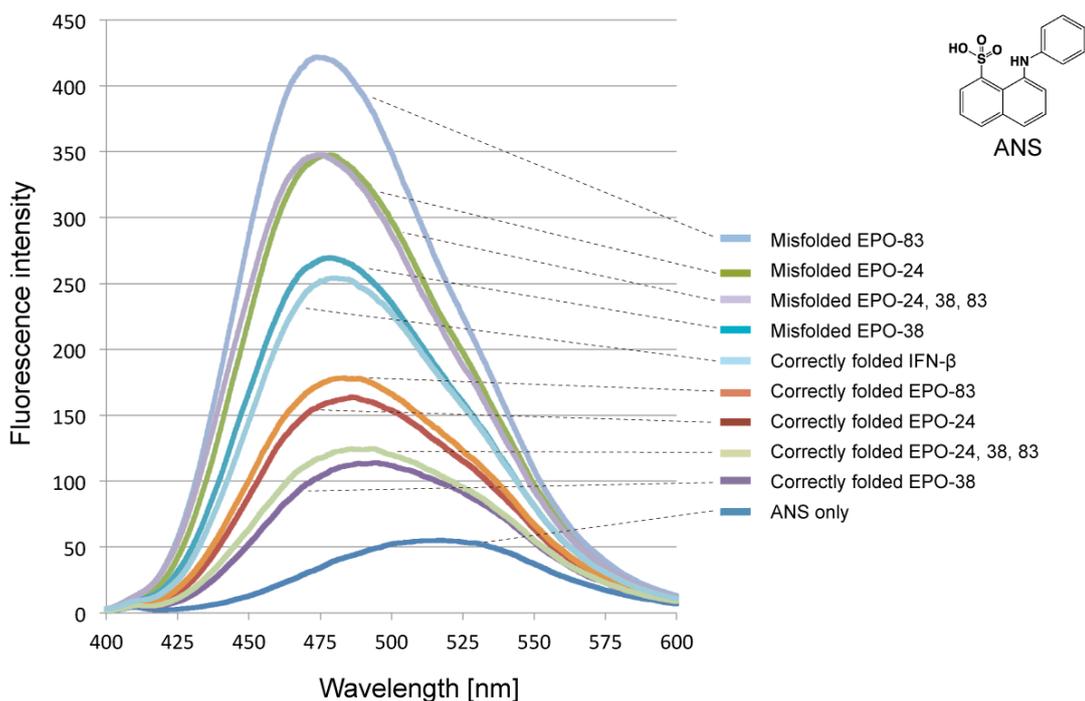


Fig. S1. ANS assay of synthetic M9-EPO analogs and synthetic M9-IFN- β . In order to evaluate hydrophobicity, ANS assay was performed. Individual protein (10 μ M) was dissolved into a solution containing ANS (50 μ M) and then measured fluorescence.

Hydrophobicity of M9-EPO and M9-IFN- β for substrate specificity of UGGT.

Eight EPOs were glucosylated by UGGT, while, M9-IFN- β did not. These results indicated that UGGT recognized EPOs as hydrophobic misfolded glycoproteins. It is known that UGGT recognized hydrophobic patch as a key characteristic nature of misfolded glycoproteins (ref 2). 1-Anilino-8-naphthalenesulfonate (ANS) can be used as a gauge of the protein surface hydrophobicity (ref 3), because ANS binds with hydrophobic protein surface and shows fluorescence.

Although EPOs showed moderate fluorescence, M9-IFN- β showed potent fluorescence than eight EPOs (Fig S1). These data indicates M9-IFN- β is prone to bind with ANS rather than EPOs.

In terms of this inconsistency between UGGT assays and ANS assays, we concluded that EPO has many hydrophobic area exposed on the surface than that of M9-IFN- β (Fig S3). M9-IFN β has also several hydrophobic areas on the surface, but these hydrophobic areas are covered by the hydrophilic amino acids (Fig S3f-i). All models were built by Paymol software and we put hydrophobic amino acids according to ref 5b. Finally, we analyzed the exposed hydrophobic surfaces by a visual inspection and determined the hydrophobic surfaces are covered by hydrophilic amino acids.

We found interesting two papers that investigated the binding mode of ANS to proteins. The first paper reported the structure of camel zeta-crystallin binding to NADPH as a ligand (ref 5a). This paper also described that ANS was found to inhibit the binding of NADPH under a competitive mode

(mixed inhibition). Interestingly, the binding site had hydrophobic area covered by hydrophilic amino acids (Fig. S2, A-D: human zeta crystallin). The second paper showed that the structure of calcium binding protein binding to ANS (Fig. S2, E-G) (ref 5c). The structure shows that naphthalene group of ANS binds to hydrophobic area consists of two Met, Phe and Pro. A sulfonate group of ANS interacts with the hydrophilic area of Asp and Gly-amide.

According to these two papers, ANS seems to have unique binding mode recognizing hydrophobic area covered by hydrophilic amino acids, because ANS has hydrophobic aromatic rings and a hydrophilic sulfonate group.

On the other hand, EPOs has many hydrophobic area exposed on the protein surface (Fig. S3a-e). These hydrophobic area (ref 5b) exposed was selected and made yellow color based on visual inspection. Although we could not use computational analysis, We easily found EPO had hydrophobic area exposed and IFN had many unique hydrophobic area covered by hydrophilic amino acids (Fig. S3a-i). Recognition of UGGT by hydrophobic interaction may prefer to bind to hydrophobic area exposed on the target protein surface.

According to these all results (Fig. S2, 3), we concluded that UGGT recognized native-form EPOs as misfolded like hydrophobic glycoproteins.

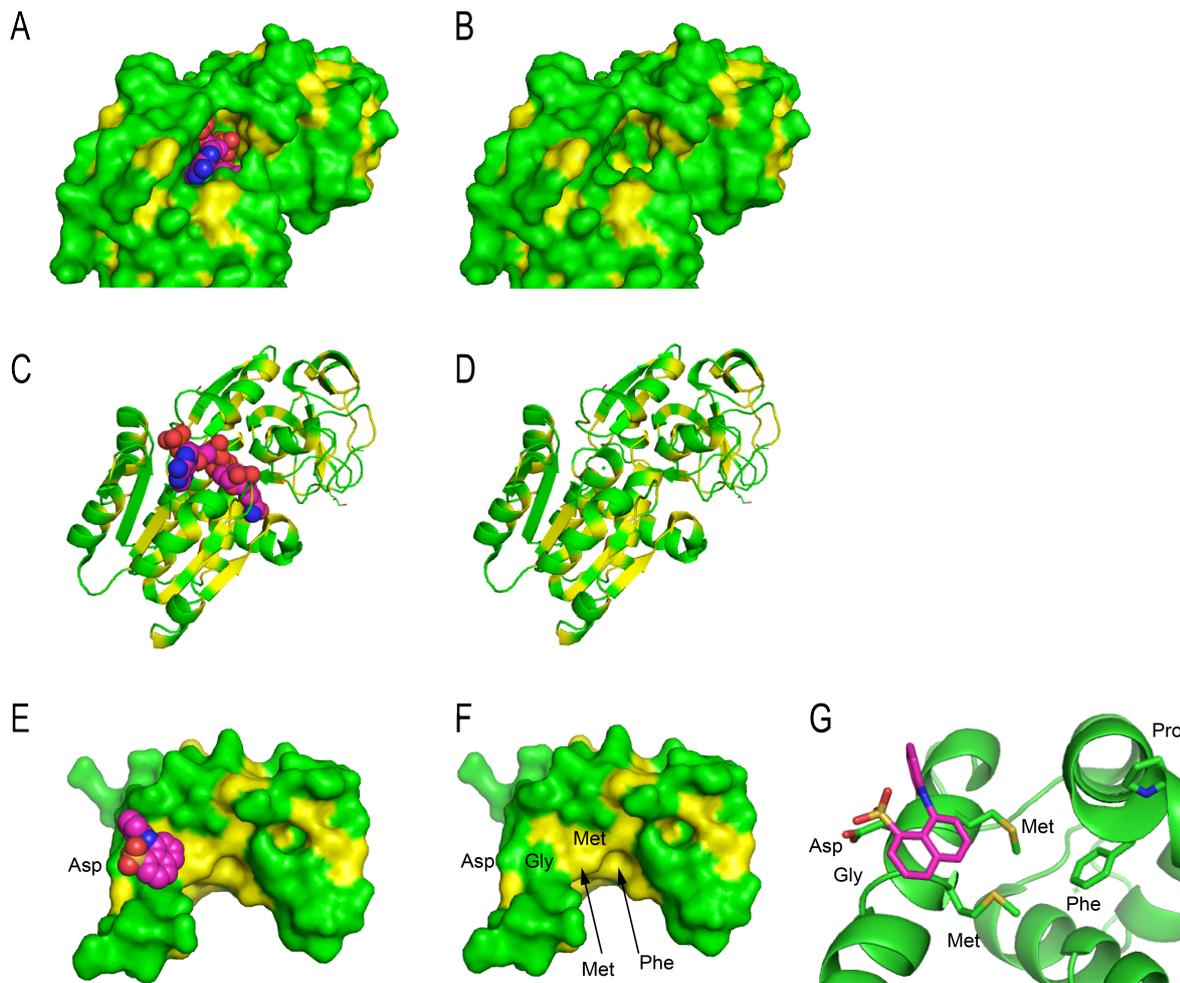


Fig. S2. Binding of 1-anilinoanthracene-8-sulfonic acid (ANS) to several proteins. A) Surface model of zeta-crystallin binding to NADPH (ref 5a). Yellow and green colors indicates hydrophobic amino acid (Ile, Ala, Phe, leu, Met, Pro, Val, and Trp: ref 5b) and hydrophilic amino acids, respectively. B) Surface model of zeta-crystallin without NADPH. *The binding of NADPH was found to be inhibited by ANS under the competitive mode (mixed inhibition) (ref 5a). The inhibition experiments indicated that ANS could bind a hydrophobic NADPH binding site. This combining site has hydrophobic surfaces covered by hydrophilic amino acids like Interferon beta surface.* C) Cartoon model of zeta-crystallin binding to NADPH. D) Cartoon model of zeta-crystallin without NADH. A and B show that NADPH binds with hydrophobic surfaces. Protein models of zeta-crystallin were built by Pymol based on PDB (1YB5). E) Surface model of calcium binding protein binding to ANS (ref 5c). F) Surface model of calcium binding protein without ANS. G) Cartoon model of calcium binding protein binding with ANS. *E, F and G show that naphthalene group of ANS binds to hydrophobic area consists of Met, Phe and Pro. A sulfonate group of ANS interacts with the hydrophilic area of Asp and Gly-amide (ref 5c).* Protein models of calcium binding protein were built by Pymol based on PDB (2WOR).

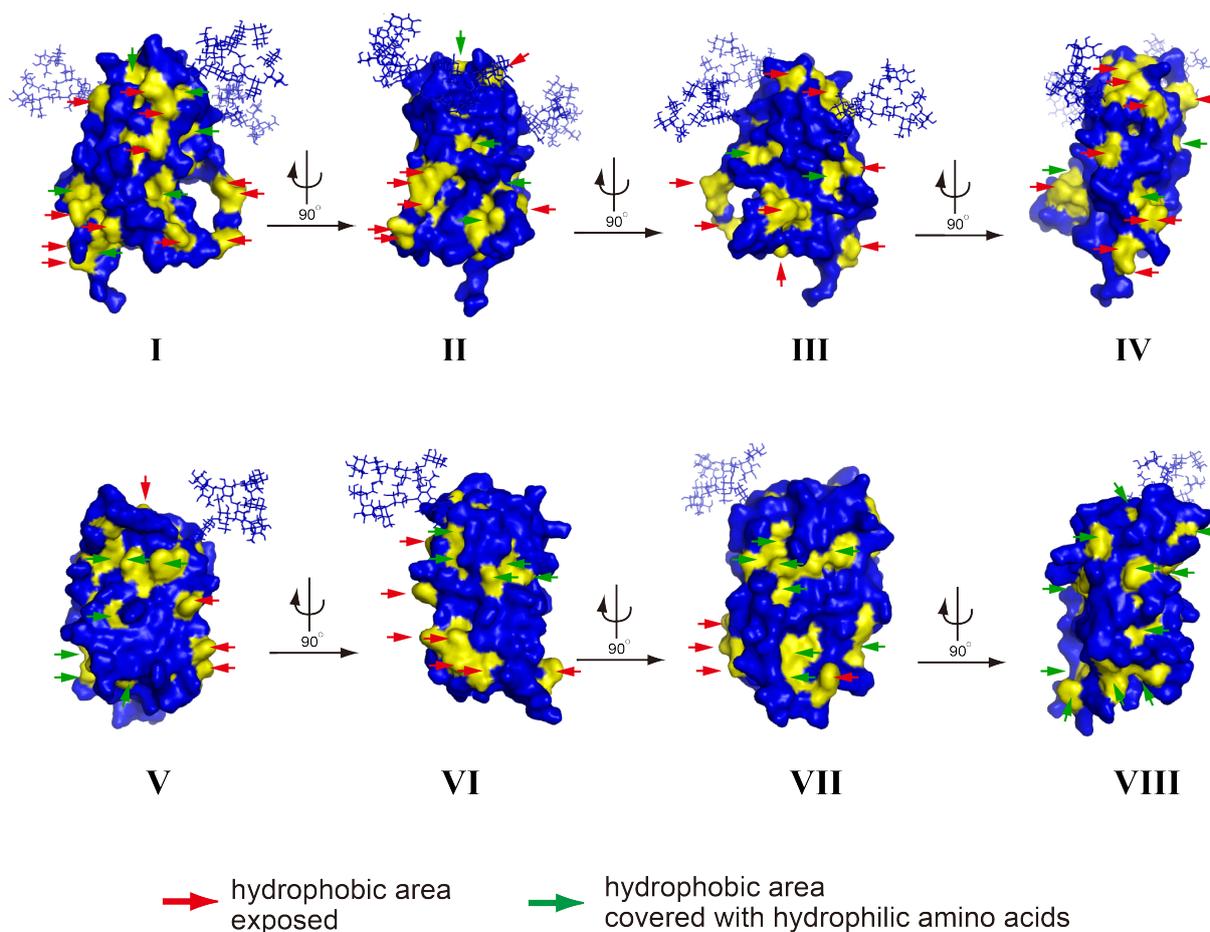


Fig. S3a. Structure of erythropoietin (EPO) having three glycans at 24, 38, and 83 positions and interferon(IFN)- β having a glycans at 80 position. Blue and yellow color indicate hydrophilic and hydrophobic amino acids (Ile, Ala, Phe, leu, Met, Pro, Val, and Trp) (ref 5b). Surface models EPO I-IV and IFN V-VIII are shown with different four angles (every 90 degree different). Protein models were built by Pymol based on PDB (EPO: 1BUY; IFN:1AU1). Red and yellow arrows indicate hydrophobic area exposed and hydrophilic area covered by hydrophilic amino acids, respectively.

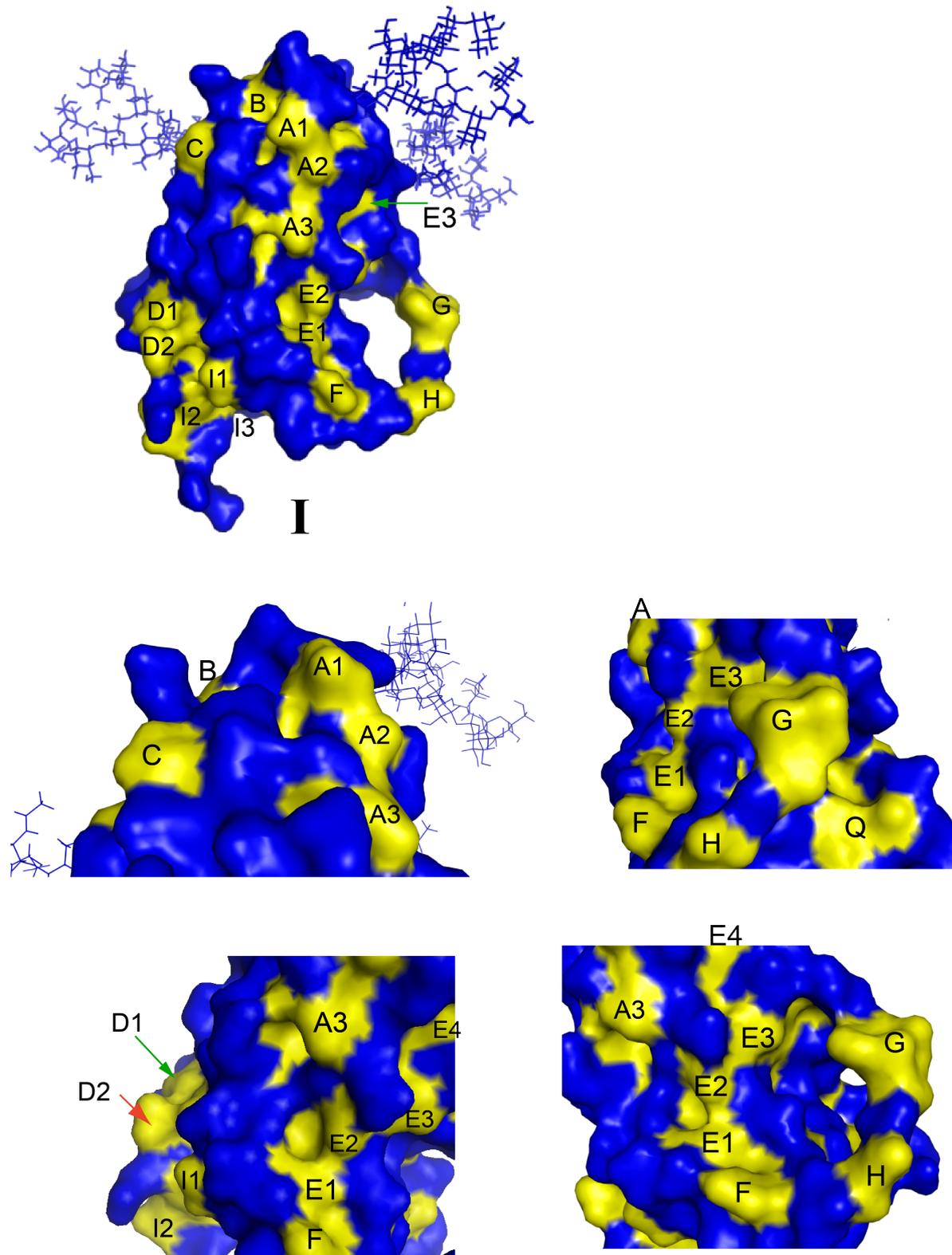


Fig. S3b. Structure I of EPO having three glycans at 24, 38, and 83 positions and their expanded surface.

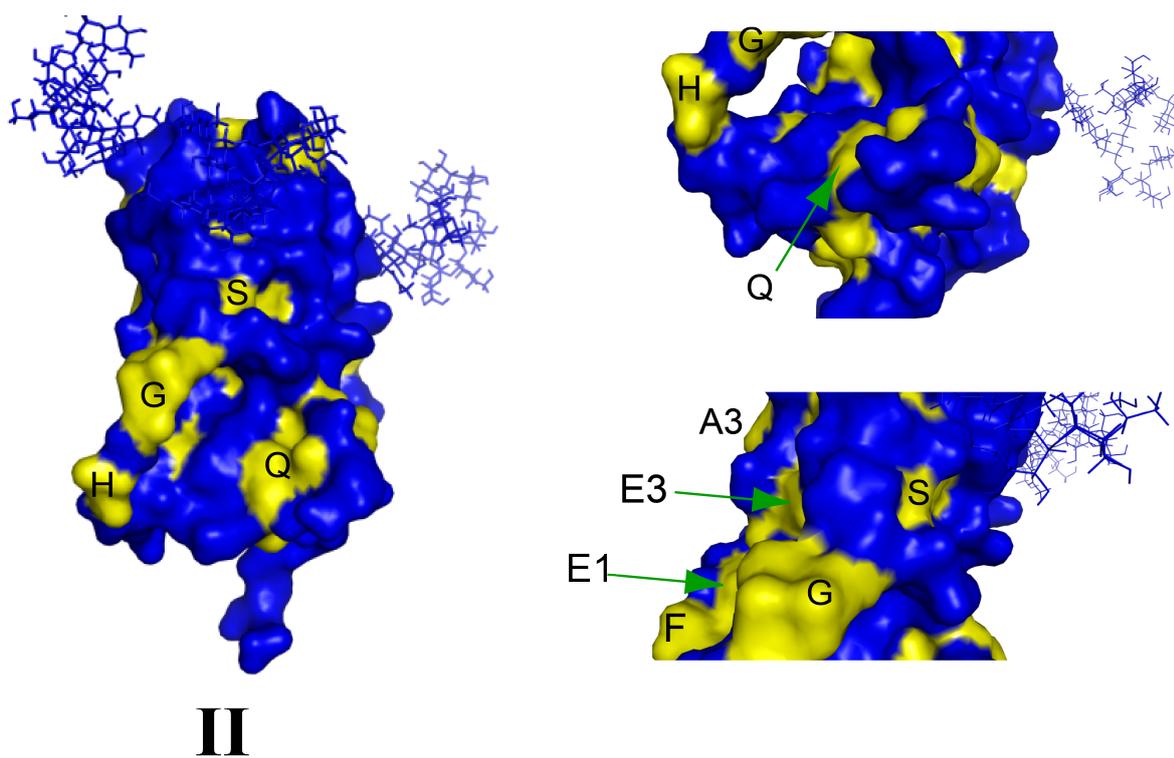


Fig. S3c. Structure II of EPO having three glycans at 24, 38, and 83 positions and their expanded surface.

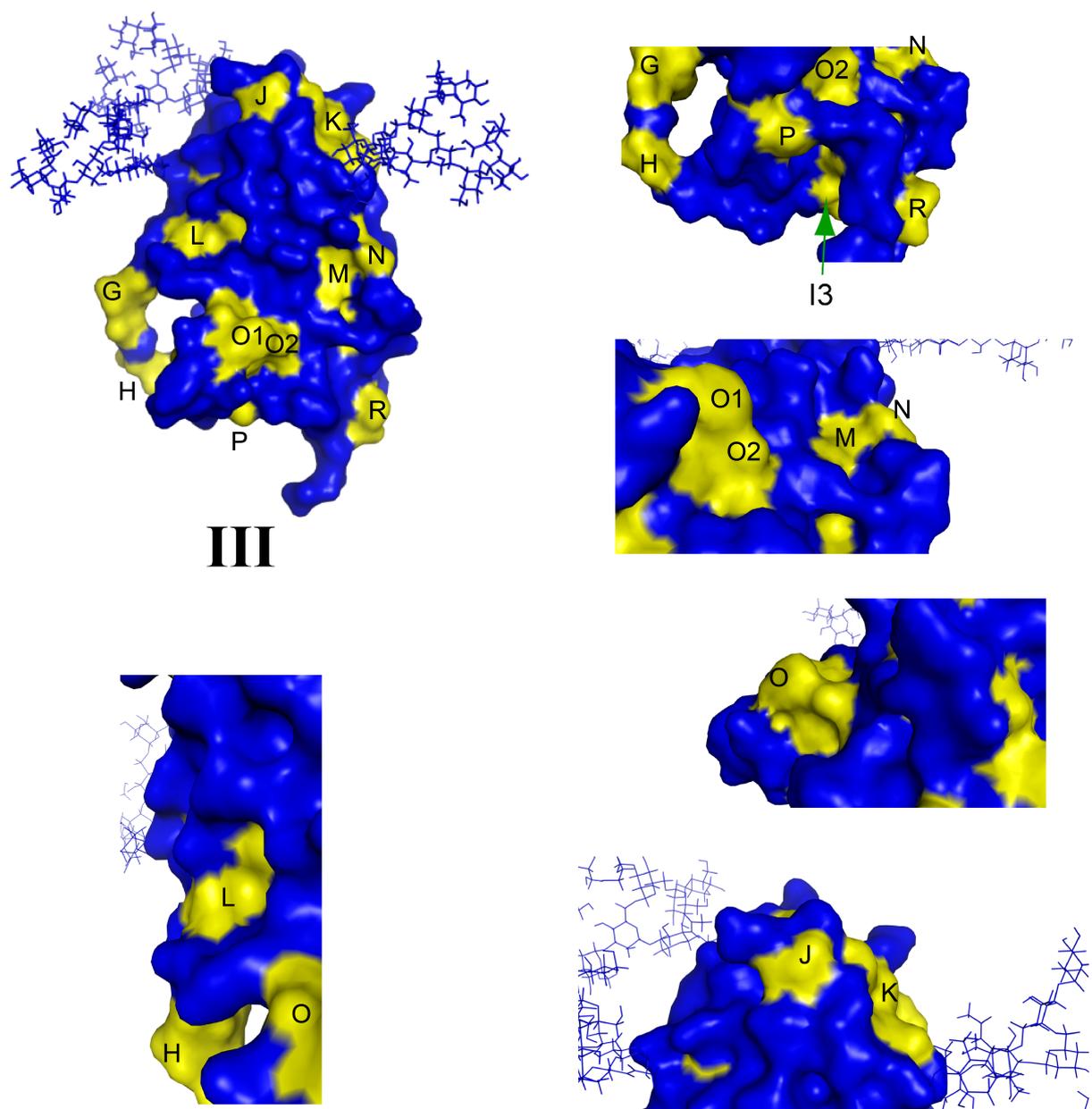
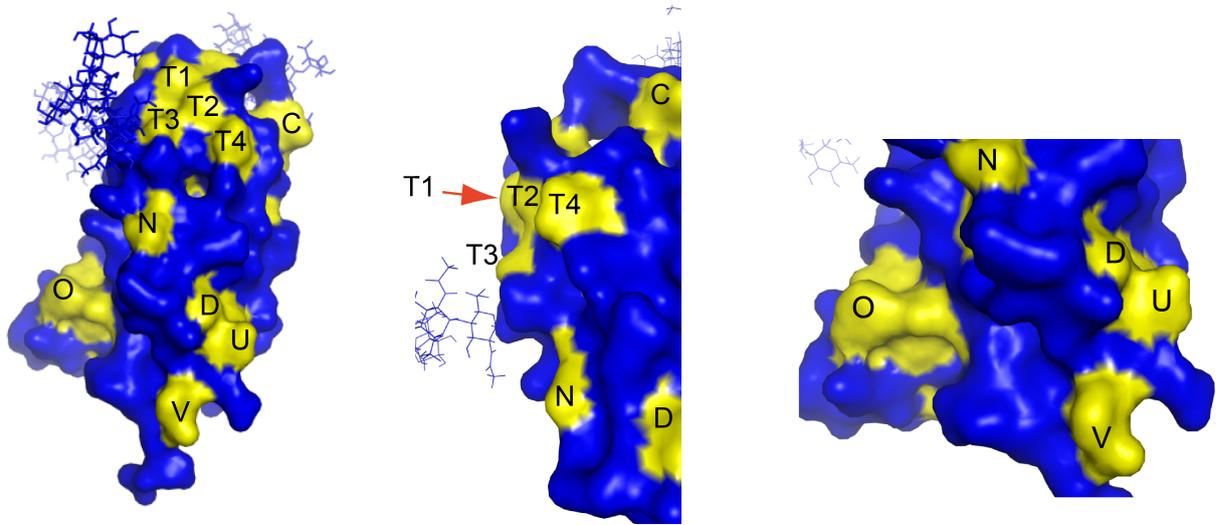


Fig. S3d. Structure III of EPO having three glycans at 24, 38, and 83 positions and their expanded surface.



IV

Fig. S3e. Structure IV of EPO having three glycans at 24, 38, and 83 positions and their expanded surface.

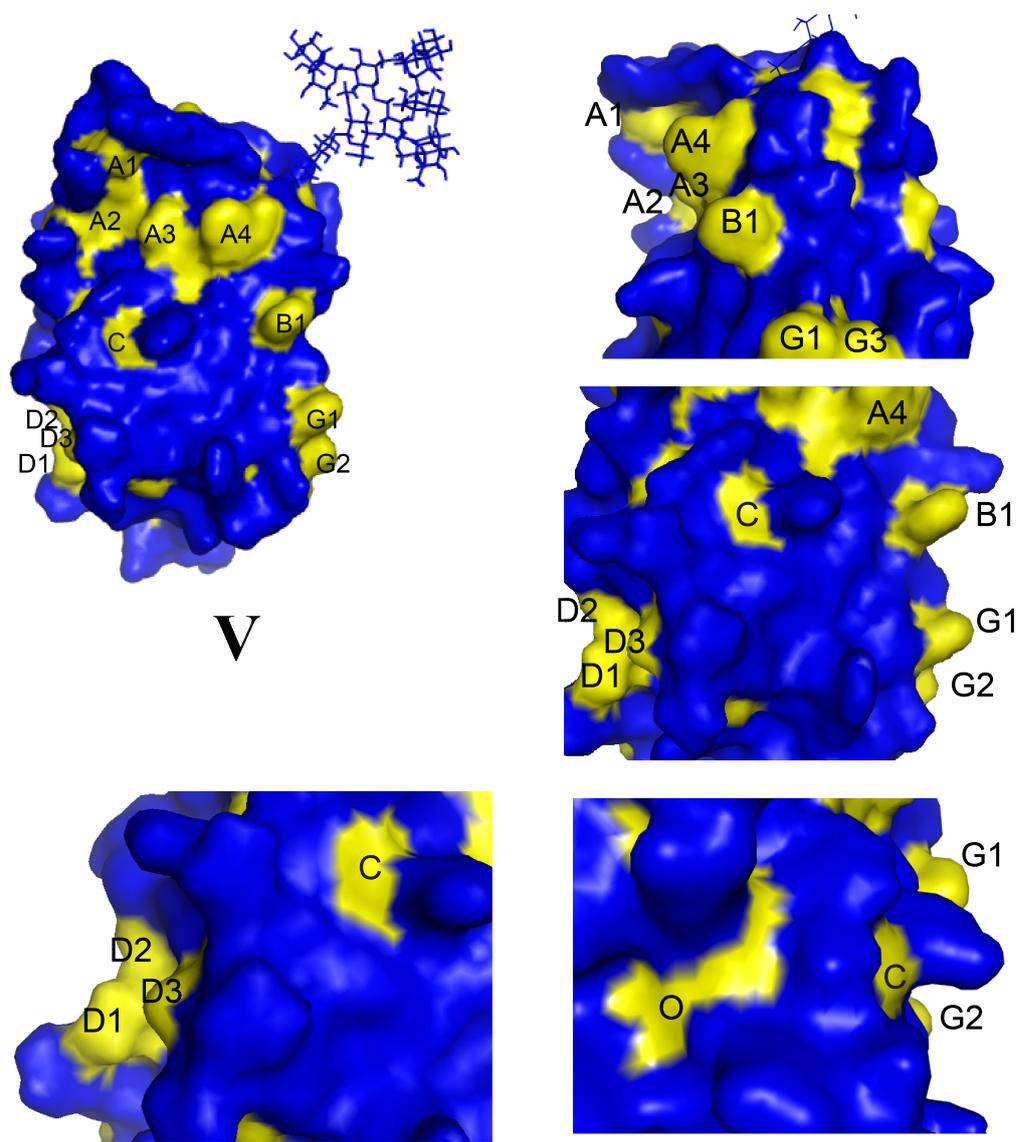


Fig. S3f. Structure V of IFN-β having a glycans at 80 position and their expanded surface.

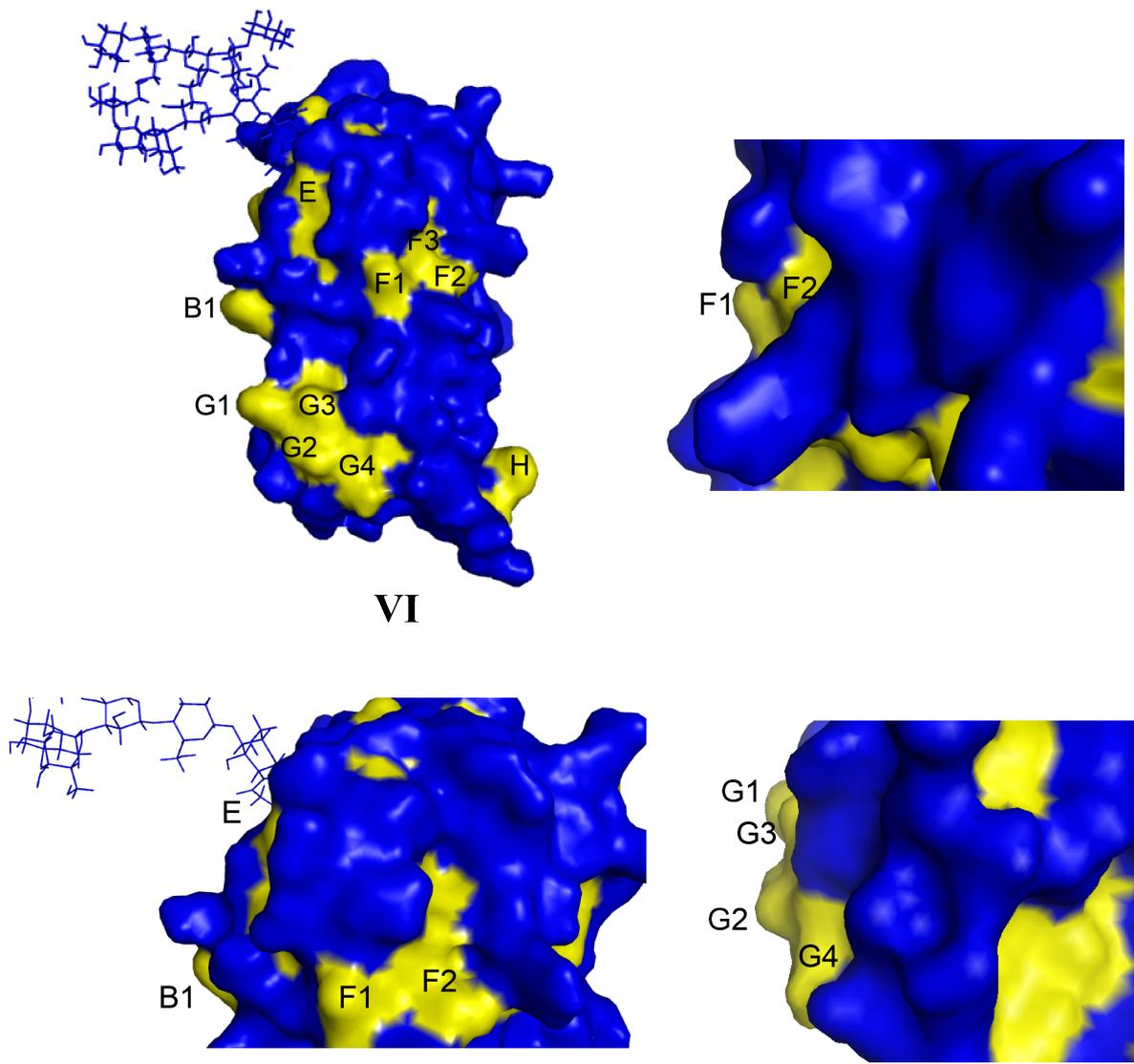
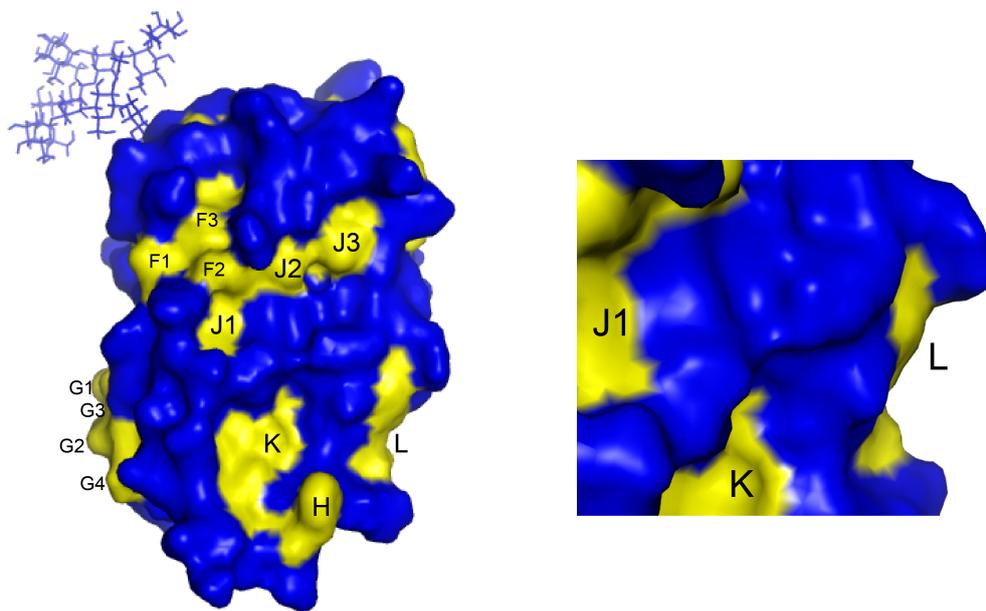


Fig. S3g. Structure VI of IFN-beta having a glycans at 80 position and their expanded surface.



VII

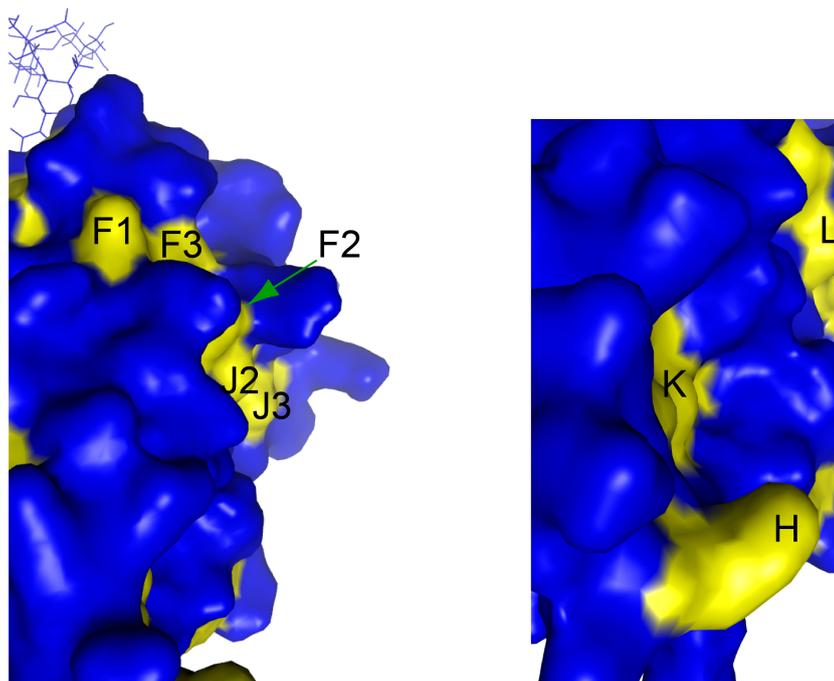


Fig. S3h. Structure VII of IFN- β having a glycans at 80 position and their expanded surface.

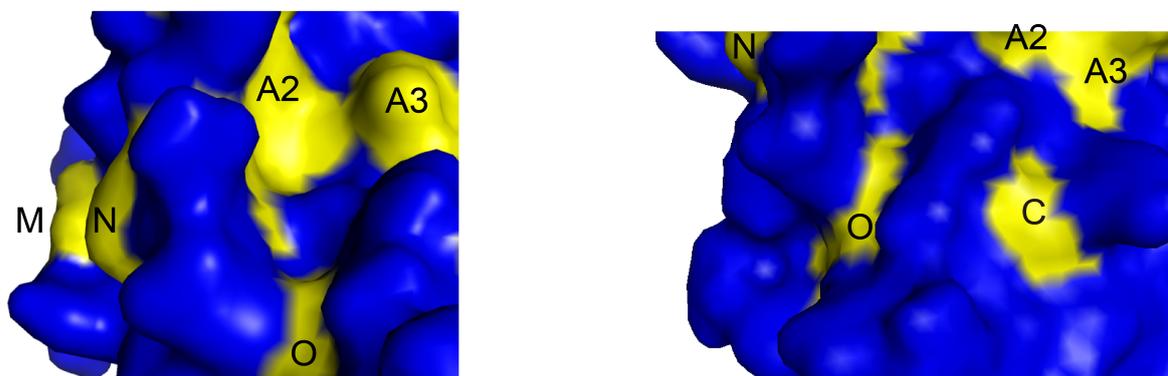
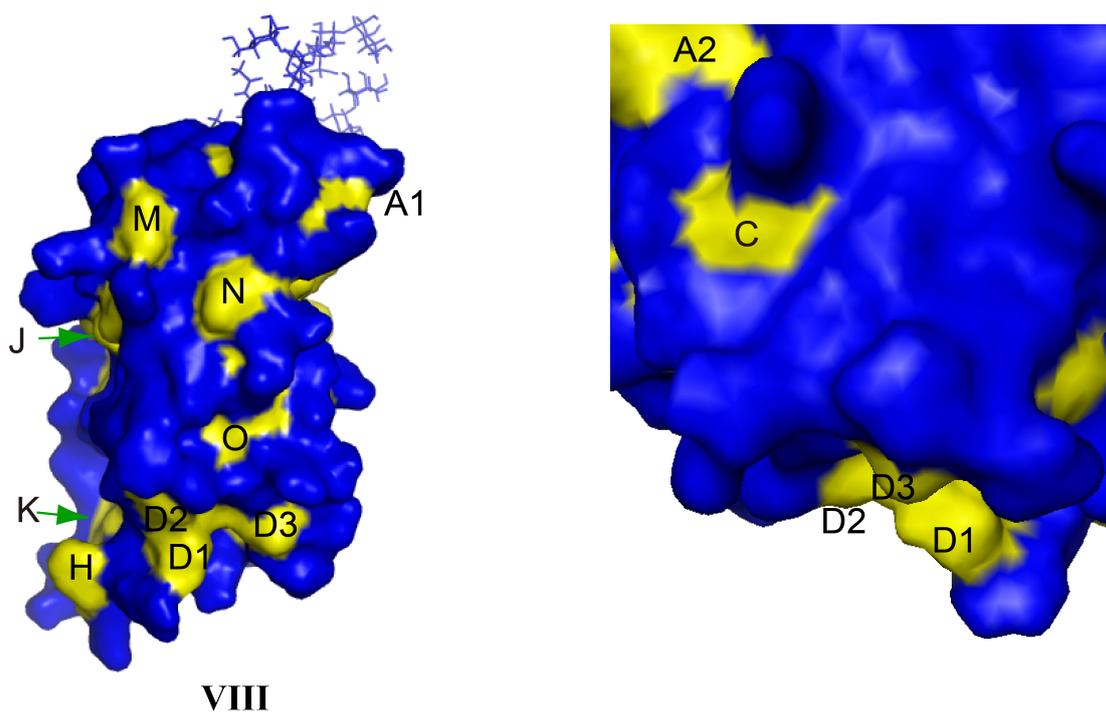


Fig. S3i. Structure VIII of IFN- β having a glycans at 80 position and their expanded surface.

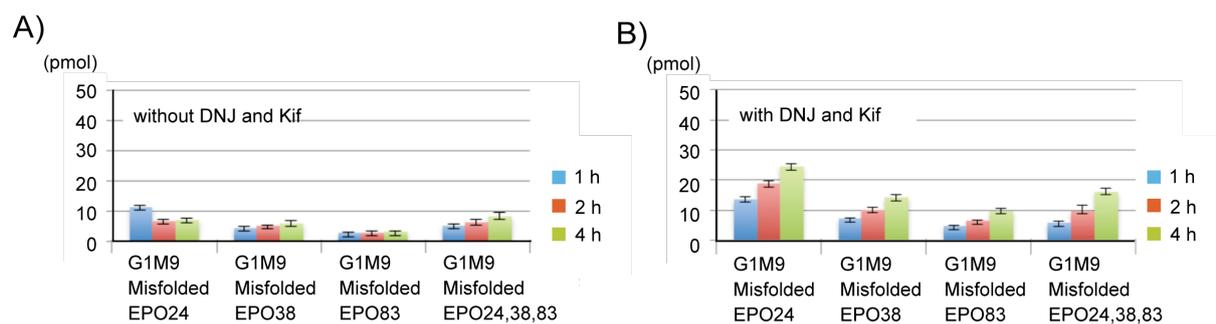


Fig. S4. The formation of G1M9-glycan in the presence or absence of an inhibitor on the ER assays. A) The rate of G1M9-EPO (misfolded) generated by ER lysate in the absent of deoxynojirimycin (DNJ) and kifunensine (Kif) ($\times 10^{-10}$ mol). B) The rate of G1M9-EPO (misfolded) generated by ER lysate in the present of DNJ and Kif ($\times 10^{-10}$ mol). Because M8 gradually generated and this resulted in difficult to analyze. In order to inhibit ER-mannosidases, we added kifunensine.

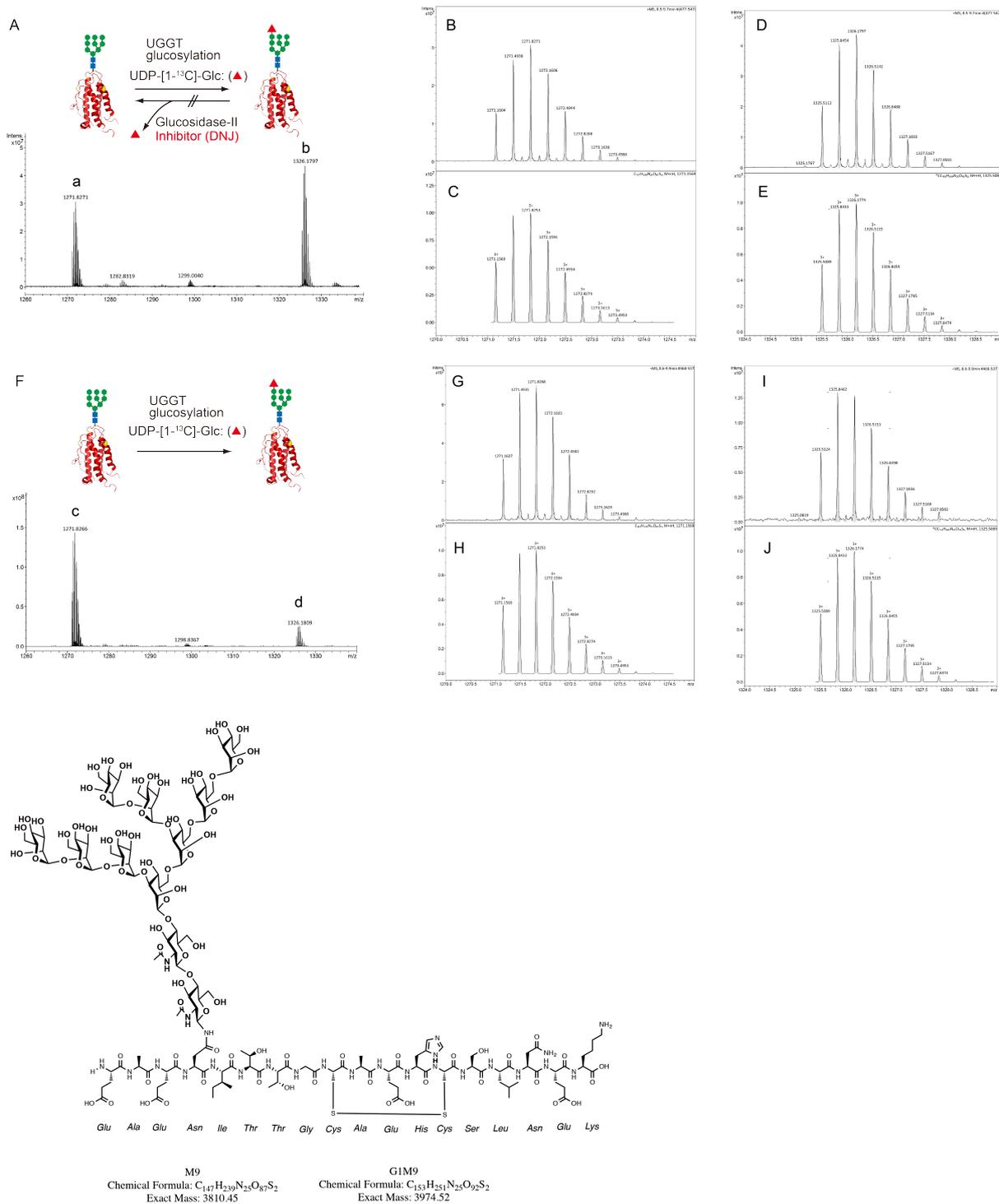


Fig. S5. Monitoring glucosylation and deglucosylation in the ER lysate by means of UDP-[1-¹³C]-glucose. After ER lysate assay in the presence of UDP-[1-¹³C]-Glc, EPO was digested with Lys-C peptidase and the resultant G1M9-peptide and M9-peptide were analyzed by FTICR high resolution mass spectroscopy. EPO substrate used had a M9-glycan at 38 position. The resultant peptide sequence was GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluLys and their HPLC profile included both G1M9-peptide and M9-peptide. (A) A High resolution mass spectrum of the glycopeptide after ER assay with UDP-[1-

^{13}C]-Glc and deoxynojirimycin (DNJ). The signals corresponding a and b are M9-peptide and $[1-^{13}\text{C}]$ -Glc-M9-peptide, respectively. (B) and (D) are the expanded signals corresponding signal a and b shown in (A), respectively. (F) A High resolution mass spectrum of the glycopeptide after ER assay with UDP- $[1-^{13}\text{C}]$ -Glc. The signals corresponding c and d are M9-peptide and $[1-^{13}\text{C}]$ -Glc-M9-peptide derived from ER assay in the absence of DNJ, respectively. (G) and (I) are the expanded signals corresponding signal c and d shown in (F), respectively. (C) and (H) are simulation spectra of M9-peptide. (E) and (J) are simulation spectra of $[1-^{13}\text{C}]$ -Glc-M9-peptide. A comparison of spectrum of M9-peptide measured with a simulation of $[1-^{13}\text{C}]$ -Glc-M9-peptide is shown in Fig. S6. Molecular formula of M9-peptide and $[1-^{13}\text{C}]$ -Glc-M9-peptide are $\text{C}_{147}\text{H}_{239}\text{N}_{25}\text{O}_{87}\text{S}_2$ and $^{13}\text{C}_{152}\text{C}_{152}\text{H}_{249}\text{N}_{25}\text{O}_{92}\text{S}_2$, respectively. All mass signals were observed as $[\text{M}+3\text{H}]^{3+}$.

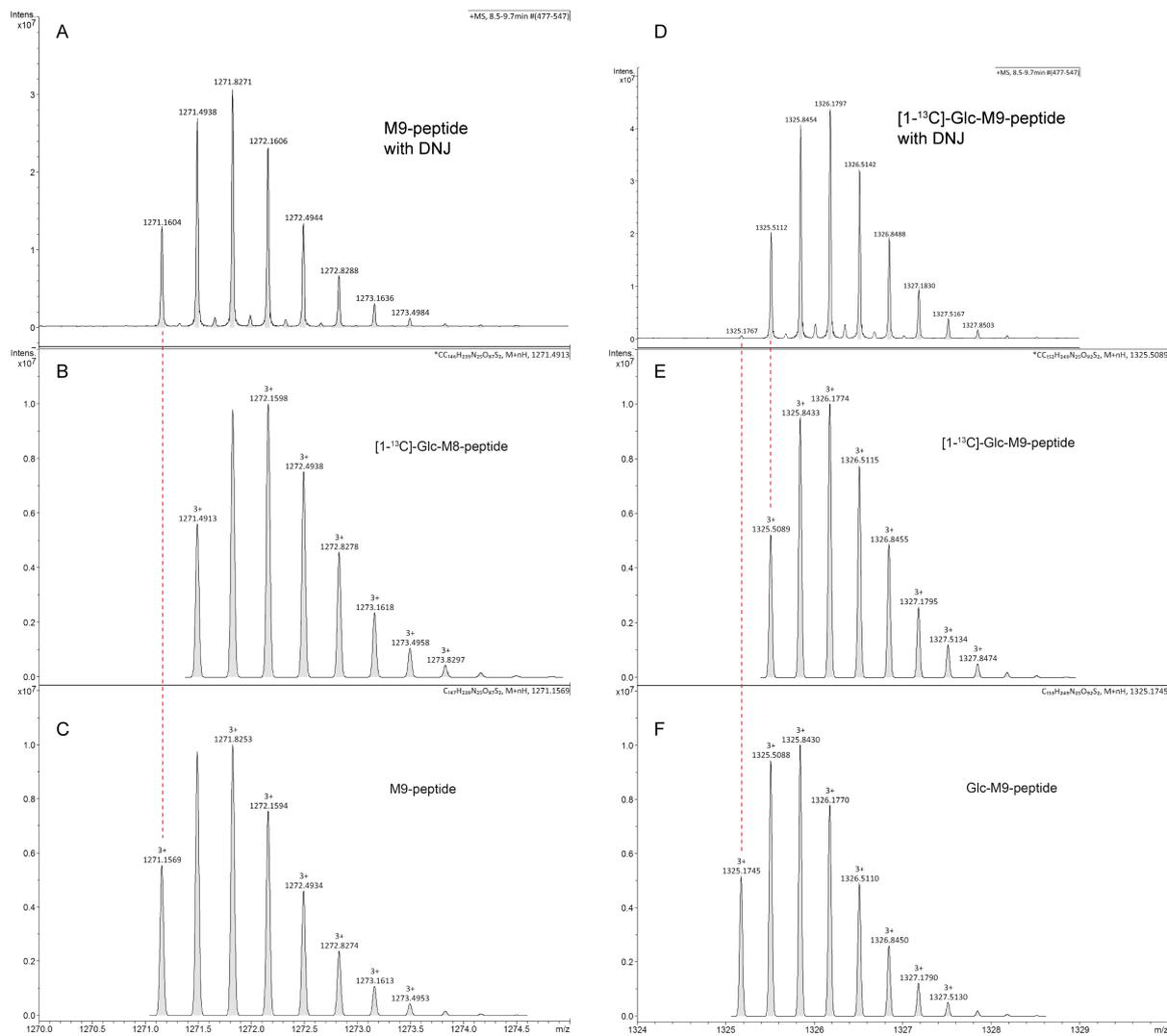


Fig. S6. Simulation of high-resolution mass spectra of G1M9-glycoprotein and $[^{13}\text{C}]$ -G1M9-glycoprotein.

(A) The observed signals of M9-peptide. This signal is identical with the signal-a in Fig. S5A. (B) Simulation signals of $[1-^{13}\text{C}]$ -Glc-M8-peptide. (C) Simulation signals of M9-peptide. (D) The observed signals of $[1-^{13}\text{C}]$ -Glc-M9-peptide. This signal is identical with the signal b in Fig. S5A. (E) Simulation signals of $[1-^{13}\text{C}]$ -Glc-M9-peptide.

peptide. (F) Simulation signals of G1M9-peptide. Molecular formula of M9-peptide, $[1-^{13}\text{C}]\text{-Glc-M9-peptide}$ and $[1-^{13}\text{C}]\text{-Glc-M8-peptide}$ are $\text{C}_{147}\text{H}_{239}\text{N}_{25}\text{O}_{87}\text{S}_2$, $^{13}\text{C}_1\text{C}_{152}\text{H}_{249}\text{N}_{25}\text{O}_{92}\text{S}_2$, and $^{13}\text{C}_1\text{C}_{146}\text{H}_{239}\text{N}_{25}\text{O}_{87}\text{S}_2$, respectively. These data and Fig. S5 clearly proved that GQC yielded M9-glycan and G1M9-glycan during the fast glucosylation and deglycosylation processes.

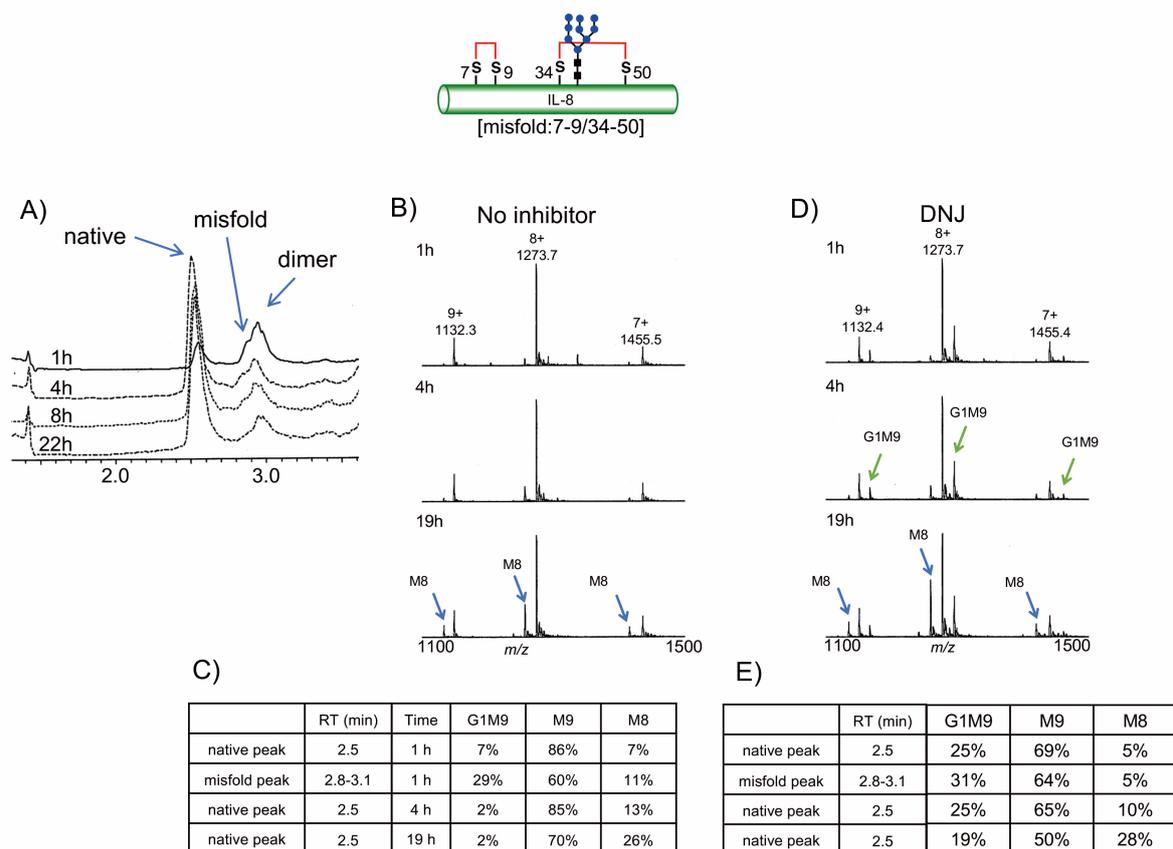


Fig. S7. Monitoring of glucosylation amount during refolding process of misfolded M9-IL8 in the ER lysate.

Refolding reactions were performed with a solution (total 20 μL) containing ER fraction 4 μL (300 ng/ μL), 5 μM glycoprotein, 5 mM inhibitors (DNJ), 0.5 mM UDP-Glc, 50 mM Tris-HCl pH 7.5, 5 mM CaCl_2 , 0.05% and Triton X-100, 37°C. A) Monitoring refolding reaction from misfolded M9-IL8 by reverse phase LS-MS. B) Mass spectra sampling at 1, 4, 19 h in the absence of inhibitor DNJ. C) G1M9-, M9 and M8-glycoprotein ratio at individual peak in the absence of inhibitor DNJ. The ratio was determined by the intensity of mass signals. D) Mass spectra sampling at 1, 4, 19 h in the presence of inhibitor DNJ. E) G1M9-, M9 and M8-glycoprotein ratio at individual peak in the presence of inhibitor DNJ. The ratio was determined by the intensity of mass signals.

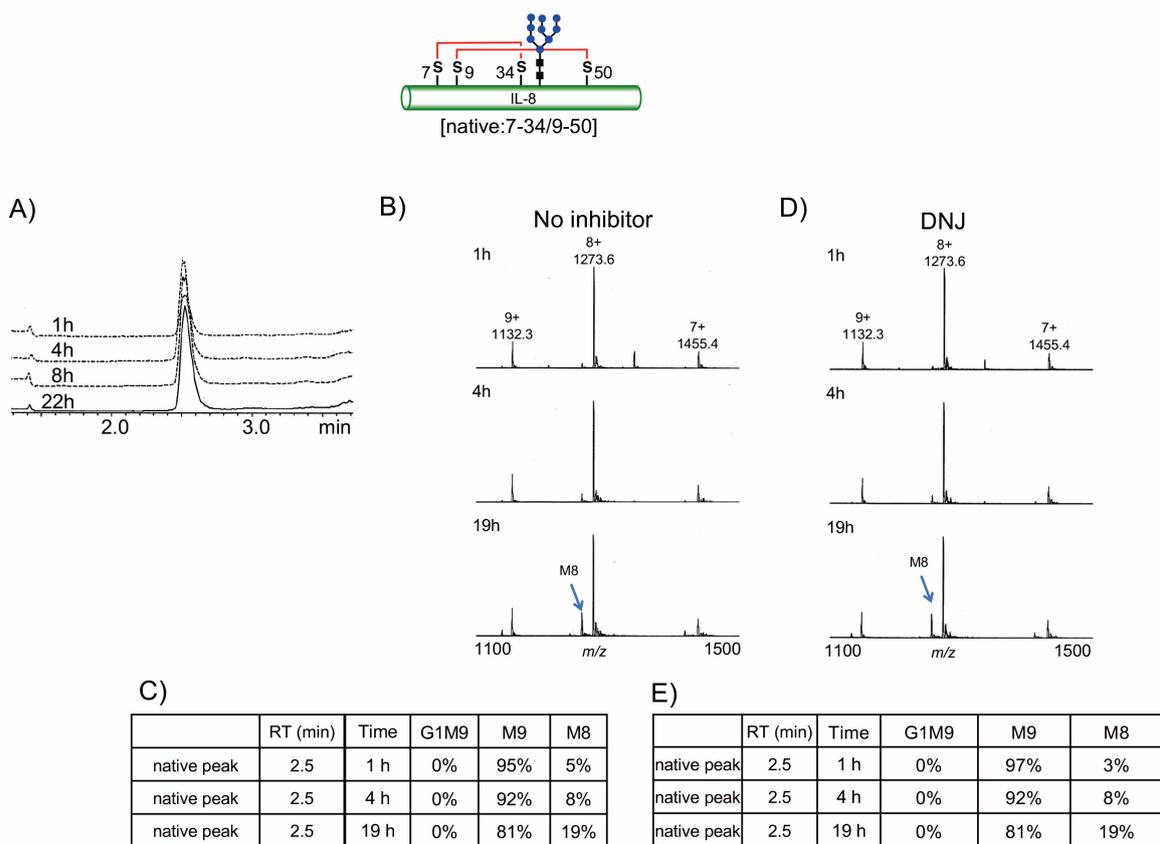


Fig. S8. Monitoring of glycosylation amount of native M9-IL8 in the ER lysate. Reactions were performed with a solution (total 20 μ L) containing ER fraction 4 μ L (300 ng/ μ L), 5 μ M glycoprotein, 5 mM inhibitors (DNJ), 0.5 mM UDP-Glc, 50 mM Tris-HCl pH 7.5, 5 mM CaCl₂, 0.05% and Triton X-100, 37°C. A) Monitoring reaction of native M9-IL8 by reverse phase LS-MS. B) Mass spectra sampling at 1, 4, 19 h in the absence of inhibitor DNJ. C) G1M9-, M9 and M8-glycoprotein ratio at individual peak in the absence of inhibitor DNJ. The ratio was determined by the intensity of mass signals. D) Mass spectra sampling at 1, 4, 19 h in the presence of inhibitor DNJ. E) G1M9-, M9 and M8-glycoprotein ratio at individual peak in the presence of inhibitor DNJ. The ratio was determined by the intensity of mass signals.

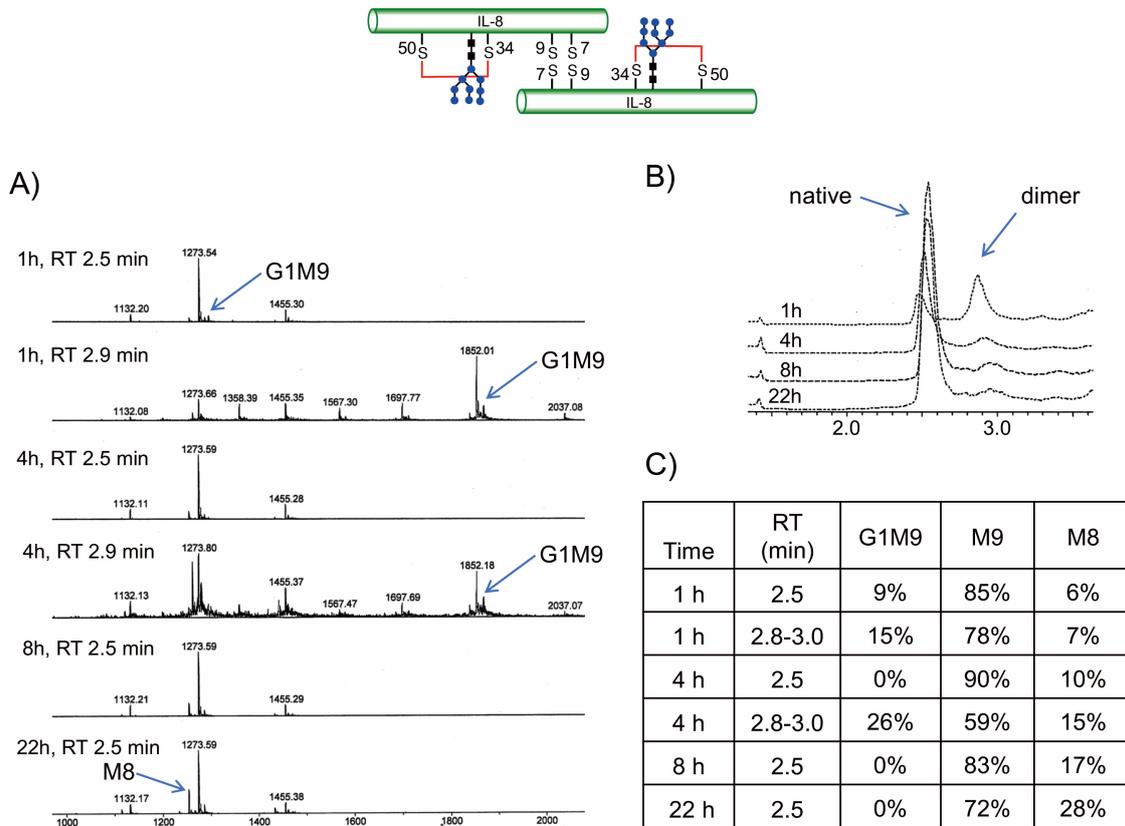


Fig. S9. Monitoring of glycosylation amount of refolding process of dimeric misfolded M9-IL8 in the ER lysate. Reactions were performed with a solution (total 20 μ L) containing ER fraction 4 μ L (300 ng/ μ L), 5 μ M glycoprotein, 5 mM inhibitors (DNJ), 0.5 mM UDP-Glc, 50 mM Tris-HCl pH 7.5, 5 mM CaCl₂, 0.05% and Triton X-100, 37°C. B) Monitoring reaction of native M9-IL8 by reverse phase LS-MS. A) Mass spectra sampling at 1, 4, 19 h in the absence of inhibitor DNJ. C) G1M9-, M9 and M8-glycoprotein ratio at individual peak in the absence of inhibitor DNJ. The ratio was determined by the intensity of mass signals.

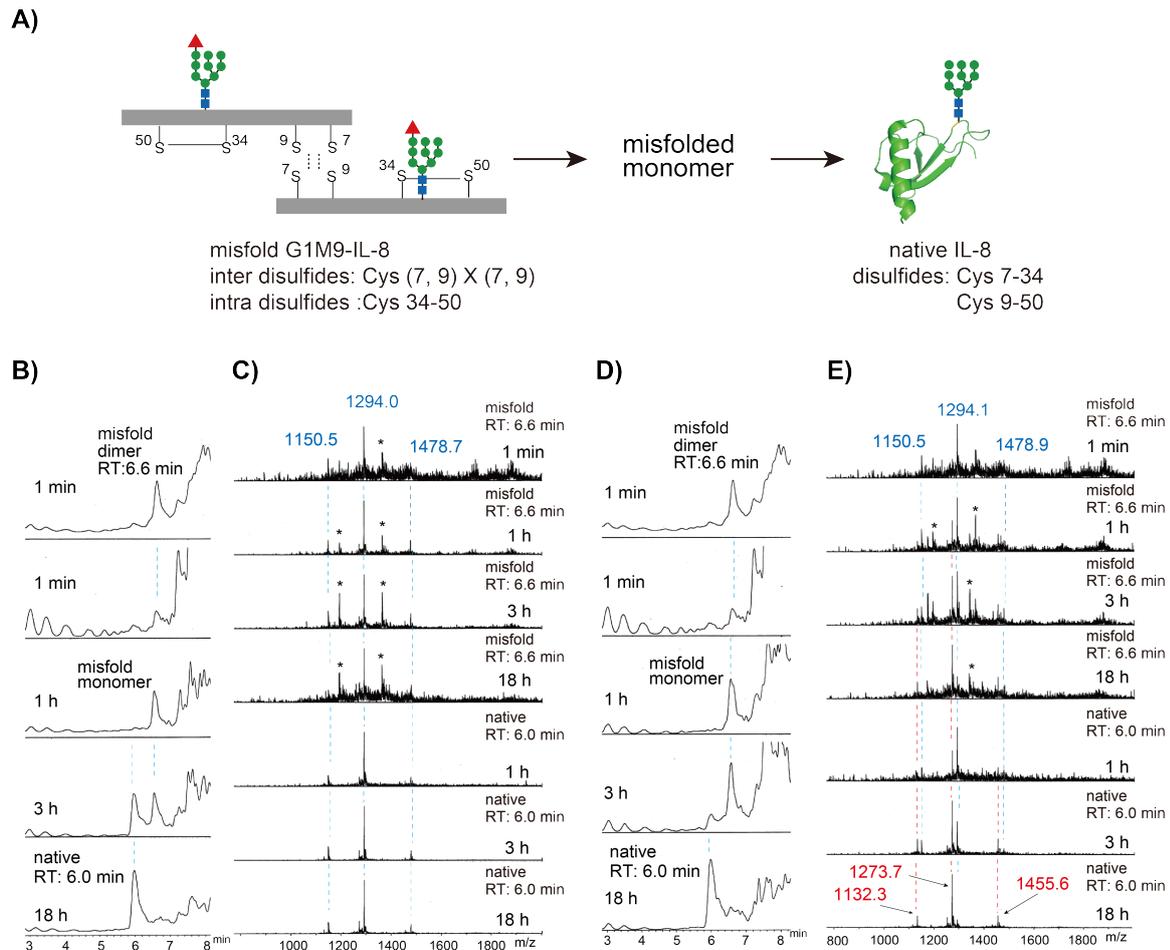


Fig. S10. Relationship between glucosylation/de-glucosylation and refolding ability of misfolded dimeric G1M9-IL8 in the ER lysate. (A) Refolding of misfolded dimeric G1M9-IL8 to native M9-IL8 in the ER lysate. Synthesis of misfolded dimeric G1M9-IL8 was previously reported (ref 2b) (B) Monitoring of refolding by LC in the presence of glucosidase-II inhibitor (DNJ). The peak observed at 6.6 min and 6.0 min were misfolded dimeric G1M9-IL8 and native G1M9-IL8. Top and the second HPLC profile at 1 min reaction time were observed by total ion chromatography and base peak chromatography. Other profiles were by base peak chromatography. Peaks observed around 7-9 min are unknown ER proteins. (C) Monitoring of refolding by LC-MS in the presence of glucosidase-II inhibitor (DNJ). The top four mass spectra were observed at 6.6 min retention time. Blue mass values and dotted line correspond with both misfolded and native G1M9-IL8s. (D) Monitoring of refolding by LC in the absent of glucosidase-II inhibitor (DNJ). Reaction conditions and monitoring system were similar to that of experiments shown in B and C. Top and the second HPLC profile at 1 min reaction time were observed by total ion chromatography and base peak chromatography. Other profiles were by base peak chromatography. Peaks observed around 7-9 min are unknown ER proteins. (E) Monitoring of refolding by LC-MS in the absent of glucosidase-II inhibitor (DNJ). The top four mass spectra were observed at 6.6 min retention time. Blue mass values and blue dotted line correspond with misfolded and native G1M9-IL8s. Red mass values and red dotted line correspond with native M9-IL8s. Misfolded M9-IL8 was not observed at 6.6 retention time. Asterisks * indicate impurities.

Synthesis:

General and abbreviation.

2-(*1H*-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 1-hydroxybenzotriazole (HOBt) and Boc-amino acids were purchased from Peptide Institute Inc. Boc-Cys(Trt) was purchased from Watanabe Chemical Ind. S-Trityl-mercaptopropionic acid was purchased from Oakwood Products Inc. Trifluoroacetic acid (TFA), triisopropylsilane (TIPS), 1,2-ethanedithiol (EDT), sodium 2-mercaptoethanesulfonate (MESNa), N,N'-diisopropylcarbodiimide (DIC), N,N-diisopropylethylamine (DIPEA), dibutyldicarbonate (Boc₂O) and tris(2-carboxyethyl)phosphine hydrochloride (TCEP) were purchased from Tokyo Chemical Ind. N,N-dimethylformamide (DMF), dichloromethane (DCM), dimethyl sulfoxide (DMSO), piperidine, m-cresol, thioanisole, trifluoromethanesulfonic acid (TfOH), diethyl ether, dimethylsulfide (DMS), 2,2'-Azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride (VA-044), Cystine, Glutathione (reduced), Glutathione-disulfide and 2-amino-2-hydroxymethyl-1,3-propanediol (Tris) were purchased from Wako Pure Chemical. Fmoc-amino acids, Boc-Arg(di-Z), Boc-His(DNP), Boc-Asn(Xan), 1-(mesitylene-2-sulfonyl)-3-nitro-*1H*-1,2,4-triazole (MSNT), amino-PEGA resin, Benzotriazole-1-yl-oxy-trispyrrolidino-phosphonium hexafluorophosphate (PyBOP) and 4-(4-Hydroxymethyl-3-methoxyphenoxy)-butyric acid (HMPB) were purchased from Novabiochem. Guanidine hydrochloride (Gn·HCl), HPLC grade acetonitrile and L-Cysteine hydrochloride hydrate were purchased from Kanto Chemical Co. Inc. 2-Bromoaceto-phenone, 3-(diethoxyphosphoryloxy)-1,2,3-benzotriazin-4(3*H*)-one (DEPBT) and 4-mercaptophenylacetic acid (MPAA) were purchased from Sigma-Aldrich.

General procedure for Boc solid phase peptide synthesis (SPPS) (ref 6)

Boc-amino acids used were Gly, Ala, Val, Leu, Ile, Pro, Ser(Bzl), Thr(Bzl), Met, Asp(Bzl), Glu(Bzl), Asn(Xan), Gln, His(DNP), Lys(Cl-Z), Arg(di-Z), Arg(Mtr), Phe, Tyr(Br-Z), Trp(CHO), Cys(Acm), Cys(MBzl) and Thz.

Peptide- α -thioesters were prepared using improved *in situ* neutralization Boc SPPS protocol on Nova PEG amino resin (ref 6). Nova PEG amino resin was pre-washed by MeOH (1min x 2), DMF (1min x 2), DCM (1min x 3), 1%TFA/DCM (1min x 3), 5% DIPEA/DCM (1min x 3) and DCM (1min x 3). HBTU (72.1 mg, 0.19 mmol), DIPEA (70.0 μ L, 0.4 mmol) and S-trityl-3-mercaptopropionic acid (80.0 mg, 0.2 mmol) were dissolved in DMF (1.0 mL) and pre-activated for 30 seconds. This solution was added to Nova PEG amino resin (50 μ mol) to incorporate 3-mercaptopropionic acid. S-Trityl group was removed by treatment of the resin with TFA for 1.5 min twice. The first Boc amino acid (Boc-AA, 0.2 mmol) was pre-activated with HBTU (72.1 mg, 0.19 mmol) and DIPEA (70.0 μ L, 0.4 mmol) in DMF

(1.0 mL) for 30 seconds and then the solution was added to the resin and the suspension was gently shaken for 20 min at ambient temperature. After the coupling of first amino acid, the following Boc-AA couplings were performed by same manner with the first Boc-AA coupling. For the Boc deprotection, the peptidyl resin was treated with TFA for 1.5 min twice followed by careful washing with DCM (ca. 70-100 mL) and DMF (ca. 70-100 mL).

After the coupling of the last amino acid residue, the resin was washed with DMF and DCM. To the resin (ca. 10 μ mol) was added a solution of TFA (1.2 mL), *m*-cresol (0.12 mL) and DMS (0.36 mL) at 0°C followed by slow addition of TfOH (0.12 mL). Then the suspension was shaken for 1 hour at 0°C. The suspension was filtered and then the resulting resin was treated again with the same manner to complete the side-chain deprotection. In the case of incompleteness of the deprotection by this process, strong acidic conditions employing TFA : TfOH : EDT : thioanisole = 20 : 2 : 1 : 2 for 1 hour at 0°C were able to use. However this strong condition was only for peptide- α -thioester and not for glycopeptide- α -thioester. An aliquot of the resin was treated with sodium phosphate buffer (0.2 M, pH 7.0, 2.0 mL) containing 6 M Gn-HCl and 0.2 M MESNa for 12 h to perform thiolysis at the ambient temperature. Purification of the resultant solution containing peptide- α -thioester by preparative HPLC (Vydac C4, C8, C18 or Proteonavi or Capcell pak (Shiseido) Φ 10 \times 250 mm, 0.1% TFA : 0.1% TFA in 90% MeCN at moderate ratio, over 30 min at the flow rate of 2.5 mL/min) afforded peptide- α -thioester.

Preparation of Boc-M9-high-mannose type oligosaccharide (Boc-Asn (M9))

Preparation of Fmoc M9-high-mannose type oligosaccharide (Fmoc-Asn (M9)) was following the method that we have already reported (ref 7). Prepared Fmoc-M9 (55 mg, 24.8 μ mol) was dissolved to DMSO-DMF (4 / 1, v / v, 3.0 ml). Piperidine (750 μ l, 7.59 mmol) was added to the solution for deprotection of Fmoc group. After 1 h, the solution was poured into chilled Et₂O, and the resultant white precipitate was collected. This precipitate was dissolved to water and lyophilized. The resultant white foam was dissolved to DMSO-DMF (4 / 1, v / v, 3.0 ml). DIPEA (41 μ l, 230 μ mol) and Boc₂O (52 μ l, 226 μ mol) were added to the solution. After 1 h, the solution was poured into chilled Et₂O, and the resultant white precipitate was collected. This precipitate was dissolved to water and lyophilized. The resultant white foam was dissolved to water and applied to ion-exchange column (Dowex W50 \times 8 i.d. 10 \times 80 mm, eluted with H₂O). The resultant fraction was lyophilized and Boc-Asn (M9) was yielded 43 mg as white foam. The yield was ca. 82.7% as the isolated yield.

Preparation of glycopeptide- α -thioester bearing M9-high-mannose type oligosaccharide (ref 7)

Peptide- α -thioesters on the resin before coupling of high-mannose-M9-oligosaccharide were prepared using general procedure for Boc SPPS as described above (ca. 4 μ mol scale). Coupling of M9-high-mannose type oligosaccharide (8.4 mg, 4 μ mol) was performed by using DEPBT (2.4 mg, 8.0 μ mol) or

PyBOP (4 mg, 8 μ mol) and DIPEA (8.7 μ L, 49 μ mol) in DMF-DMSO (1/1, v/v, 133 μ L) for 14 h at ambient temperature. After this coupling step, following Boc deprotection were performed according to general procedure. The coupling of next Boc-AA was performed by using diluted condition such as Boc-AA (40 μ mol, 40 mM), HBTU (38 μ mol) and DIPEA (80 μ mol) in DMF (1 mL) for 4 μ mol scale synthesis. This coupling was performed for 20 min. After the coupling of the last amino acid residue, the protecting groups of side chain were deprotected. The glycopeptide- α -thioester was cleaved from the resin and analyzed by the same method mentioned above. Then the desired compound was obtained by purification by RP-HPLC.

Typical procedure for the synthesis of EPO derivatives: Synthesis of EPO-24, 38, 83 (compound 1 and compound 5)

First ligation (ref 8). Peptide segment F (compound **10**, 20.0 mg, 4.3 μ mol) was dissolved in a ligation buffer (1070 μ L) containing 6 M Gn-HCl, 200 mM sodium phosphate, 40 mM MPAA (7.1 mg, 42 μ mol) and 20 mM TCEP (6.1 mg, 21 μ mol) at pH 6.8. The peptide- α -thioester segment E (compound **9**, 14.0 mg, 4.4 μ mol) was added to the solution and this solution was left for 3 h at ambient temperature. The reaction was monitored by RP-HPLC (Cadenza CD-C18 Φ 4.6 \times 10 mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 70 : 30 to 30 : 70 for 15 min by 1 mL/min flow rate). After completion of the reaction, *N*-terminal thiazolidine (Thz) moiety of the product was converted into Cys residue as a one-pot reaction by adjusting pH to 4.0 with methoxyamine-HCl. After 2 h, the reaction mixture was subjected to RP-HPLC purification (Vydac C18 Φ 10 \times 250 mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 100 : 0 to 62 : 38 to 30 : 70 for 35 min at the flow rate of 3 mL/min). The product (glycopeptide segment EF, compound **11**) thus obtained was characterized by ESI-MS. This product was lyophilized and afforded 14.0 mg as white foam. The yield of the first NCL between segment E (compound **9**) and F (compound **10**) was ca. 41.8% as the isolated yield. ESI-MS: *m/z* calcd. for C₃₃₄H₅₅₅N₁₀₀O₉₇S₃: [M+H]⁺ 7619.8, found 7620.2 (deconvoluted).

Second ligation. The M9-glycopeptide- α -thioester segment D (compound **12**, 5.6 mg, 1.05 μ mol) was dissolved in a ligation buffer (524 μ L) containing 6 M Gn-HCl, 200 mM sodium phosphate, 40 mM MPAA (3.5 mg, 21 μ mol) and 20 mM TCEP (3.0 mg, 10 μ mol) at pH 6.9. The peptide segment EF (compound **11**, 7.9 mg, 1.04 μ mol) was added to the solution and this solution was left for 4 h at ambient temperature. The reaction was monitored by RP-HPLC (Proteonavi Φ 4.6 \times 250 mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 70 : 30 to 35 : 65 for 30 min by 1 mL/min flow rate). After completion of the reaction, *N*-terminal thiazolidine (Thz) moiety of the product was converted into Cys residue followed with the same manner as mentioned above. After 2 h, the reaction mixture was subjected to RP-HPLC purification (Vydac C4 Φ 10 \times 250 mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 100 : 0 to 57: 43 to 40 :

60 for 35 min by 2.5 mL/min flow rate). The product (glycopeptide segment DEF, compound **14**) thus obtained was characterized by ESI-MS. This product was lyophilized and afforded 5.0 mg as white foam. The yield of the second NCL between segment D (compound **12**) and EF (compound **11**) was ca. 37.6% as the isolated yield. ESI-MS: m/z calcd. for $C_{553}H_{912}N_{143}O_{195}S_4$: $[M+H]^+$ 12812.3, found 12812.5 (deconvoluted).

Third ligation. The peptide- α -thioester segment C (compound **16**, 1.5 mg, 0.66 μ mol) was dissolved in a ligation buffer (195 μ L) containing 6 M Gn-HCl, 200 mM sodium phosphate, 40 mM MPAA (1.3 mg, 7.8 μ mol) and 20 mM TCEP (1.1 mg, 7.8 μ mol) at pH 6.9. The glycopeptide segment DEF (compound **14**, 5 mg, 0.39 μ mol) was added to the solution and this solution was left for 2.5 h at ambient temperature. The reaction was monitored by RP-HPLC (Protonavi $\Phi 4.6 \times 250$ mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 70 : 30 to 30 : 70 for 30 min by 1 mL/min flow rate). After completion of the reaction, *N*-terminal thiazolidine (Thz) moiety of the product was converted into Cys residue followed with the same manner as mentioned above. After 2 h, the reaction mixture was subjected to RP-HPLC purification (Vydac C4 $\Phi 10 \times 250$ mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 100 : 0 to 52 : 48 to 35 : 65 for 35 min by 2.5 mL/min flow rate). The product (glycopeptide segment CDEF, compound **17**) thus obtained was characterized by ESI-MS. This product was lyophilized and afforded 3.6 mg as white foam. The yield of the third NCL between segment C (compound **16**) and DEF (compound **14**) was ca. 61.7% as the isolated yield. ESI-MS: m/z calcd. for $C_{648}H_{1057}N_{170}O_{220}S_6$: $[M+H]^+$ 14941.8, found 14942.8 (deconvoluted).

Fourth ligation. The M9-glycopeptide- α -thioester segment B (compound **19**, 1.2 mg, 0.26 μ mol) was dissolved in a ligation buffer (120 μ L) containing 6 M Gn-HCl, 200 mM sodium phosphate, 40 mM MPAA (0.81 mg, 4.8 μ mol) and 40 mM TCEP (1.37 mg, 4.8 μ mol) at pH 7.0. The M9-glycopeptide segment CDEF (compound **17**, 3.6 mg, 0.24 μ mol) was added to the solution and this solution was left for 5 h at ambient temperature. The reaction was monitored by RP-HPLC (Protonavi $\Phi 4.6 \times 250$ mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 60 : 40 to 30 : 70 for 30 min by 1 mL/min flow rate). After completion of the reaction, the reaction mixture was subjected to RP-HPLC purification (Protonavi $\Phi 10 \times 250$ mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 100 : 0 to 55 : 45 to 35 : 65 for 35 min by 2.5 mL/min flow rate). The product (glycopeptide segment BCDEF, compound **21**) thus obtained was characterized by ESI-MS. This product was lyophilized and afforded 2.0 mg as white foam. The yield of the fourth NCL between segment B (compound **19**) and CDEF (compound **17**) was ca. 47.6% as the isolated yield. ESI-MS: m/z calcd. for $C_{827}H_{1338}N_{201}O_{311}S_8$: $[M+H]^+$ 19329.2, found 19329.2 (deconvoluted).

Desulfurization of segment BCDEF (ref 9) (EPO (29-166)-M9-glycopeptide, compound 21)

M9-glycopeptide BCDEF (compound **21**, 1.0 mg, 0.052 μmol) was dissolved in a 200 mM sodium phosphate buffer (pH 7.0, 220 μL) containing 7.5 M Gn-HCl and 0.25 M TCEP. To the solution, 2-methyl-2-propanethiol (19.8 μL), 0.1 M VA-044 (39.8 μL) was added. The reaction mixture was stirred for 3 h at 37°C. The reaction was monitored by RP-HPLC (Proteonavi $\Phi 4.6 \times 250$ mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 55 : 45 to 32 : 68 for 30 min at the flow rate of 1 mL/min). After completion of the reaction, the reaction mixture was subjected to RP-HPLC purification (Proteonavi $\Phi 4.6 \times 250$ mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 100 : 0 to 55 : 45 to 38 : 62 for 35 min by 1 mL/min flow rate). The product (Reduced M9-glycopeptide segment BCDEF, compound **25**) thus obtained was characterized by ESI-MS. This product was lyophilized and afforded 0.5 mg as white foam. The yield of the reduction of segment BCDEF (compound **21**) was ca. 50.4% as the isolated yield. ESI-MS: m/z calcd. for $\text{C}_{827}\text{H}_{1338}\text{N}_{201}\text{O}_{311}\text{S}_4$: $[\text{M}+\text{H}]^+$ 19200.9, found 19200.9 (deconvoluted).

Deprotection of Acetamidomethyl (Acm) group of segment BCDEF (EPO (29-166)-M9-glycopeptide, compound 25)

The desulfurized M9-glycopeptide BCDEF (compound **25**, 0.65 mg, 0.034 μmol) was dissolved in 90% acetic acid solution (141 μL) containing AgOAc (0.82 mg) and the mixture was stirred for 4 h at ambient temperature. After centrifugation of the reaction mixture, the supernatant was collected in a new tube and was added DTT (0.95 mg). After stirring for 5 min, the resultant precipitate was removed by centrifugation and the supernatant was subjected to RP-HPLC purification (Proteonavi $\Phi 10 \times 250$ mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 100 : 0 to 55 : 45 to 38 : 62 for 35 min at the flow rate of 2.5 mL/min). The purified M9-glycopeptide BCDEF (compound **29**) was characterized by ESI-MS. This product was lyophilized and afforded 0.4 mg as white foam. The yield of the segment BCDEF (compound **29**), which was deprotected Acm groups was ca. 62.5% as the isolated yield. ESI-MS: m/z calcd. for $\text{C}_{818}\text{H}_{1323}\text{N}_{198}\text{O}_{308}\text{S}_4$: $[\text{M}+\text{H}]^+$ 18987.7, found 18987.4 (deconvoluted).

Final ligation. The M9-glycopeptide- α -thioester segment A (compound **33**, ca. 0.4 mg, 0.78 μmol) was dissolved in a ligation buffer (20 μL) containing 6 M Gn-HCl, 200 mM sodium phosphate, 40 mM MPAA (ca. 0.13 mg, 0.8 μmol) and 40 mM TCEP (ca. 0.23 mg, 0.8 μmol) at pH 7.0. The glycopeptide segment BCDEF (compound **29**, 0.8 mg, 0.042 μmol) was added to the solution and this solution was left for 10 h at ambient temperature. The reaction was monitored by RP-HPLC (Proteonavi $\Phi 4.6 \times 250$ mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 70 : 30 to 20 : 80 for 30 min by 1 mL/min flow rate). After completion of the reaction, the reaction mixture was subjected to RP-HPLC purification (Proteonavi $\Phi 10 \times 250$ mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 100 : 0 to 55 : 45 to 38 : 62 for 35 min by 2.5 mL/min flow rate). The product (EPO glycopeptide segment ABCDEF, compound **35**) thus obtained

was characterized by ESI-MS. This product was lyophilized and afforded as white foam. ESI-MS: m/z calcd. for $C_{1025}H_{1666}N_{239}O_{406}S_5$: $[M+H]^+$ 23993.9, found 23995.7 (deconvoluted).

Folding of EPO M9-glycosylated polypeptide (ref 10) (compound 35)

Folding of EPO M9-glycosylated polypeptide ABCDEF (compound 35) was performed by use of stepwise dialysis method. EPO M9-glycosylated polypeptide ABCDEF (compound 35) was dissolved in a Tris-HCl buffer (100 mM, pH 7.5) containing 6 M Gn-HCl. The concentration of the EPO M9-glycosylated polypeptide ABCDEF (compound 35) was adjusted to suitable concentration (ca. 0.1 mg/mL). This solution was poured into the dialysis tubing (MWCO at 8,000 Spectra/Por®) and then dialyzed against the first folding buffer (3 M Gn-HCl, 100 mM Tris-HCl, pH 8.5) containing 4 mM cysteine and 0.5 mM cystine for redox system and left for 12 h at 4°C. Then, the external buffer solution was replaced to the second folding buffer solution (1 M Gn-HCl, 100 mM Tris-HCl, pH 8.0) and dialysis was performed for 20 h. Finally, the external buffer was discarded and changed again to the third folding buffer solution (10 mM Tris-HCl, pH 7.0) The dialysis was performed for 24 h. The correctly folded and misfolded EPO analogs were purified by RP-HPLC (Protonavi $\Phi 10 \times 250$ mm, 0.1% TFA : 0.1% TFA in 90% MeCN = 100 : 0 to 60 : 40 to 30 : 70 for 35 min at the flow rate of 2.5 mL/min). Fractions containing correctly folded EPO 1 and misfolded EPO 2 were collected and then lyophilized. The purity of correctly folded EPO 1 and misfolded EPO 2 was confirmed by HPLC and ESI-MS spectroscopy. The isolated amounts of correctly folded EPO 1 and misfolded EPO 2 were ca. 124.7 μ g and ca. 25.8 μ g respectively, estimated by the Bradford protein assay. FT-ICR-MS: m/z calcd. for $C_{1025}H_{1662}N_{239}O_{406}S_5$: $[M+H]^+$ 23989.5730, found 23989.5022 (deconvoluted).

Disulfide mapping of correctly folded and misfolded EPO analogs (compound 1-8)

In order to determine the position of disulfide bonds, we employed a standard strategy using trypsin digestion. The correctly folded or misfolded EPO analogs (compound 1-8, ca. 1.4 μ g) were dissolved in phosphate buffer (pH 7.0, 100 mM, 50 μ L), and then an appropriate amount of trypsin (ca. 0.1 μ g) was added. After 11 h, the resultant fragments were analyzed by LC-MS. In order to determine whether the peptide fragments include disulfide bonds, the aliquot of this solution was also treated by 20 mM TCEP for 1 h and then subjected to analytical LC-MS to compare HPLC profile of non-reduced peptide fragments. Fig. S49-S56 show HPLC profiles of both non-reduced peptide fragments and reduced peptide fragments.

Measurement of Circular dichroism (CD) spectra of correctly folded and misfolded EPO analogs (compound 1-8)

Far-UV CD spectra of correctly folded and misfolded EPO analogs (compound 1-8) were measured with

a JASCO-j820 CD spectropolarimeter. The concentrations of EPO analogs were ca. 1.7 μ M.



Fig. S11. Amino acids sequence of erythropoietin

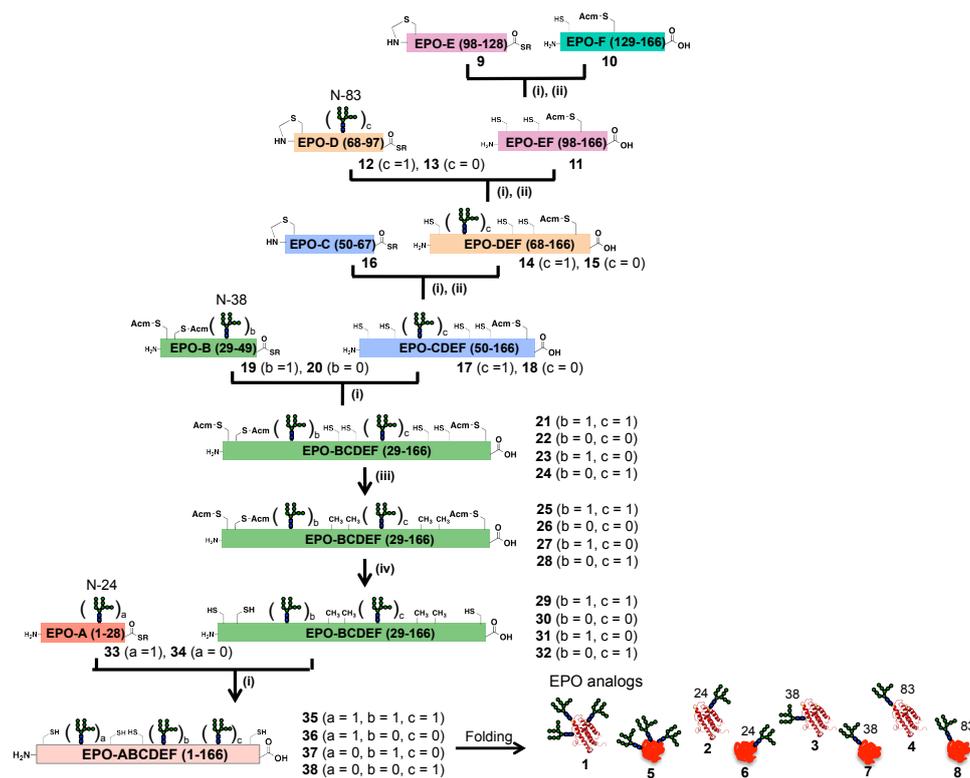


Fig. S12 Synthetic strategy of erythropoietin glycoforms. Conditions: (i) native chemical ligation; (ii) removal of thiazolidine; (iii) desulfurization; (iv) removal of the Acm groups. Small a, b or c = 1 and a, b or c = 0 indicate glycosylation and no-glycosylation, respectively.

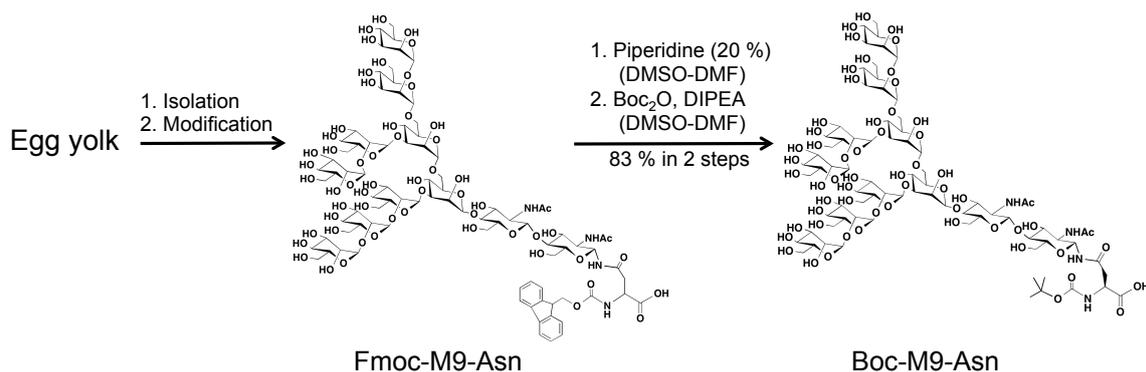


Fig. S13. Isolation of homogeneous M9-high-mannose type oligosaccharide from egg yolk and modification by Boc group. Isolation method has been published (ref 7). DMSO; dimethyl sulfoxide, DMF; dimethylformamide, Boc₂O; di-tert-butyl dicarbonate, DIPEA; diisopropylethylamine

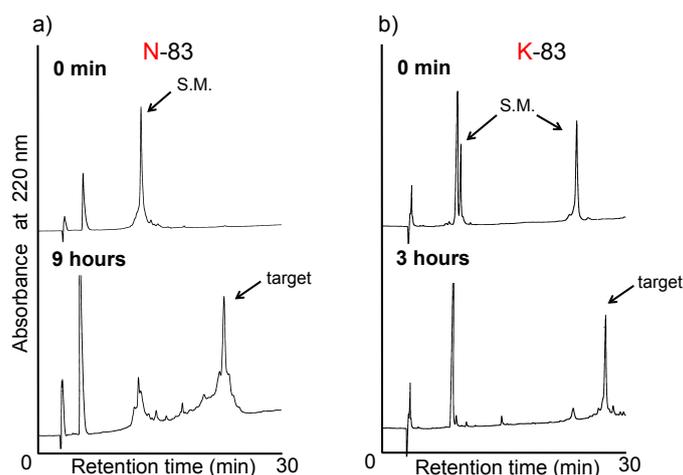


Fig. S14. Hydrophobicity of EPO segment (50-166 peptide) and their mutant N83K. Because non-glycosylated peptide showed low solubility in buffer solution, we changed Asn at the non-glycosylation sites to Lys in order for peptide to have solubility (ref 11, 12). This figure shows an example of comparison of Asn and Lys mutant. a) After 9 hours for the EPO (50-166) having Asn83. b) After 3 hours for the EPO (50-166) having Lys83. The Mutation of potential glycosylation site from N to K dramatically improved solubility and the HPLC profile of NCL for synthesis of EPO (50-166).

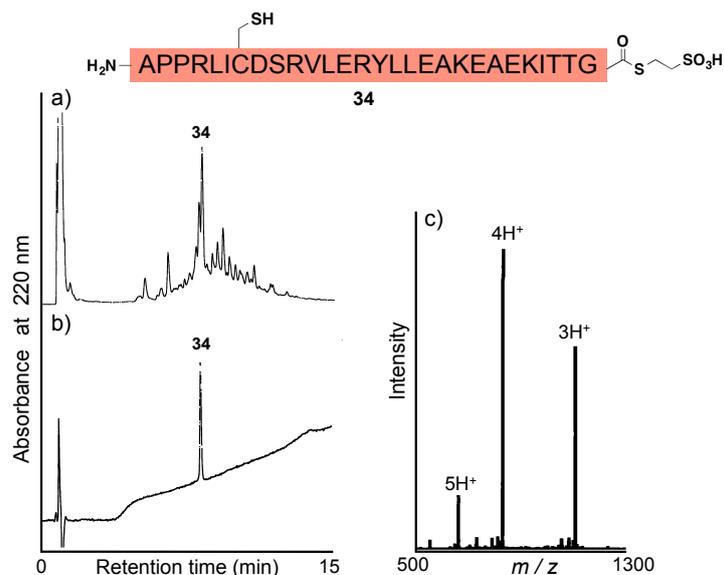


Fig. S15 HPLC profiles and ESI-MS spectrum of purified peptide segment A (ref 13). a) HPLC profile of crude compound **34**. b) HPLC profile of purified compound **34** after detachment of peptide from solid support. c) ESI-MS of compound **34**. ESI-MS: m/z calcd. for $C_{141}H_{240}N_{39}O_{45}S_3$: $[M+H]^+$ 3297.9, found 3297.2 (deconvoluted).

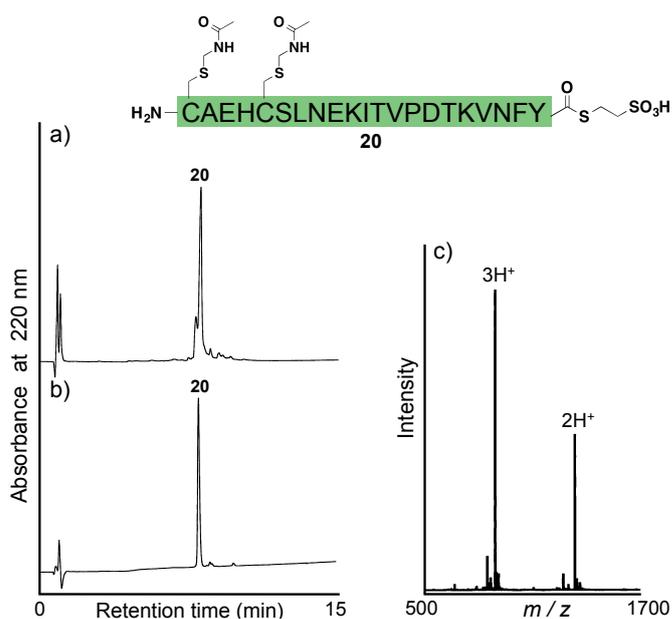


Fig. S16. HPLC profiles and ESI-MS spectrum of purified peptide segment B. a) HPLC profile of crude compound **20**. b) HPLC profile of purified compound **20** after detachment of peptide from solid support. c) ESI-MS of compound **20**. ESI-MS: m/z calcd. for $C_{113}H_{178}N_{29}O_{38}S_4$: $[M+H]^+$ 2677.2, found 2677.8 (deconvoluted).

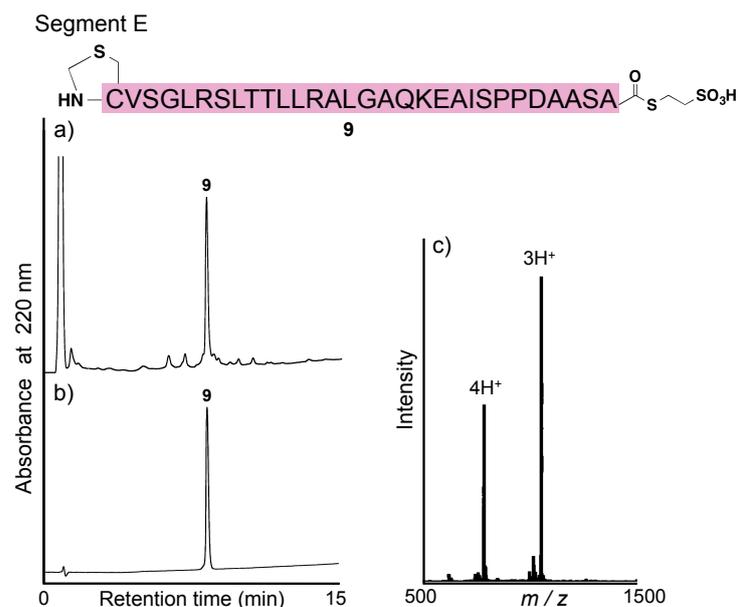


Fig. S19. HPLC profiles and ESI-MS spectrum of purified peptide segment E. a) HPLC profile of crude compound **9**. b) HPLC profile of purified compound **9** after detachment of peptide from solid support. c) ESI-MS of compound **9**. ESI-MS: m/z calcd. for $C_{131}H_{227}N_{38}O_{44}S_3$: $[M+H]^+$ 3134.6 found 3133.6 (deconvoluted).

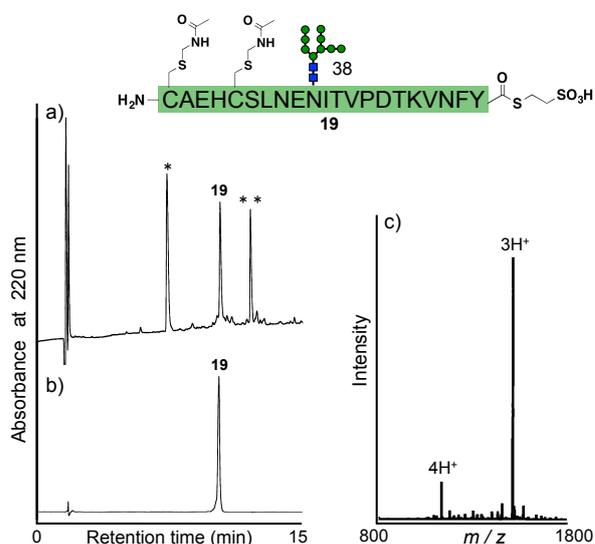


Fig. S20. HPLC profiles and ESI-MS spectrum of purified glycopeptide segment B. a) HPLC profile of crude compound **19**. b) HPLC profile of purified compound **19** after detachment of peptide from solid support. c) ESI-MS of compound **19**. ESI-MS: m/z calcd. for $C_{181}H_{288}N_{31}O_{94}S_4$: $[M+H]^+$ 4530.6 found 4529.4 (deconvoluted). *DNP protecting group removed of histidine during deprotection step. **EPO-B (21-49) lacking a Asn-M9-glycan.

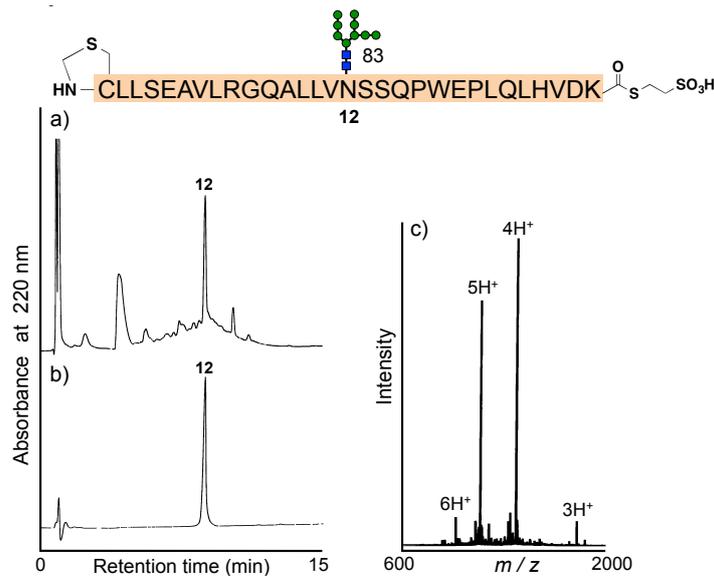


Fig. S21. HPLC profiles and ESI-MS spectrum of purified glycopeptide segment D. a) HPLC profile of crude compound **12**. b) HPLC profile of purified compound **12** after detachment of peptide from solid support. c) ESI-MS of compound **12**. ESI-MS: m/z calcd. for $C_{222}H_{364}N_{43}O_{101}S_3$: $[M+H]^+$ 5346.7 found 5347.0 (deconvoluted).

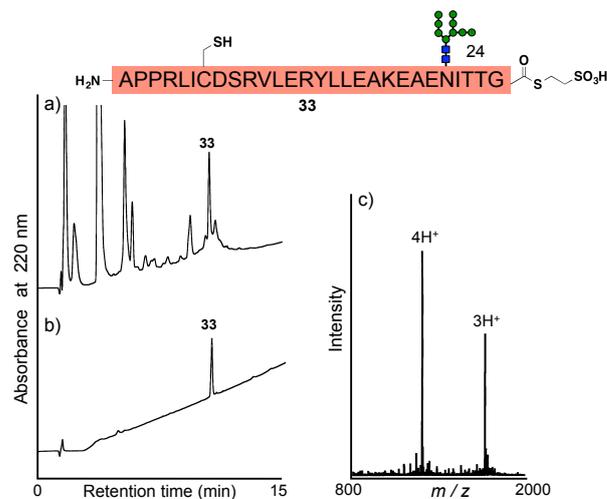


Fig. S22. HPLC profiles and ESI-MS spectrum of purified glycopeptide segment A. a) HPLC profile of crude compound **33**. b) HPLC profile of purified compound **33** after detachment of peptide from solid support. c) ESI-MS of compound **33**. ESI-MS: m/z calcd. for $C_{209}H_{350}N_{41}O_{101}S_3$: $[M+H]^+$ 5149.4 found 5148.3 (deconvoluted).

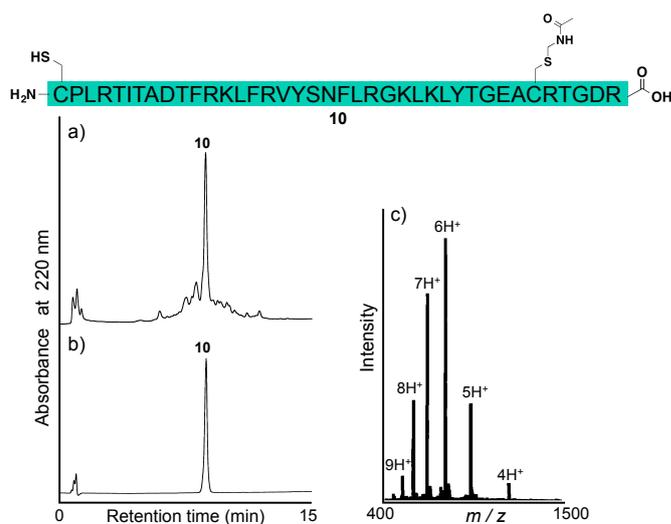


Fig. S23. HPLC profiles and ESI-MS spectrum of purified peptide segment F. a) HPLC profile of crude compound **10**. b) HPLC profile of purified compound **10** after detachment of peptide from solid support. c) ESI-MS of compound **10**. ESI-MS: m/z calcd. for $C_{206}H_{335}N_{62}O_{56}S_2$: $[M+H]^+$ 4640.4 found 4639.9 (deconvoluted).

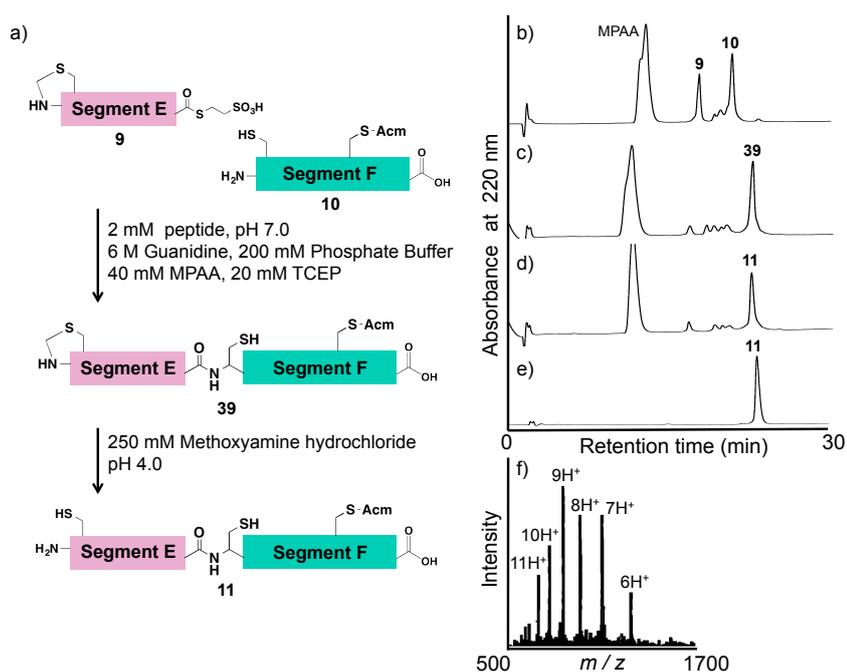


Fig. S24. 1st-NCL for the construction of EPO bearing three M9-glycans. a) Peptide ligation reaction between compound **9** and **10**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 1 h. d) HPLC profile after treatment with methoxyamine-HCl. e) HPLC profile of purified compound **11**. f) ESI-MS of compound **11**. ESI-MS: m/z calcd. for $C_{334}H_{555}N_{100}O_{97}S_3$: $[M+H]^+$ 7619.8, found 7620.2 (deconvoluted).

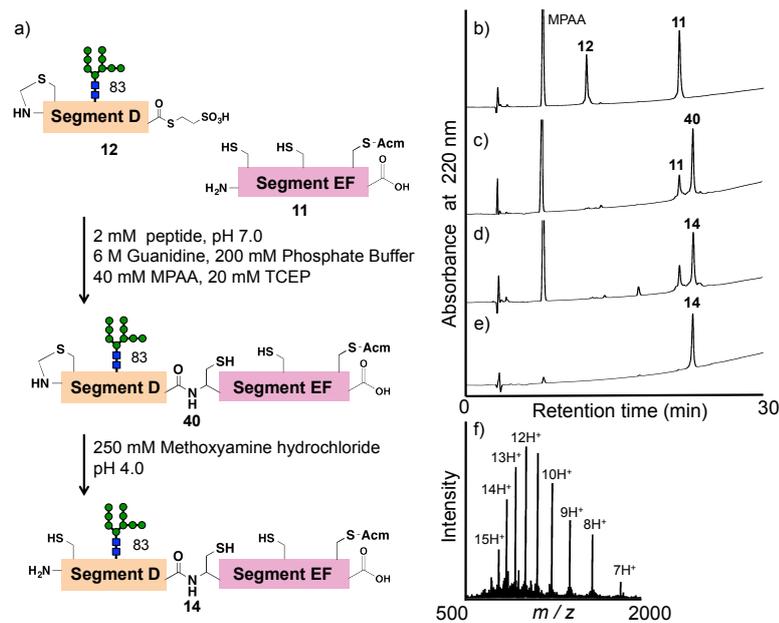


Fig. S25. 2nd-NCL for the construction of EPO bearing three M9-glycans. a) Peptide ligation reaction between compound **11** and **12**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 4 h. d) HPLC profile after treatment with methoxyamine-HCl. e) HPLC profile of purified compound **14**. f) ESI-MS of compound **14**. ESI-MS: m/z calcd. for $C_{553}H_{912}N_{143}O_{195}S_4$: $[M+H]^+$ 12812.3, found 12812.5 (deconvoluted).

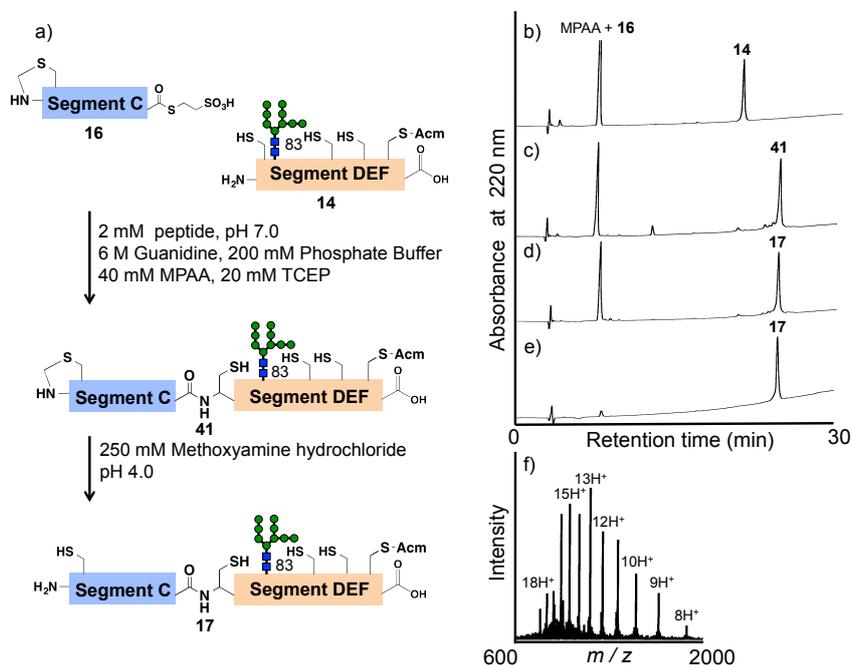


Fig. S26. 3rd-NCL for the construction of EPO bearing three M9-glycans. a) Peptide ligation reaction between compound **14** and **16**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 2.5 h. d) HPLC profile after treatment with methoxyamine-HCl. e) HPLC profile of purified compound **17**. f) ESI-MS of compound **17**. ESI-MS: m/z calcd. for $C_{648}H_{1057}N_{170}O_{220}S_6$: $[M+H]^+$ 14941.8, found 14942.8 (deconvoluted).

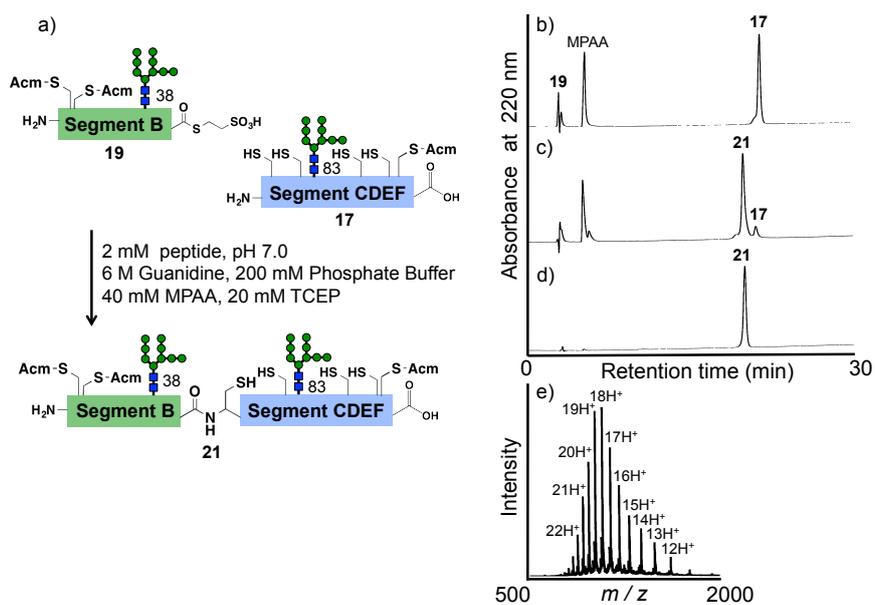


Fig. S27. 4th-NCL for the construction of EPO bearing three M9- glycans. a) Peptide ligation reaction between compound 17 and 19. This analytical condition showed the profile of hydrophilic glycopeptide 19 in the void fraction because of the monitoring of the hydrophobic product. b) HPLC profile for stating point ($t < 1$ min). c) HPLC profile after 5 h. d) HPLC profile of purified compound 21. e) ESI-MS spectrum of compound 21. ESI-MS: m/z calcd. for C₈₂₇H₁₃₃₈N₂₀₁O₃₁₁S₈: [M+H]⁺ 19329.2, found 19329.2 (deconvoluted).

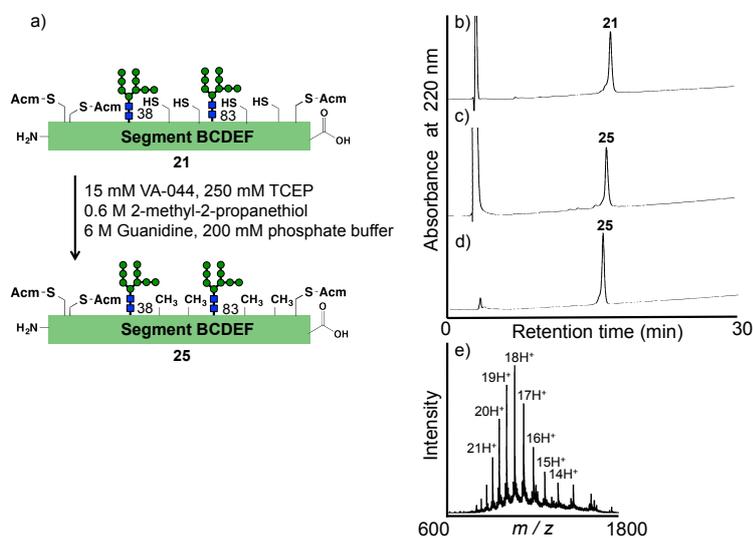


Fig. S28. Desulfurization of glycopeptide segment BCDEF for the construction of EPO bearing three M9- glycans. a) Desulfurization reaction of compound 21. b) HPLC profile for stating point ($t < 1$ min). c) HPLC profile after 4 h. d) HPLC profile of purified compound 25. e) ESI-MS spectrum of compound 25. ESI-MS: m/z calcd. for C₈₂₇H₁₃₃₈N₂₀₁O₃₁₁S₄: [M+H]⁺ 19200.9, found 19200.9 (deconvoluted).

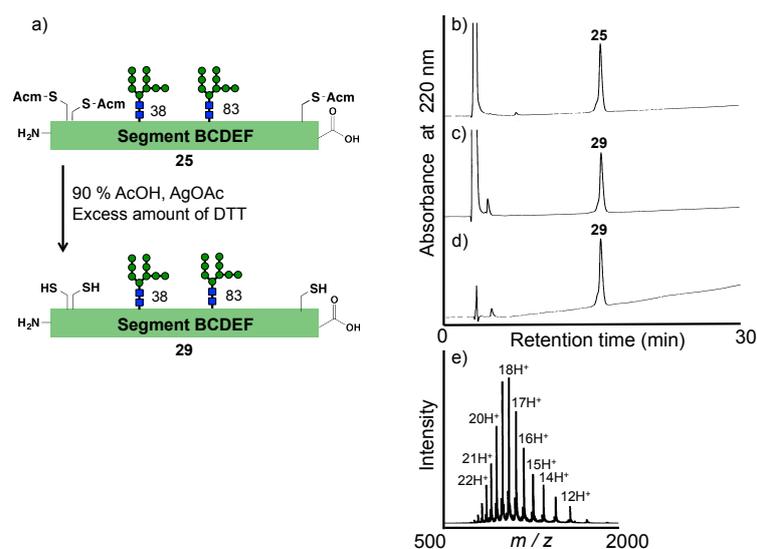


Fig. S29. Deprotection of Ac groups of glycopeptide segment BCDEF for the construction of EPO bearing three M9-glycans. a) Deprotection of Ac groups of compound **25**. b) HPLC profile for starting point (t < 1 min). c) HPLC profile after 4 h. d) HPLC profile of purified compound **29**. e) ESI-MS spectrum after purification. ESI-MS: m/z calcd. for C₈₁₈H₁₃₂₃N₁₉₈O₃₀₈S₄: [M+H]⁺ 18987.7 found 18987.4 (deconvoluted).

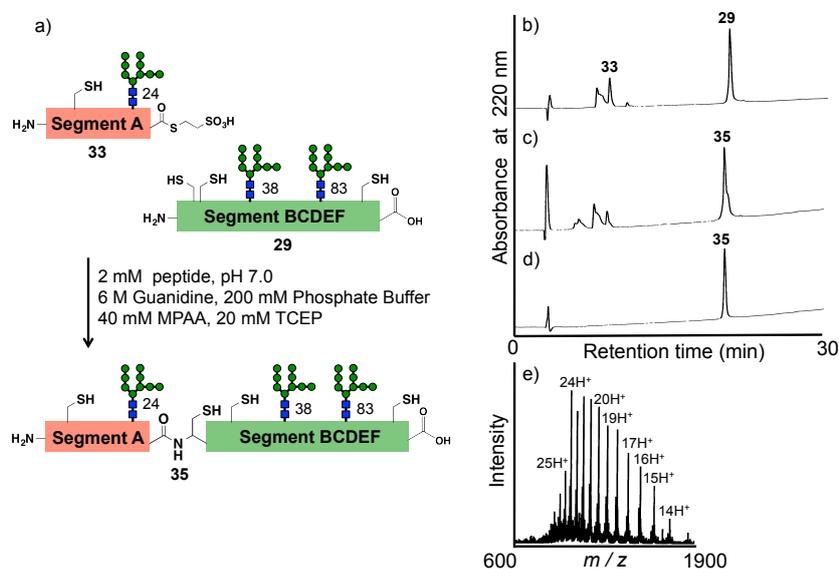


Fig. S30. Final NCL for the construction of EPO bearing three M9-glycans. a) Peptide ligation reaction between compound **29** and **33**. b) HPLC profile for starting point (t < 1 min). c) HPLC profile after 10 h. d) HPLC profile of purified compound **35**. e) ESI-MS spectrum of compound **35**. ESI-MS: m/z calcd. for C₁₀₂₅H₁₆₆₆N₂₃₉O₄₀₆S₅: [M+H]⁺ 23993.9, found 23995.7 (deconvoluted).

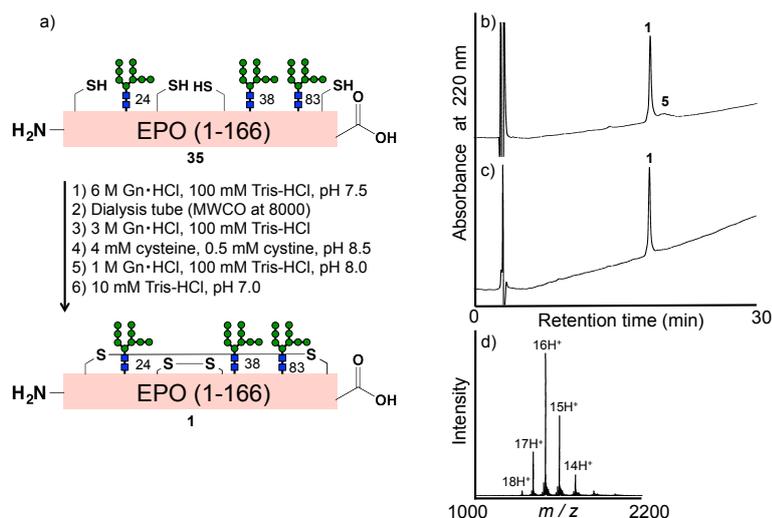


Fig. S31. Folding of segment ABCDEF for the construction of EPO bearing three M9-glycans. a) Folding experiment of compound **35**. b) HPLC profile after folding process. c) HPLC profile of purified native folded EPO (compound **1**). d) ESI-MS spectrum of compound **1**. ESI-MS: m/z calcd. for $C_{1025}H_{1662}N_{239}O_{406}S_5$: $[M+H]^+$ 23989.5730, found 23989.5022 (deconvoluted). This high-resolution mass spectrometry analysis of purified EPOs was performed with direct injections into the mass spectrometer without LC. The mass spectra were not obtained from a single time point of the LC-MS run.

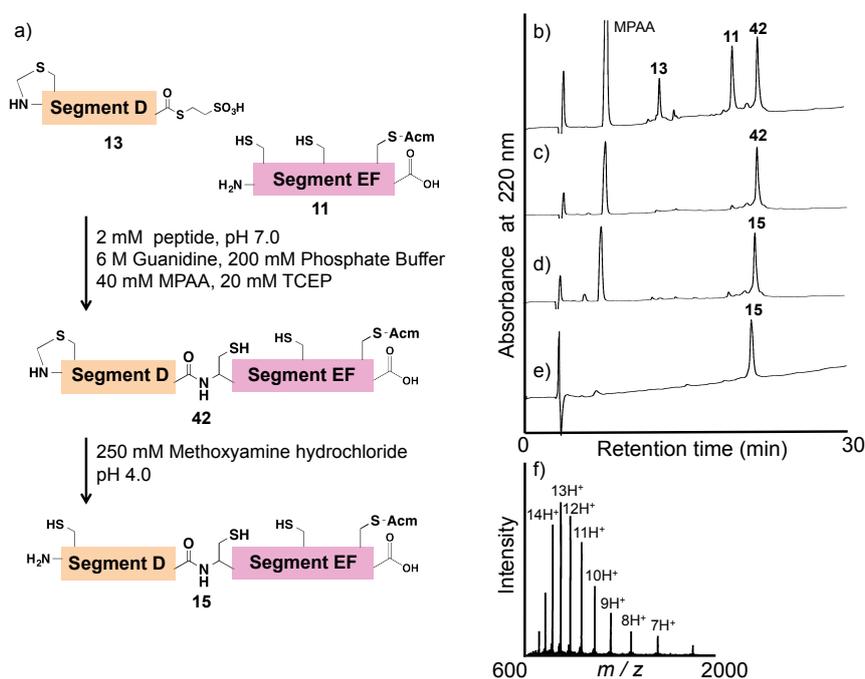


Fig. S32. 2nd-NCL for the construction of EPO bearing an M9-glycans at 24. a) Peptide ligation reaction between compound **11** and **13**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 2.5 h. d) HPLC profile after treatment with methoxyamine-HCl. e) HPLC profile of purified compound **15**. f) ESI-MS spectrum of compound **15**. ESI-MS: m/z calcd. for $C_{485}H_{802}N_{141}O_{139}S_4$: $[M+H]^+$ 10960.7, found 10960.4 (deconvoluted).

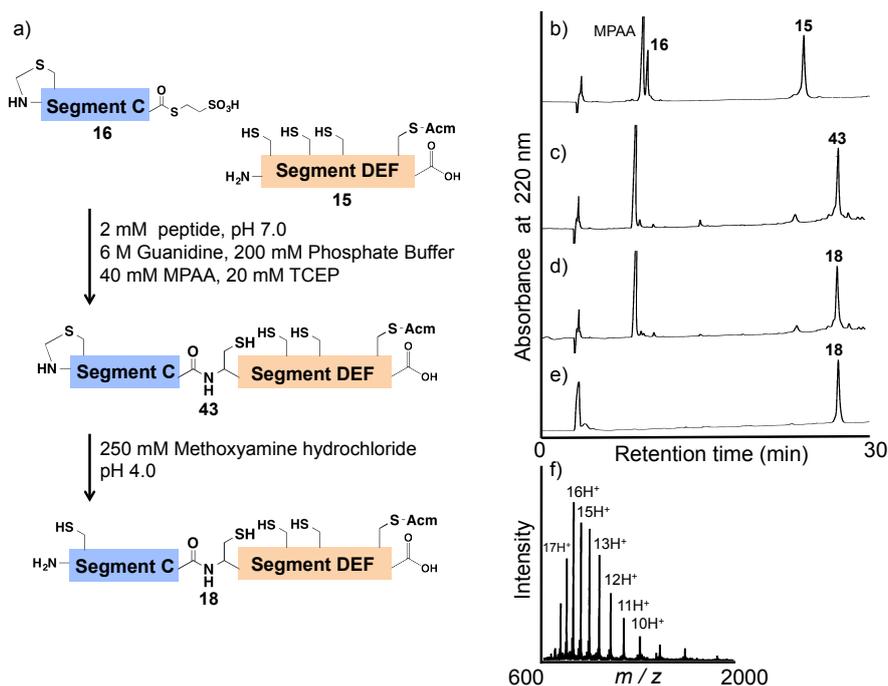


Fig. S33. 3rd-NCL for the construction of EPO bearing an M9-glycans at 24. a) Peptide ligation reaction between compound **15** and **16**. b) HPLC profile for starting point (t < 1 min). c) HPLC profile after 3 h. d) HPLC profile after treatment with methoxyamine-HCl. e) HPLC profile of compound **18**. f) ESI-MS spectrum of compound **18**. ESI-MS: *m/z* calcd. for C₅₈₀H₉₄₇N₁₆₈O₁₆₄S₆ : [M+H]⁺ 13089.1, found 13088.6 (deconvoluted).

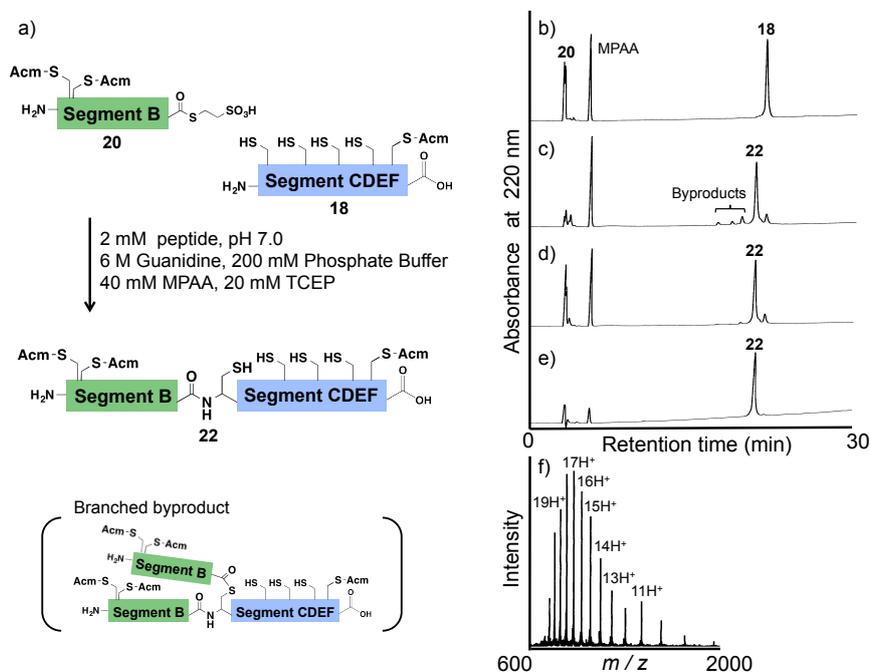


Fig. S34. 4th-NCL for the construction of EPO bearing an M9-glycans at 24. a) Peptide ligation reaction between compound **18** and **20**. b) HPLC profile for starting point (t < 1 min). c) HPLC profile after 3.5 h. d) HPLC profile after treatment with MESNa to remove the branched byproducts. e) HPLC profile of compound **22**. f) ESI-MS spectrum of compound **22**. ESI-MS: *m/z* calcd. for C₅₈₀H₉₄₇N₁₆₈O₁₆₄S₆ : [M+H]⁺ 13089.1, found 13088.6 (deconvoluted).

ESI-MS spectrum of compound **22**. ESI-MS: m/z calcd. for $C_{691}H_{1118}N_{197}O_{199}S_8$: $[M+H]^+$ 15626.0, found 15626.0 (deconvoluted).

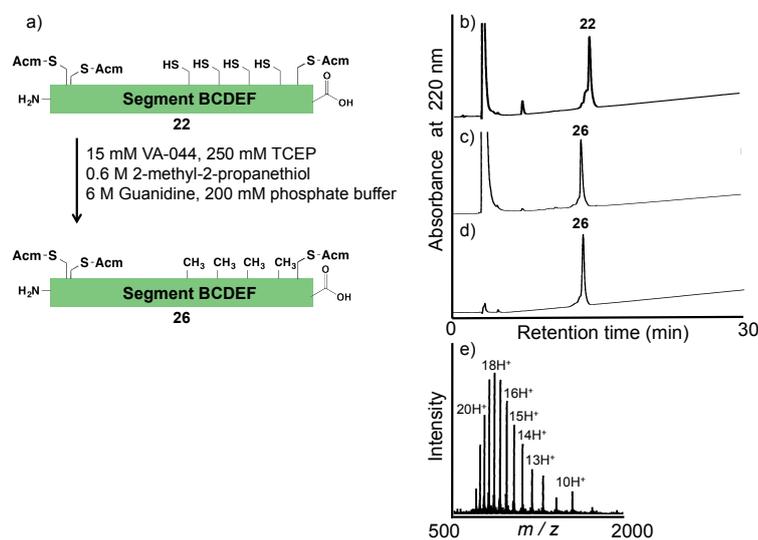


Fig. S35. Desulfurization of glycopeptide segment BCDEF for the construction of EPO bearing an M9-glycans at 24. a) Desulfurization reaction of compound **22**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 5 h. d) HPLC profile of purified compound **26**. e) ESI-MS spectrum of compound **26**. ESI-MS: m/z calcd. for $C_{691}H_{1118}N_{197}O_{199}S_4$: $[M+H]^+$ 15497.8, found 15498.6 (deconvoluted).

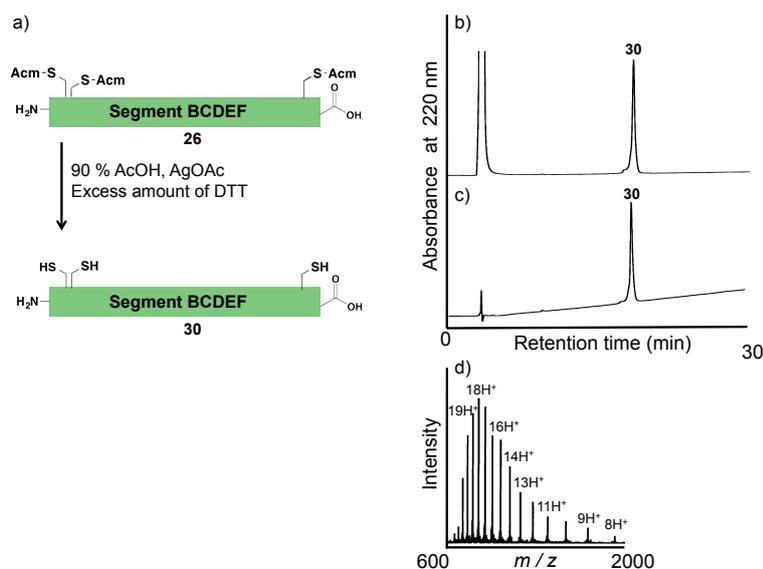


Fig. S36. Deprotection of Acm groups of glycopeptide segment BCDEF for the construction of EPO bearing an M9-glycans at 24. a) Deprotection of Acm groups of compound **26**. b) HPLC profile after 4 h. c) HPLC profile of purified compound **30**. d) ESI-MS spectrum of compound **30**. ESI-MS: m/z calcd. for $C_{682}H_{1103}N_{194}O_{196}S_4$: $[M+H]^+$ 15284.5, found 15284.6 (deconvoluted).

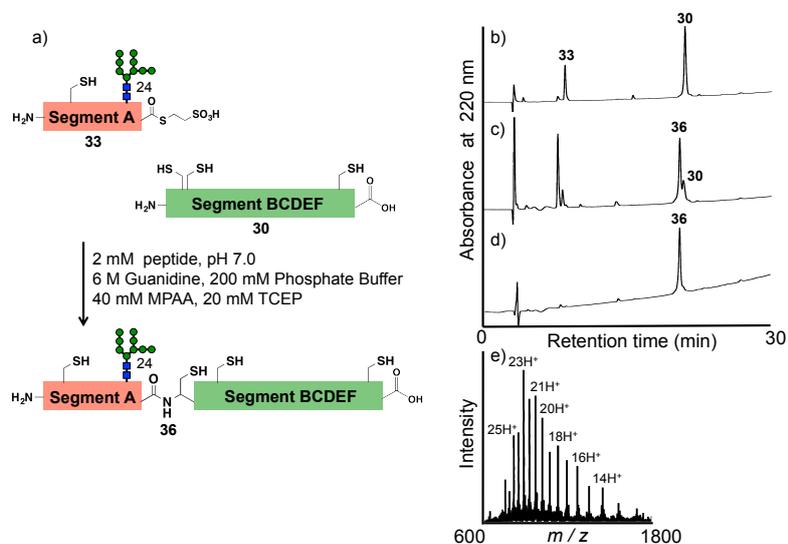


Fig. S37. Final NCL for the construction of EPO bearing an M9-glycans at 24. a) Peptide ligation reaction between compound **30** and **33**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 10 h. d) HPLC profile of purified compound **36**. e) ESI-MS spectrum of compound **36**. ESI-MS: m/z calcd. for $C_{889}H_{1446}N_{235}O_{294}S_5$: $[M+H]^+$ 20290.8, found 20290.8 (deconvoluted).

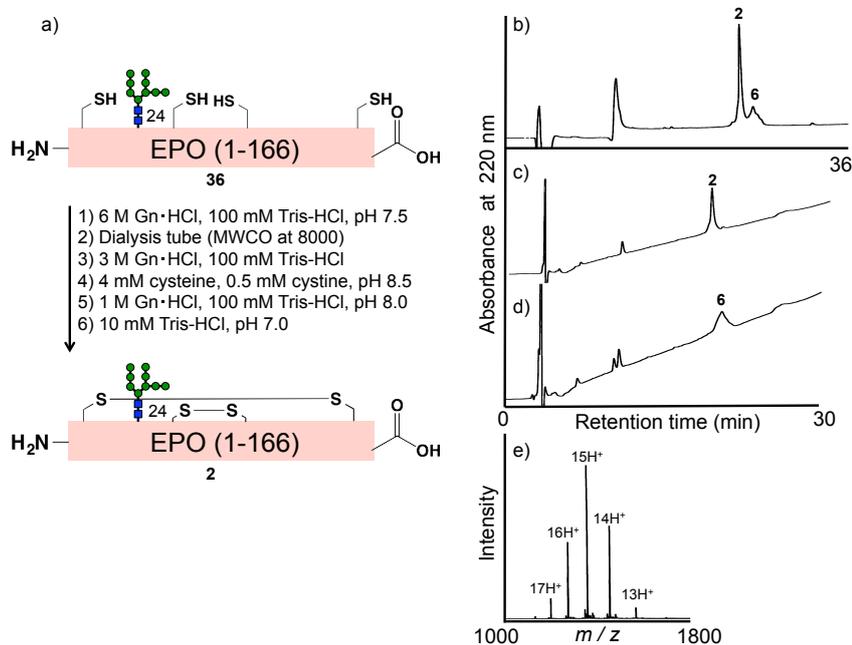


Fig. S38. Folding of segment ABCDEF for the construction of EPO bearing an M9-glycans at 24. a) Folding experiment of compound **36**. b) HPLC profile after folding process. c) HPLC profile of purified native folded EPO (compound **2**). d) HPLC profile of purified misfolded EPO (compound **6**). e) ESI-MS spectrum of compound **2**. FT-ICR-MS: m/z calcd. for $C_{889}H_{1442}N_{235}O_{294}S_5$: $[M+H]^+$ 20285.3948, found 20285.4012 (deconvoluted). This high-resolution mass spectrometry analysis of purified EPOs was performed with direct injections into

the mass spectrometer without LC. The mass spectra were not obtained from a single time point of the LC-MS run.

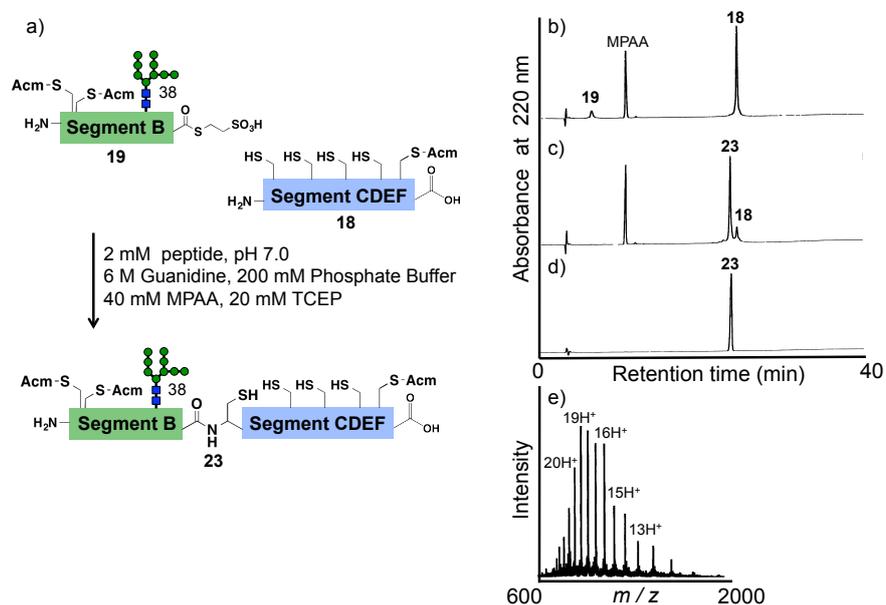


Fig. S39. 4th-NCL for the construction of EPO bearing an M9-glycans at 38. a) Peptide ligation reaction between compound **18** and **19**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 7 h. d) HPLC profile after treatment with methoxyamine-HCl. e) HPLC profile of purified compound **23**. f) ESI-MS spectrum of compound **23**. ESI-MS: m/z calcd. for $C_{759}H_{1228}N_{199}O_{255}S_8$: $[M+H]^+$ 17477.6, found 17476.0 (deconvoluted).

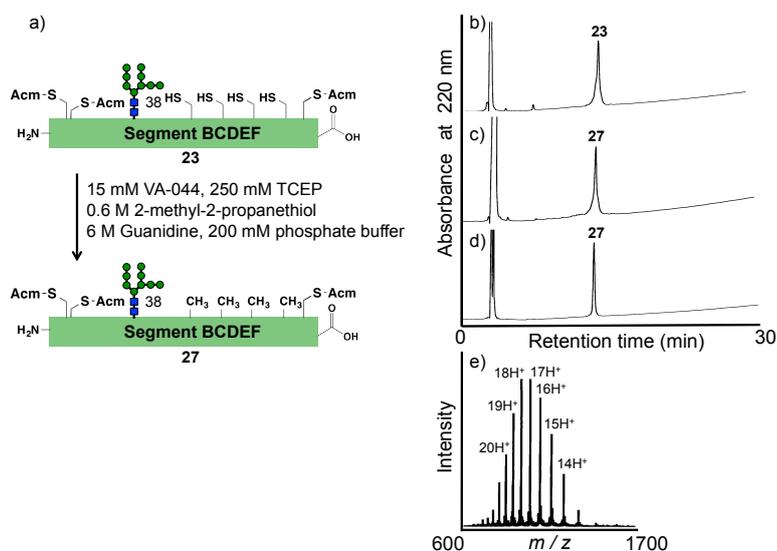


Fig. S40. Desulfurization of glycopeptide segment BCDEF for the construction of EPO bearing an M9-glycans at 38. a) Desulfurization reaction of compound **23**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 3.5 h. d) HPLC profile of purified compound **27**. e) ESI-MS spectrum of compound **27**. ESI-MS: m/z calcd. for $C_{759}H_{1228}N_{199}O_{255}S_4$: $[M+H]^+$ 17349.4, found 17348.3 (deconvoluted).

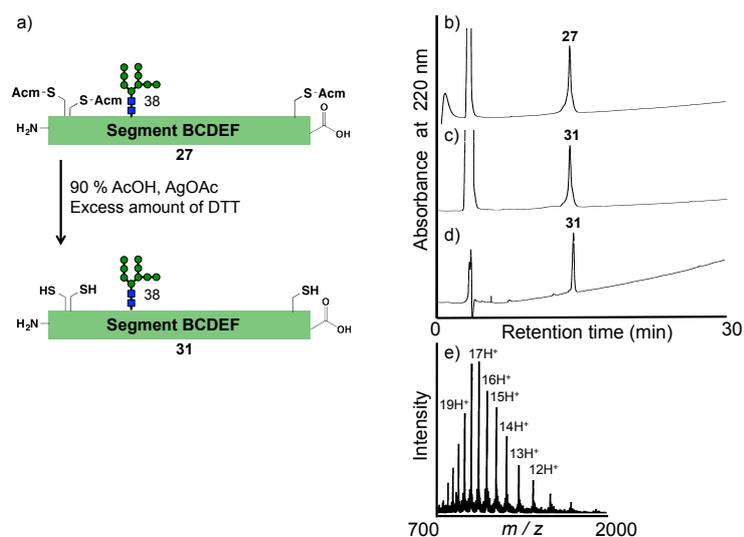


Fig. S41. Deprotection of Ac-m groups of glycopeptide segment BCDEF for the construction of EPO bearing an M9-glycans at 38. a) Deprotection of Ac-m groups of compound 27. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 3.5 h. d) HPLC profile of purified compound 31. e) ESI-MS spectrum of compound 31. ESI-MS: m/z calcd. for $C_{750}H_{1213}N_{196}O_{252}S_4$: $[M+H]^+$ 17136.1, found 17134.7 (deconvoluted).

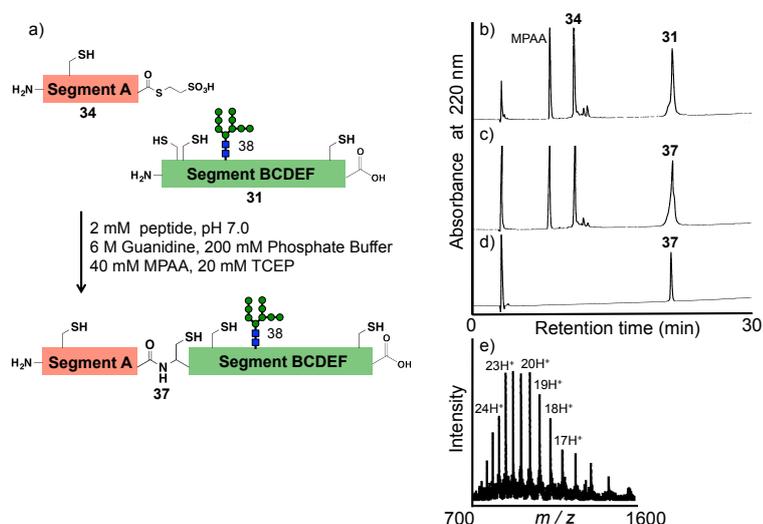


Fig. S42. Final NCL for the construction of EPO bearing an M9-glycans at 38. a) Peptide ligation reaction between compound 31 and 34. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 5.5 h. d) HPLC profile of purified compound 37. e) ESI-MS spectrum of compound 37. ESI-MS: m/z calcd. for $C_{889}H_{1446}N_{235}O_{294}S_5$: $[M+H]^+$ 20290.8, found 20289.2 (deconvoluted).

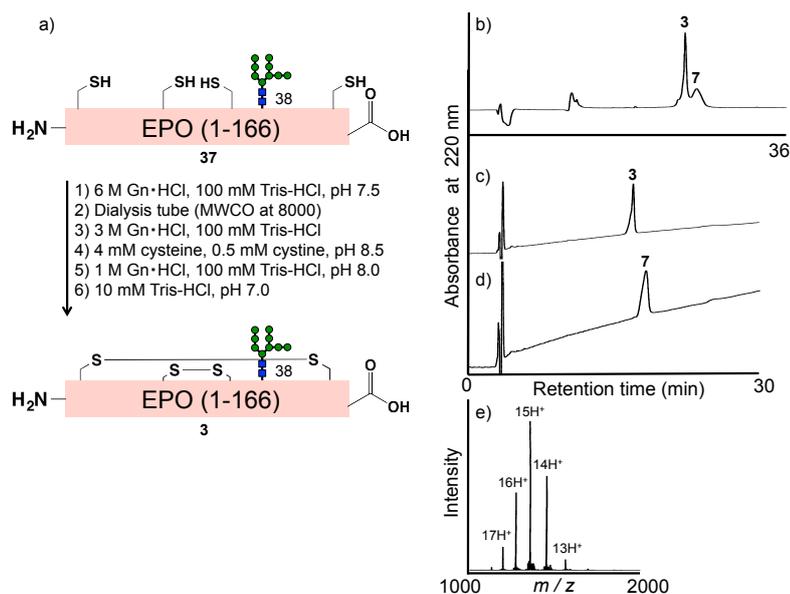


Fig. S43. Folding of segment ABCDEF for the construction of EPO bearing an M9-glycans at 38. a) Folding experiment of compound **37**. b) HPLC profile after folding process. c) HPLC profile of purified native folded EPO (compound **3**). d) HPLC profile of purified misfolded EPO (compound **7**). e) ESI-MS spectrum of compound **3**. FT-ICR-MS: m/z calcd. for $C_{889}H_{1442}N_{235}O_{294}S_5$: $[M+H]^+$ 20285.3948, found 20285.3562 (deconvoluted). This high-resolution mass analysis of purified EPOs was performed with direct injections into the mass spectrometer without LC. The mass spectra were not obtained from a single time point of the LC-MS run.

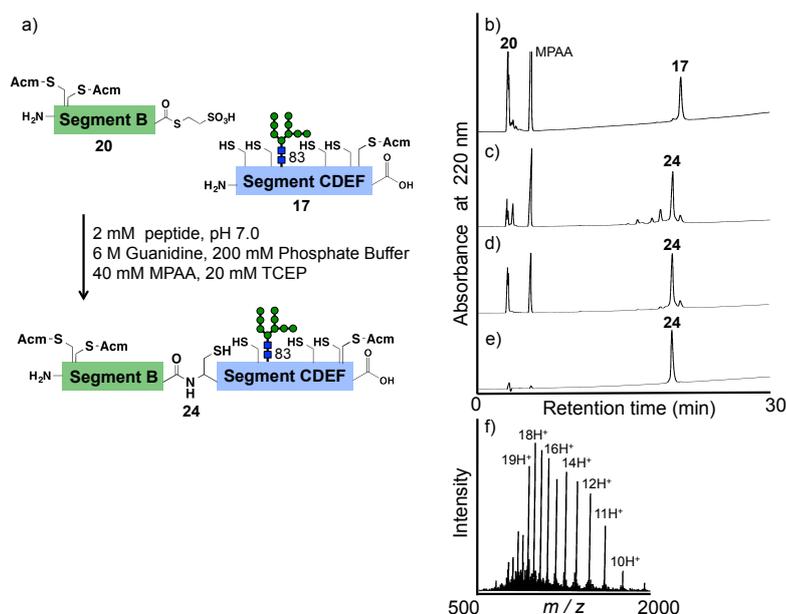


Fig. S44. 4th-NCL for the construction of EPO bearing an M9-glycans at 83. a) Peptide ligation reaction between compound **17** and **20**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 4.5 h. d) HPLC profile after treatment with MESNa. e) HPLC profile of purified compound **24**. f) ESI-MS spectrum of compound **24**. ESI-MS: m/z calcd. for $C_{759}H_{1228}N_{199}O_{255}S_8$: $[M+H]^+$ 17477.6, found 17478.9 (deconvoluted).

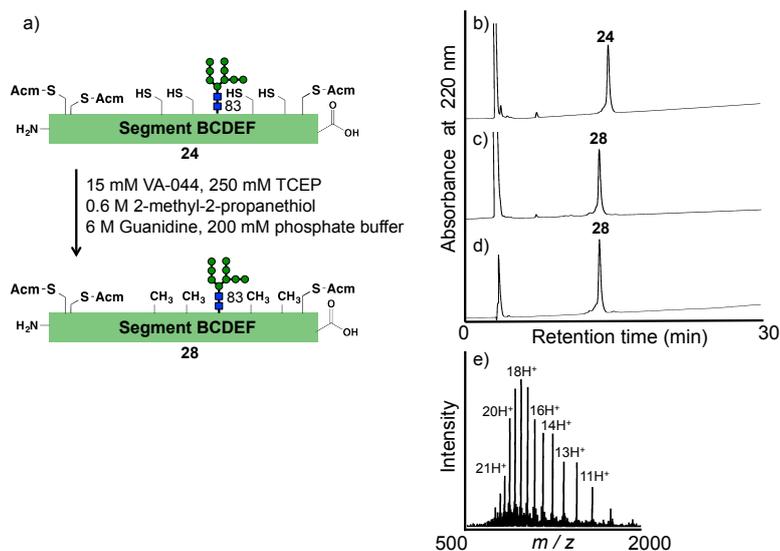


Fig. S45. Desulfurization of glycopeptide segment BCDEF for the construction of EPO bearing an M9-glycans at 83. a) Desulfurization reaction of compound **24**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 3 h. d) HPLC profile of purified compound **28**. e) ESI-MS spectrum of compound **28**. ESI-MS: m/z calcd. for $C_{759}H_{1228}N_{199}O_{255}S_4$: $[M+H]^+$ 17349.4, found 17349.7 (deconvoluted).

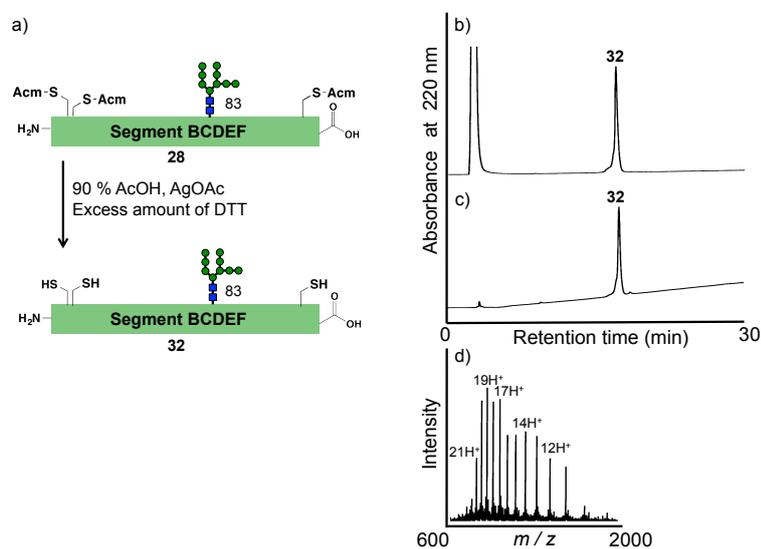


Fig. S46. Deprotection of Acm groups of glycopeptide segment BCDEF for the construction of EPO bearing an M9-glycans at 83. a) Deprotection of Acm group of compound **28**. b) HPLC profile after 4 h. c) HPLC profile of purified compound **32**. d) ESI-MS spectrum of compound **32**. ESI-MS: m/z calcd. for $C_{750}H_{1213}N_{196}O_{252}S_4$: $[M+H]^+$ 17136.1, found 17136.5 (deconvoluted).

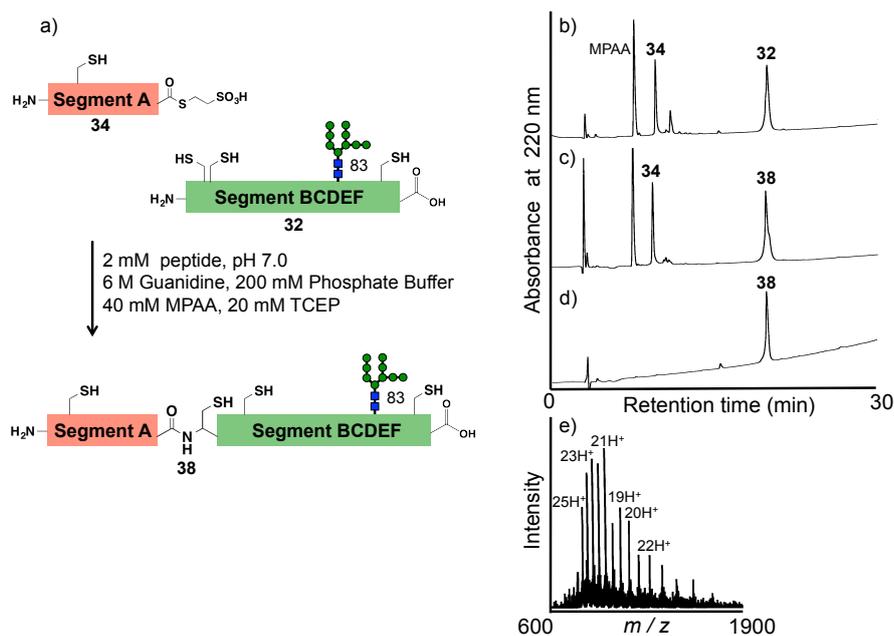


Fig. S47. Final NCL for the construction of EPO bearing an M9-glycans at 83. a) Peptide ligation reaction between compound 32 and 34. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 10 h. d) HPLC profile of purified compound 38. e) ESI-MS spectrum of compound 38. ESI-MS: m/z calcd. for $\text{C}_{889}\text{H}_{1446}\text{N}_{235}\text{O}_{294}\text{S}_5$: $[\text{M}+\text{H}]^+$ 20290.8, found 20290.9 (deconvoluted).

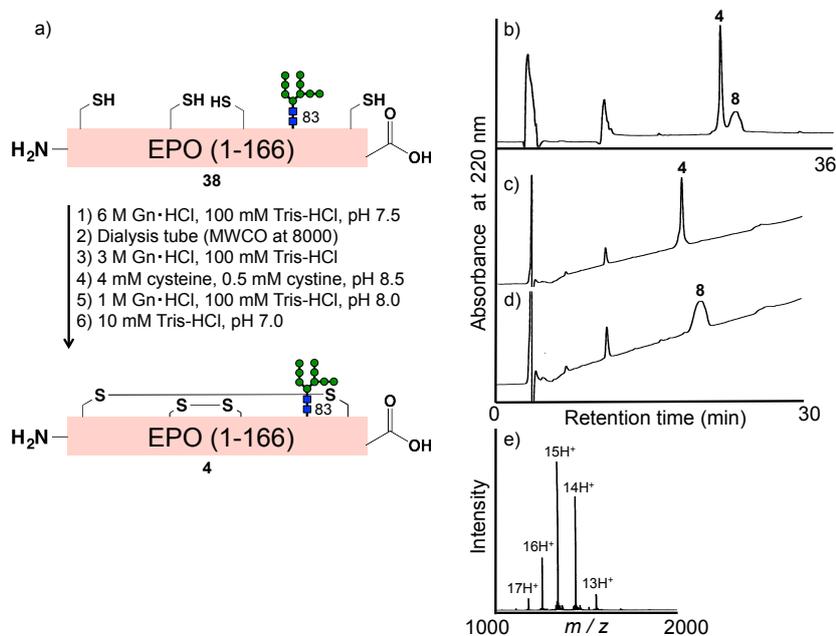


Fig. S48. Folding of segment ABCDEF for the construction of EPO bearing an M9-glycans at 83. a) Folding experiment of compound 38. b) HPLC profile after folding process. c) HPLC profile of purified native folded EPO (compound 4). d) HPLC profile of purified misfolded EPO (compound 8). e) ESI-MS spectrum of compound 4.

FT-ICR-MS: m/z calcd. for $C_{889}H_{1442}N_{235}O_{294}S_5$: $[M+H]^+$ 20285.3948, found 20285.3514 (deconvoluted). This high-resolution mass spectrometry analysis of purified EPOs was performed with direct injections into the mass spectrometer without LC. The mass spectra were not obtained from a single time point of the LC-MS run.

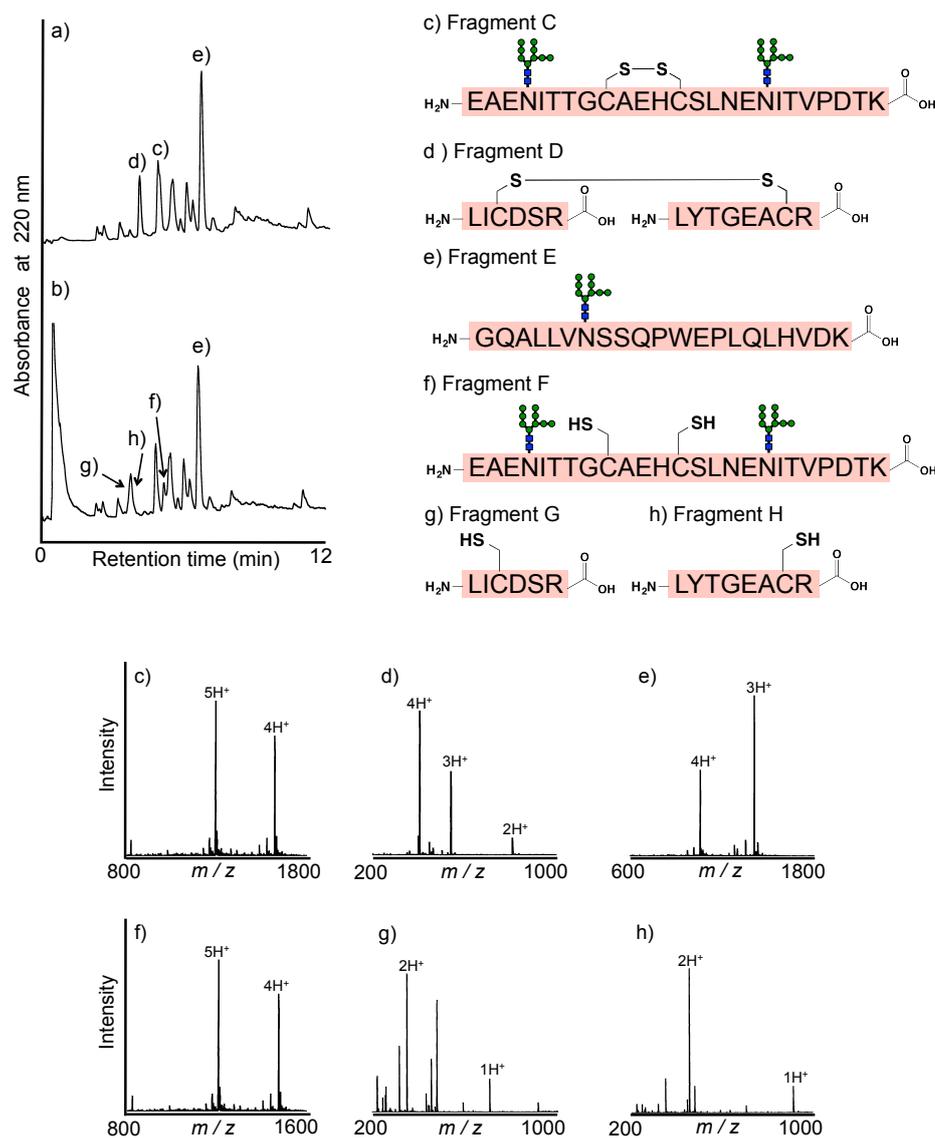


Fig. S49. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from native folded EPO bearing three glycans (compound 1). a) HPLC profile after trypsin digestion. b) HPLC profile after TCEP treatment. c) ESI-MS spectrum of Fragment C: ESI-MS: m/z calcd. for $C_{249}H_{408}N_{35}O_{154}S_2$: $[M+H]^+$ 6420.2, found 6419.8 (deconvoluted). d) ESI-MS spectrum of Fragment D: ESI-MS: m/z calcd. for $C_{66}H_{111}N_{20}O_{23}S_2$: $[M+H]^+$ 1616.8, found 1616.3 (deconvoluted). e) ESI-MS spectrum of Fragment E: ESI-MS: m/z calcd. for $C_{176}H_{284}N_{31}O_{87}$: $[M+H]^+$ 4223.9, found 4225.3 (deconvoluted). f) ESI-MS spectrum of Fragment F: ESI-MS: m/z calcd. for $C_{249}H_{410}N_{35}O_{154}S_2$: $[M+H]^+$ 6422.2, found 6421.8 (deconvoluted). g) ESI-MS spectrum of

Fragment G: ESI-MS: m/z calcd. for $C_{28}H_{52}N_9O_{10}S$: $[M+H]^+$ 706.8, found 706.7 (deconvoluted). h) ESI-MS spectrum of Fragment H: ESI-MS: m/z calcd. for $C_{38}H_{62}N_{11}O_{13}S$: $[M+H]^+$ 912.4, found 912.7 (deconvoluted).

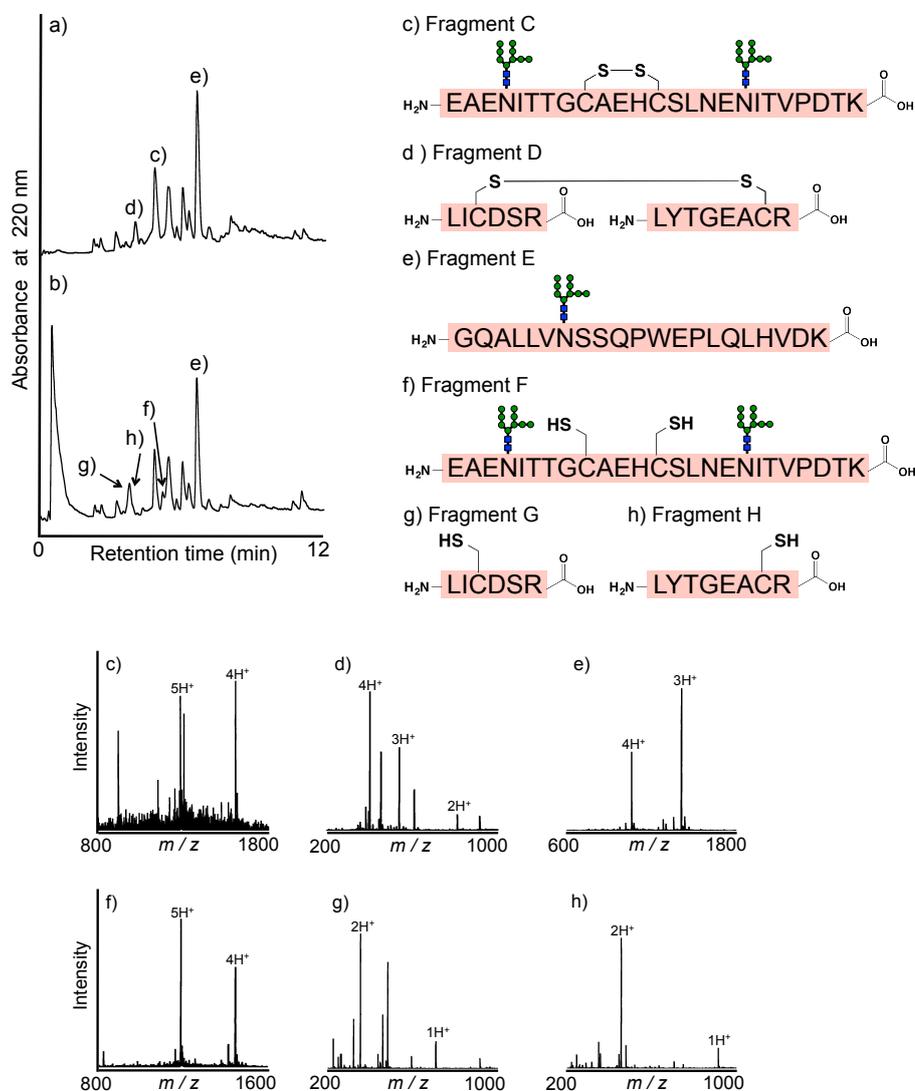


Fig. S50. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from misfolded EPO bearing three glycans (compound 5). a) HPLC profile after trypsin digestion. b) HPLC profile after TCEP treatment. c) ESI-MS spectrum of Fragment C: ESI-MS: m/z calcd. for $C_{249}H_{408}N_{35}O_{154}S_2$: $[M+H]^+$ 6420.2, found 6419.8 (deconvoluted). d) ESI-MS spectrum of Fragment D: ESI-MS: m/z calcd. for $C_{66}H_{111}N_{20}O_{23}S_2$: $[M+H]^+$ 1616.8, found 1616.3 (deconvoluted). e) ESI-MS spectrum of Fragment E: ESI-MS: m/z calcd. for $C_{176}H_{284}N_{31}O_{87}$: $[M+H]^+$ 4223.9, found 4225.7 (deconvoluted). f) ESI-MS spectrum of Fragment F: ESI-MS: m/z calcd. for $C_{249}H_{410}N_{35}O_{154}S_2$: $[M+H]^+$ 6422.2, found 6421.7 (deconvoluted). g) ESI-MS spectrum of Fragment G: ESI-MS: m/z calcd. for $C_{28}H_{52}N_9O_{10}S$: $[M+H]^+$ 706.8, found 706.6 (deconvoluted). h) ESI-MS spectrum of Fragment H: ESI-MS: m/z calcd. for $C_{38}H_{62}N_{11}O_{13}S$: $[M+H]^+$ 912.4, found 912.7 (deconvoluted).

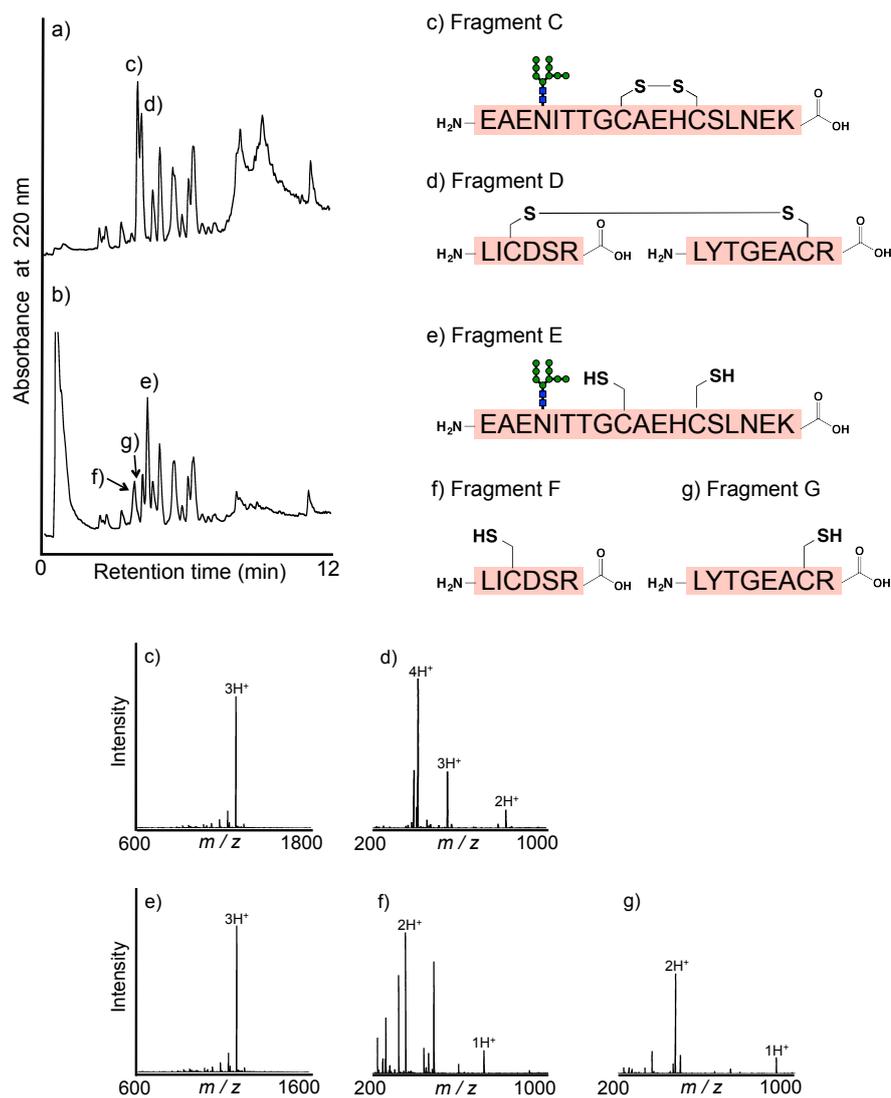


Fig. S51. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from native folded EPO bearing a glycan at 24 (compound 2). a) HPLC profile after trypsin digestion. b) HPLC profile after TCEP treatment. c) ESI-MS spectrum of Fragment C: ESI-MS: m/z calcd. for $C_{147}H_{240}N_{25}O_{87}S_2$: $[M+H]^+$ 3813.7, found 3815.6 (deconvoluted). d) ESI-MS spectrum of Fragment D: ESI-MS: m/z calcd. for $C_{66}H_{111}N_{20}O_{23}S_2$: $[M+H]^+$ 1616.8, found 1616.3 (deconvoluted). e) ESI-MS spectrum of Fragment E: ESI-MS: m/z calcd. for $C_{147}H_{242}N_{25}O_{87}S_2$: $[M+H]^+$ 3815.7, found 3817.6 (deconvoluted). f) ESI-MS spectrum of Fragment F: ESI-MS: m/z calcd. for $C_{28}H_{52}N_9O_{10}S$: $[M+H]^+$ 706.8, found 706.7 (deconvoluted). g) ESI-MS spectrum of Fragment G: ESI-MS: m/z calcd. for $C_{38}H_{62}N_{11}O_{13}S$: $[M+H]^+$ 912.4, found 912.7 (deconvoluted).

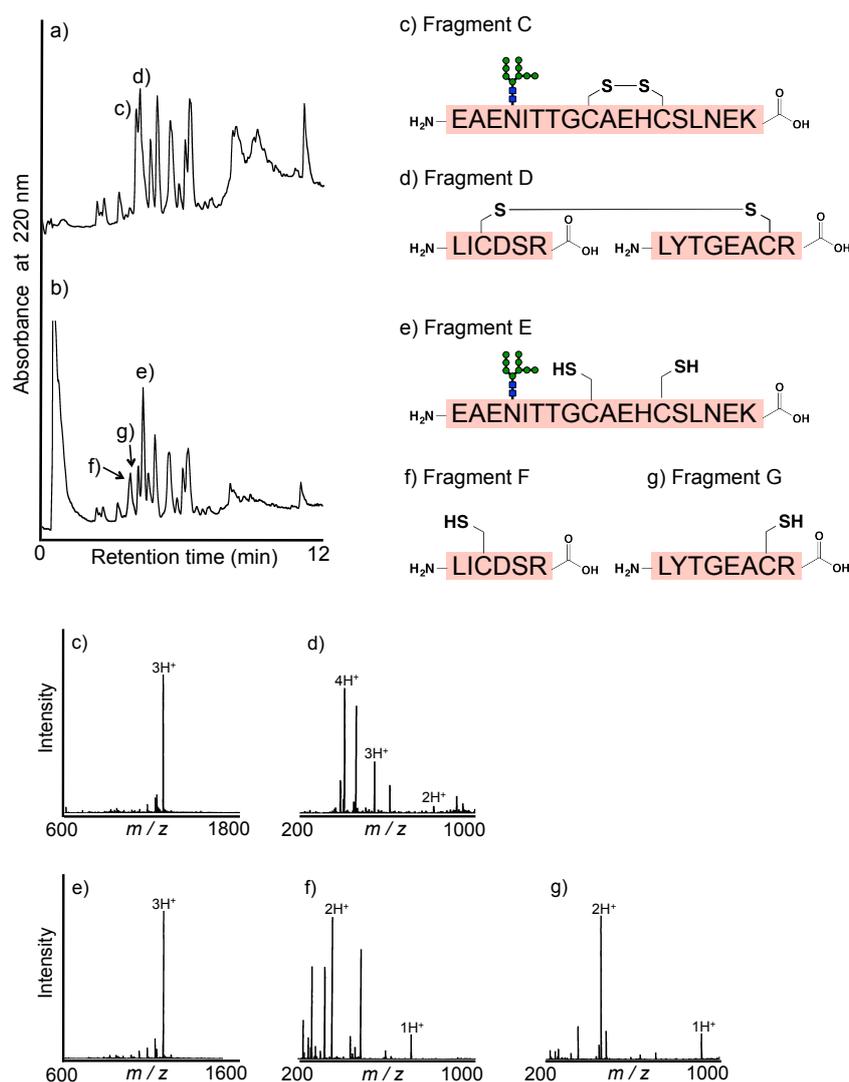


Fig. S52. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from misfolded EPO bearing a glycan at 24 (compound 6). a) HPLC profile after trypsin digestion. b) HPLC profile after TCEP treatment. c) ESI-MS spectrum of Fragment C: ESI-MS: m/z calcd. for $C_{147}H_{240}N_{25}O_{87}S_2$: $[M+H]^+$ 3813.7, found 3815.6 (deconvoluted). d) ESI-MS spectrum of Fragment D: ESI-MS: m/z calcd. for $C_{66}H_{111}N_{20}O_{23}S_2$: $[M+H]^+$ 1616.8, found 1616.4 (deconvoluted). e) ESI-MS spectrum of Fragment E: ESI-MS: m/z calcd. for $C_{147}H_{242}N_{25}O_{87}S_2$: $[M+H]^+$ 3815.7, found 3817.7 (deconvoluted). f) ESI-MS spectrum of Fragment F: ESI-MS: m/z calcd. for $C_{28}H_{52}N_9O_{10}S$: $[M+H]^+$ 706.8, found 706.6 (deconvoluted). g) ESI-MS spectrum of Fragment G: ESI-MS: m/z calcd. for $C_{38}H_{62}N_{11}O_{13}S$: $[M+H]^+$ 912.4, found 912.7 (deconvoluted).

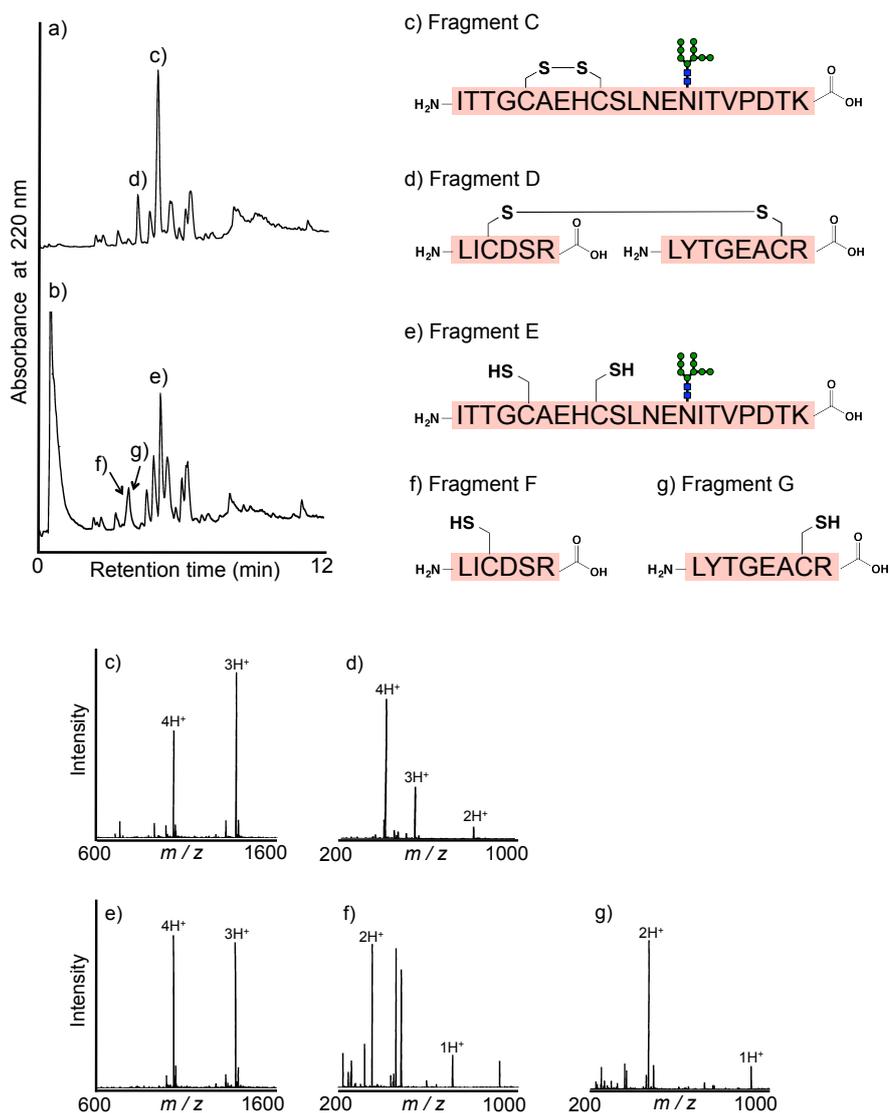


Fig. S53. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from native folded EPO bearing a glycan at 38 (compound 3). a) HPLC profile after trypsin digestion. b) HPLC profile after TCEP treatment. c) ESI-MS spectrum of Fragment C: ESI-MS: m/z calcd. for $C_{162}H_{267}N_{28}O_{90}S_2$: $[M+H]^+$ 4108.7, found 4809.6 (deconvoluted). d) ESI-MS spectrum of Fragment D: ESI-MS: m/z calcd. for $C_{66}H_{111}N_{20}O_{23}S_2$: $[M+H]^+$ 1616.8, found 1616.3 (deconvoluted). e) ESI-MS spectrum of Fragment E: ESI-MS: m/z calcd. for $C_{162}H_{269}N_{28}O_{90}S_2$: $[M+H]^+$ 4110.7, found 4111.6 (deconvoluted). f) ESI-MS spectrum of Fragment F: ESI-MS: m/z calcd. for $C_{28}H_{51}N_9O_{10}S$: $[M+H]^+$ 706.8, found 706.7 (deconvoluted). g) ESI-MS spectrum of Fragment G: ESI-MS: m/z calcd. for $C_{38}H_{61}N_{11}O_{13}S$: $[M+H]^+$ 912.4, found 912.7 (deconvoluted).

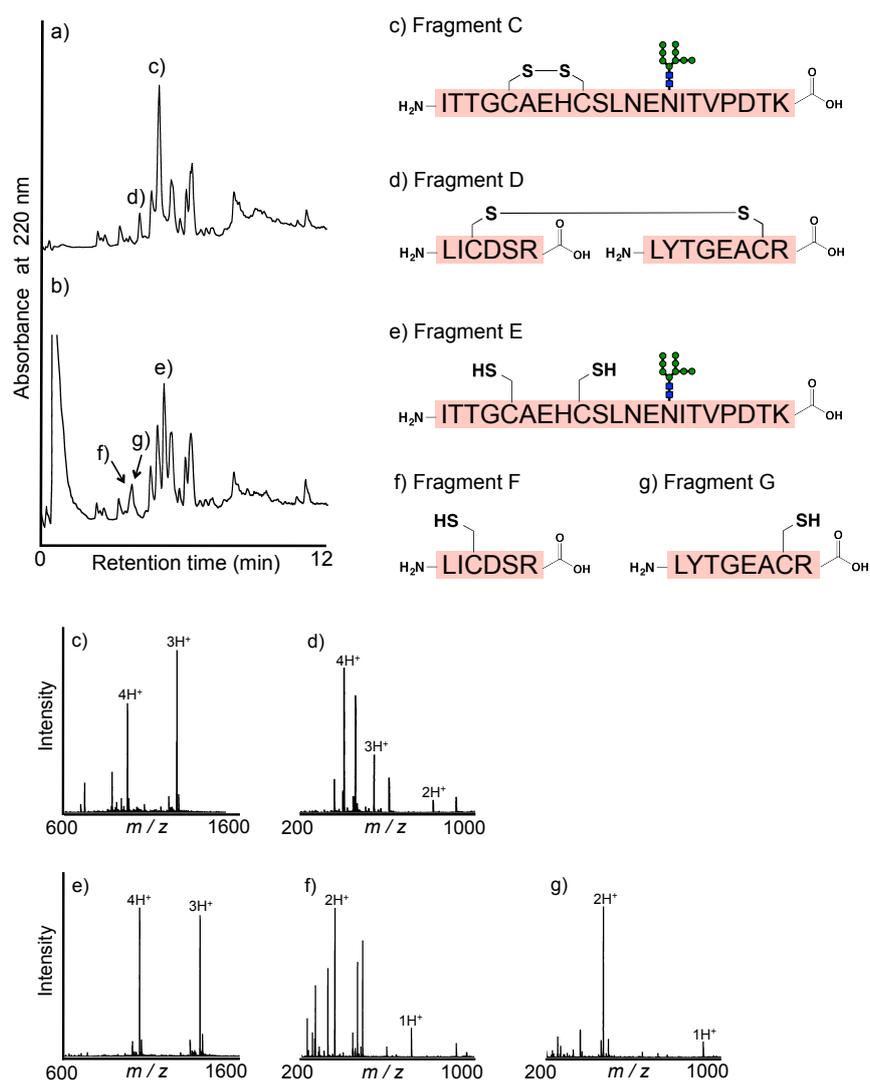


Fig. S54. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from misfolded EPO bearing a glycan at 38 (compound 7). a) HPLC profile after trypsin digestion. b) HPLC profile after TCEP treatment. c) ESI-MS spectrum of Fragment C: ESI-MS: m/z calcd. for $C_{162}H_{267}N_{28}O_{90}S_2$: $[M+H]^+$ 4108.7, found 4809.1 (deconvoluted). d) ESI-MS spectrum of Fragment D: ESI-MS: m/z calcd. for $C_{66}H_{111}N_{20}O_{23}S_2$: $[M+H]^+$ 1616.8, found 1616.3 (deconvoluted). e) ESI-MS spectrum of Fragment E: ESI-MS: m/z calcd. for $C_{162}H_{269}N_{28}O_{90}S_2$: $[M+H]^+$ 4110.7, found 4112.1 (deconvoluted). f) ESI-MS spectrum of Fragment F: ESI-MS: m/z calcd. for $C_{28}H_{52}N_9O_{10}S$: $[M+H]^+$ 706.8, found 706.7 (deconvoluted). g) ESI-MS spectrum of Fragment G: ESI-MS: m/z calcd. for $C_{38}H_{62}N_{11}O_{13}S$: $[M+H]^+$ 912.4, found 912.8 (deconvoluted).

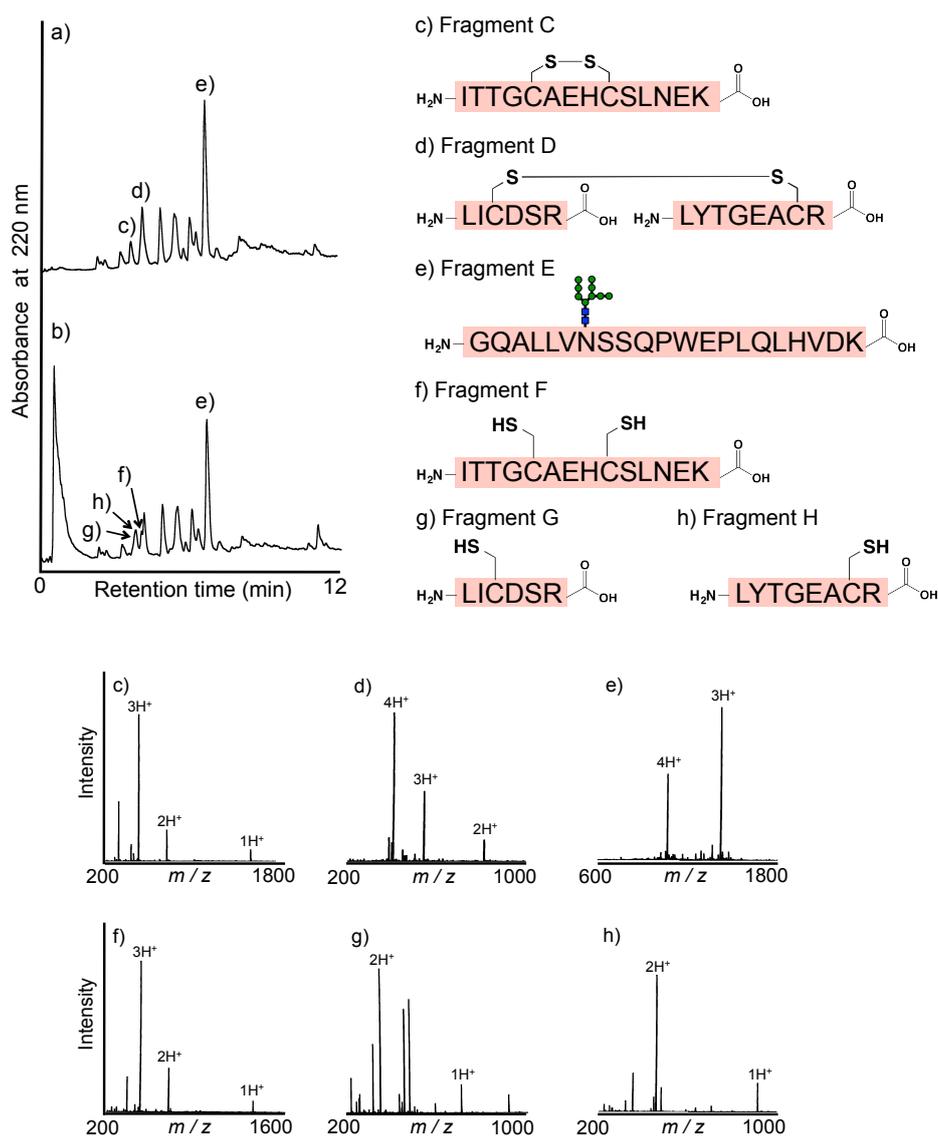


Fig. S55. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from correctly native EPO bearing a glycan at 83 (compound 4). a) HPLC profile after trypsin digestion. b) HPLC profile after TCEP treatment. c) ESI-MS spectrum of Fragment C: ESI-MS: m/z calcd. for $C_{60}H_{99}N_{18}O_{23}S_2$: $[M+H]^+$ 1504.7, found 1504.1 (deconvoluted). d) ESI-MS spectrum of Fragment D: ESI-MS: m/z calcd. for $C_{66}H_{111}N_{20}O_{23}S_2$: $[M+H]^+$ 1616.8, found 1616.3 (deconvoluted). e) ESI-MS spectrum of Fragment E: ESI-MS: m/z calcd. for $C_{176}H_{284}N_{31}O_{87}$: $[M+H]^+$ 4223.9, found 4224.8 (deconvoluted). f) ESI-MS spectrum of Fragment F: ESI-MS: m/z calcd. for $C_{60}H_{101}N_{18}O_{23}S_2$: $[M+H]^+$ 1506.7, found 1506.1 (deconvoluted). g) ESI-MS spectrum of Fragment G: ESI-MS: m/z calcd. for $C_{28}H_{52}N_9O_{10}S$: $[M+H]^+$ 706.8, found 706.6 (deconvoluted). h) ESI-MS spectrum of Fragment H: ESI-MS: m/z calcd. for $C_{38}H_{62}N_{11}O_{13}S$: $[M+H]^+$ 912.4, found 912.7 (deconvoluted).

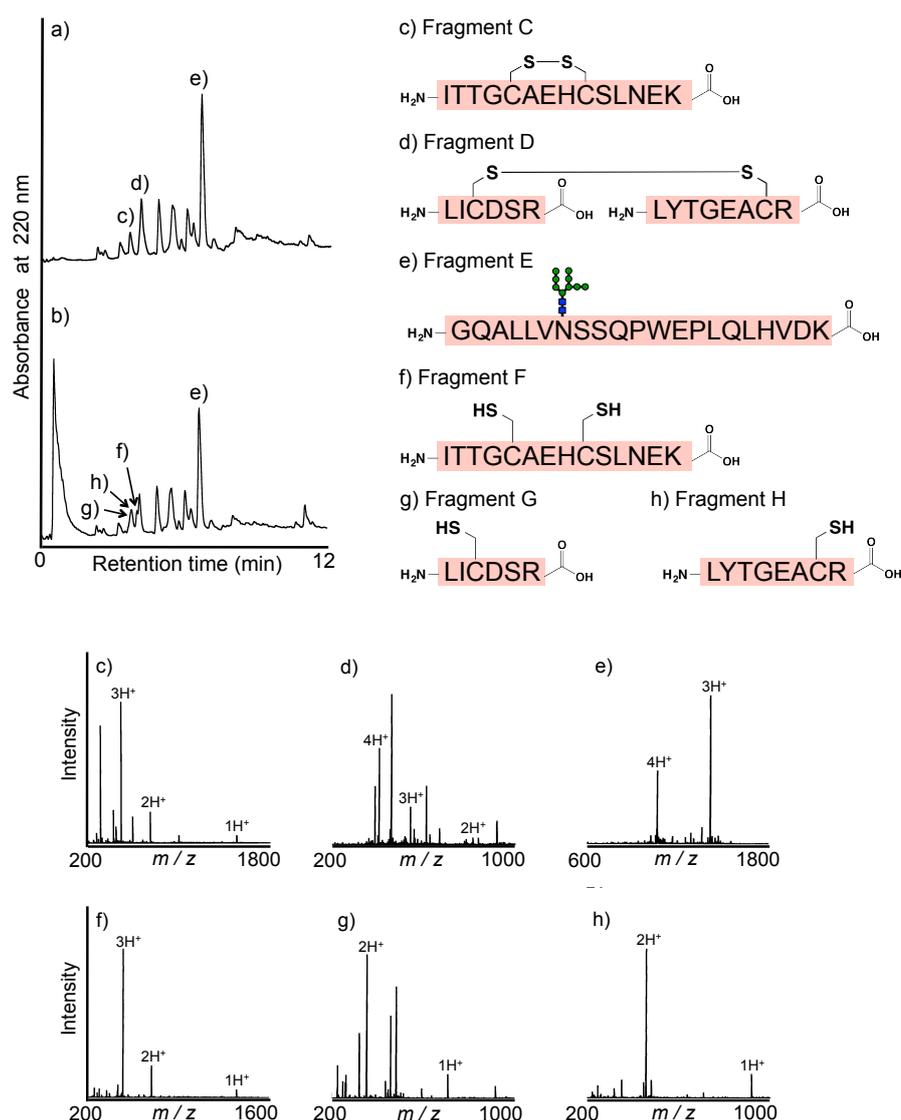


Fig. S56. HPLC profiles and ESI-MS spectra of the resultant peptide fragments by trypsin digestion from misfolded EPO bearing a glycan at 83 (compound 8). a) HPLC profile after trypsin digestion. b) HPLC profile after TCEP treatment. c) ESI-MS spectrum of Fragment C: ESI-MS: m/z calcd. for $C_{60}H_{99}N_{18}O_{23}S_2$: $[M+H]^+$ 1504.7, found 1504.1 (deconvoluted). d) ESI-MS spectrum of Fragment D: ESI-MS: m/z calcd. for $C_{66}H_{111}N_{20}O_{23}S_2$: $[M+H]^+$ 1616.8, found 1616.4 (deconvoluted). e) ESI-MS spectrum of Fragment E: ESI-MS: m/z calcd. for $C_{176}H_{284}N_{31}O_{87}$: $[M+H]^+$ 4223.9, found 4224.7 (deconvoluted). f) ESI-MS spectrum of Fragment F: ESI-MS: m/z calcd. for $C_{60}H_{101}N_{18}O_{23}S_2$: $[M+H]^+$ 1506.7, found 1506.1 (deconvoluted). g) ESI-MS spectrum of Fragment G: ESI-MS: m/z calcd. for $C_{28}H_{52}N_9O_{10}S$: $[M+H]^+$ 706.8, found 706.7 (deconvoluted). h) ESI-MS spectrum of Fragment H: ESI-MS: m/z calcd. for $C_{38}H_{62}N_{11}O_{13}S$: $[M+H]^+$ 912.4, found 912.7 (deconvoluted).

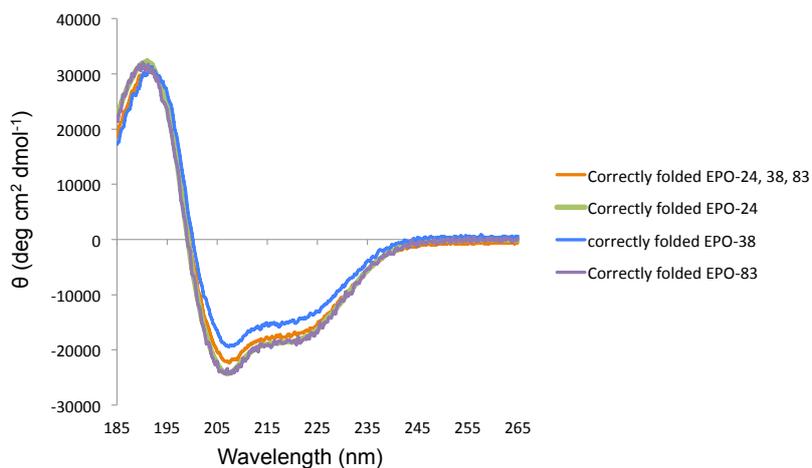


Fig. S57. Circular dichroism (CD) spectra of native folded EPOs bearing one or three glycans.

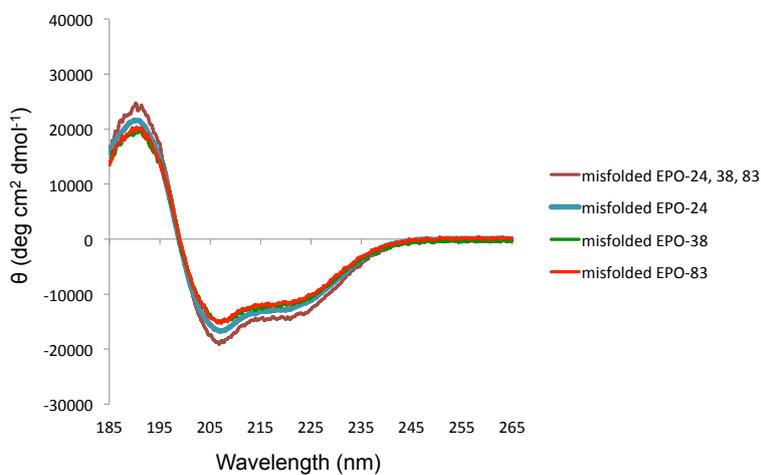


Fig. S58. Circular dichroism (CD) spectra of misfolded EPOs bearing one or three glycans

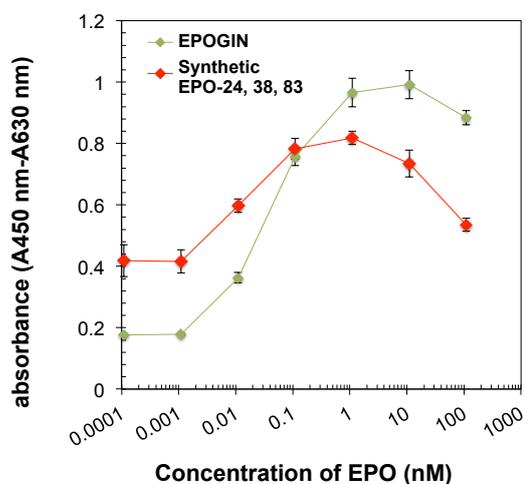


Fig. S59. Biological activity of commercially available EPO and synthetic native folded EPO bearing three glycans (compound 1).

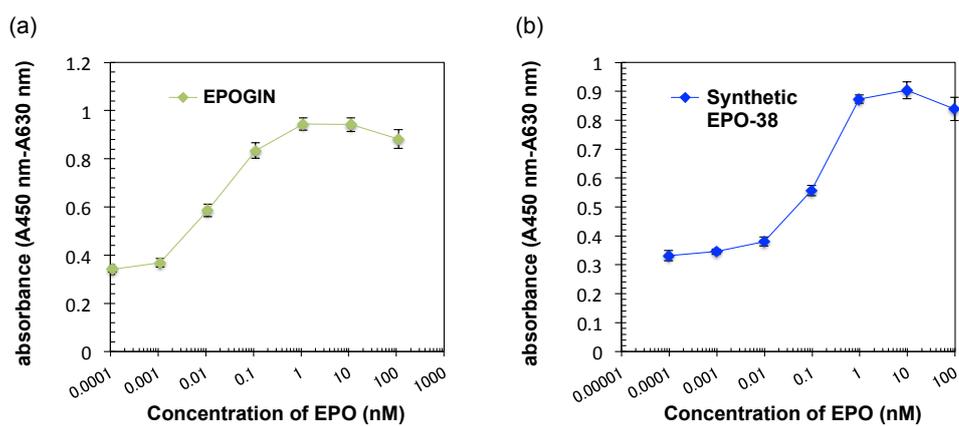


Fig. S60. a) Biological activity of commercially available EPO. b) Biological activity of synthetic EPO bearing a glycan at 38 (compound 3).

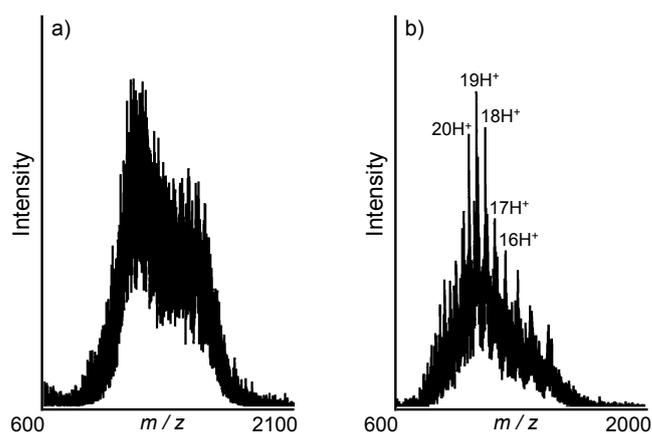


Fig. S61. Denaturing of misfolded EPO bearing a glycan at 38 (compound 7). a) ESI-MS spectrum of compound 7. b) ESI-MS spectrum of compound 7 after TCEP and guanidine treatment. However, this treatment did not give a linear glycosylpolypeptide in good yield.

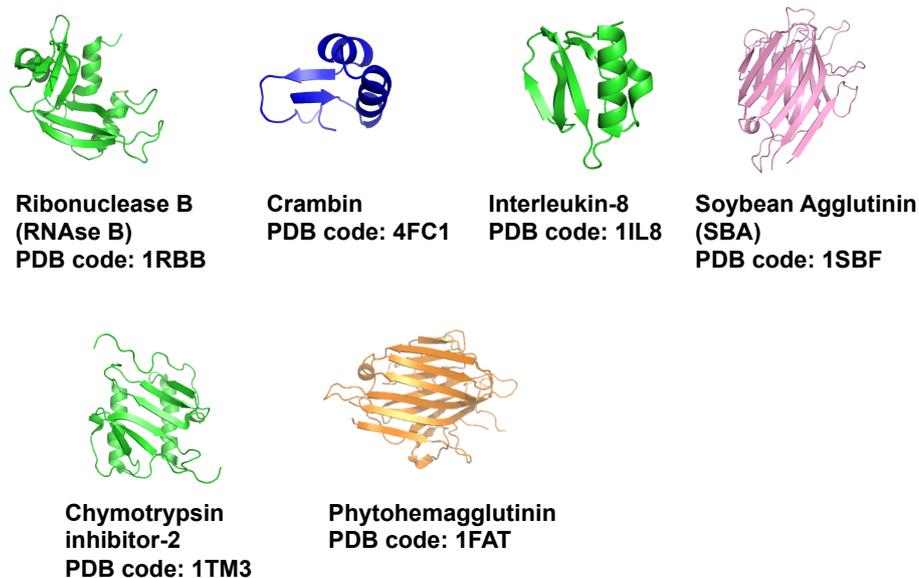


Fig. S62. Glycoproteins for UGGT substrates. The results of UGGT assays toward these glycoprotein acceptors were reported: RNase B (ref 14), Crambin (ref 15), IL-8 (ref 2), SBA (ref 16), Chymotrypsin Inhibitor-2 (ref 17), Phytohemagglutinin (ref 18).



Fig. S63. Amino acids sequence of M9-IFN β

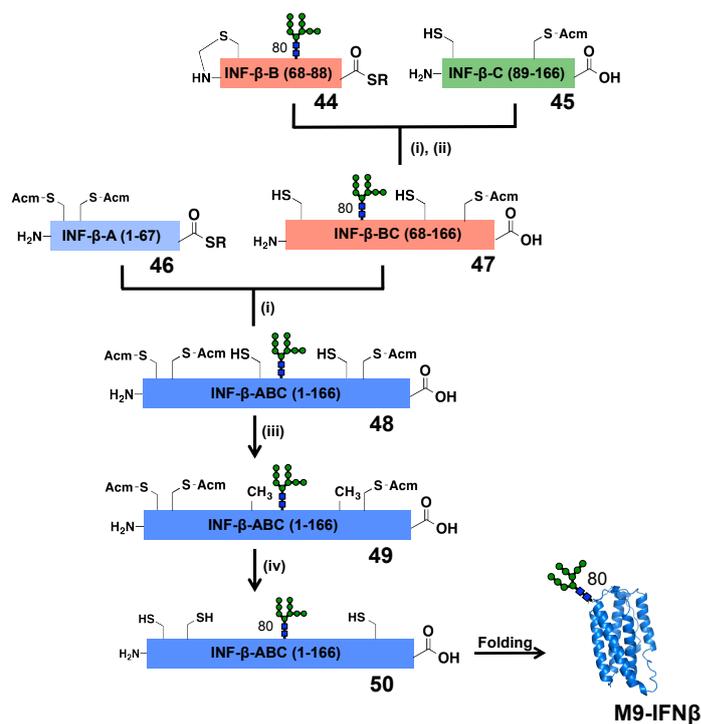


Fig. S64. Synthetic strategy of M9-IFN β . Conditions: (i) native chemical ligation; (ii) removal of thiazolidine; (iii) desulfurization; (iv) removal of the AcM groups. Synthetic procedure and synthesis of segments A and C were reported (ref 19).

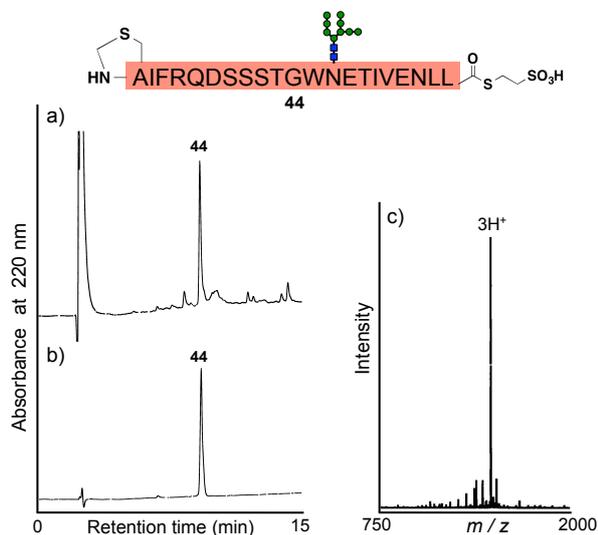


Fig. S65. HPLC profile and ESI-MS spectrum of purified glycopeptide segment B. a) HPLC profile of crude compound 44. b) HPLC profile of purified compound 44. c) ESI-MS of compound 44: m/z calcd. for $C_{177}H_{283}N_{30}O_{93}S_3$: $[M+H]^+$ 4415.5 found 4414.0 (deconvoluted).

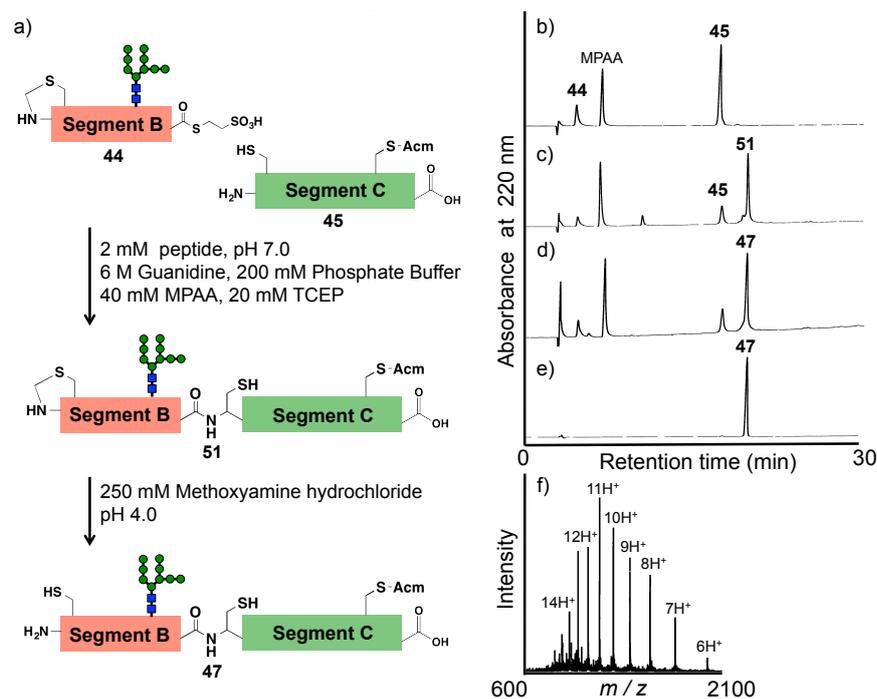


Fig. S66. 1st-NCL for the construction of M9-IFN β . a) Peptide ligation between compound **44** and **45**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 26 h. d) HPLC profile after treatment with methoxyamine-HCl. e) HPLC profile of purified compound **47**. e) ESI-MS spectrum of compound **47**. ESI-MS: m/z calcd. for $C_{615}H_{965}N_{154}O_{204}S_4$: $[M+H]^+$ 13908.4, found 13905.9 (deconvoluted).

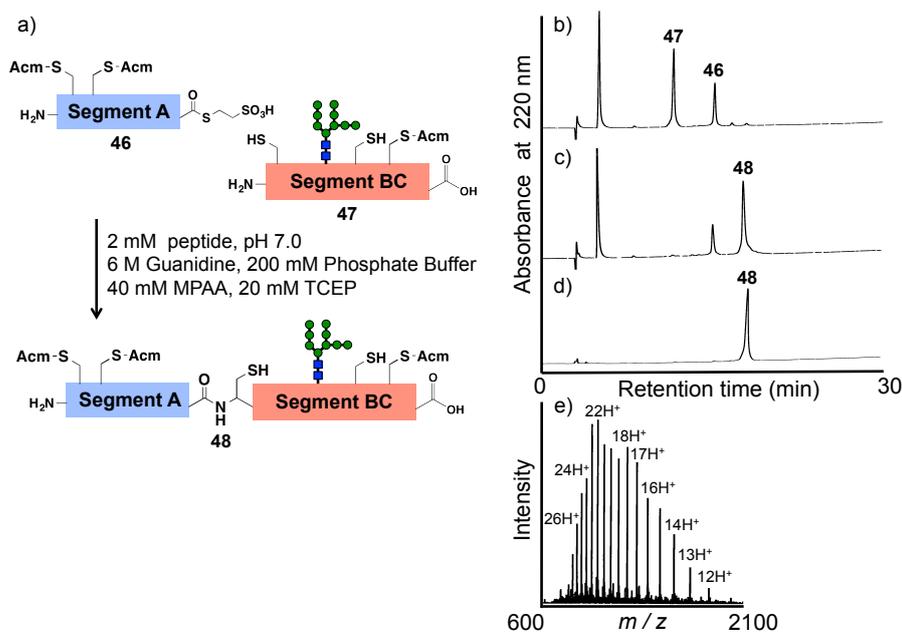


Fig. S67. 2nd-NCL for the construction of M9-IFN β . a) Peptide ligation between compound **46** and **47**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 19 h. d) HPLC profile of purified compound **48**. e) ESI-

MS spectrum of compound **48**. ESI-MS: m/z calcd. for $C_{987}H_{1540}N_{251}O_{310}S_9$: $[M+H]^+$ 22170.9, found 22168.4 (deconvoluted).

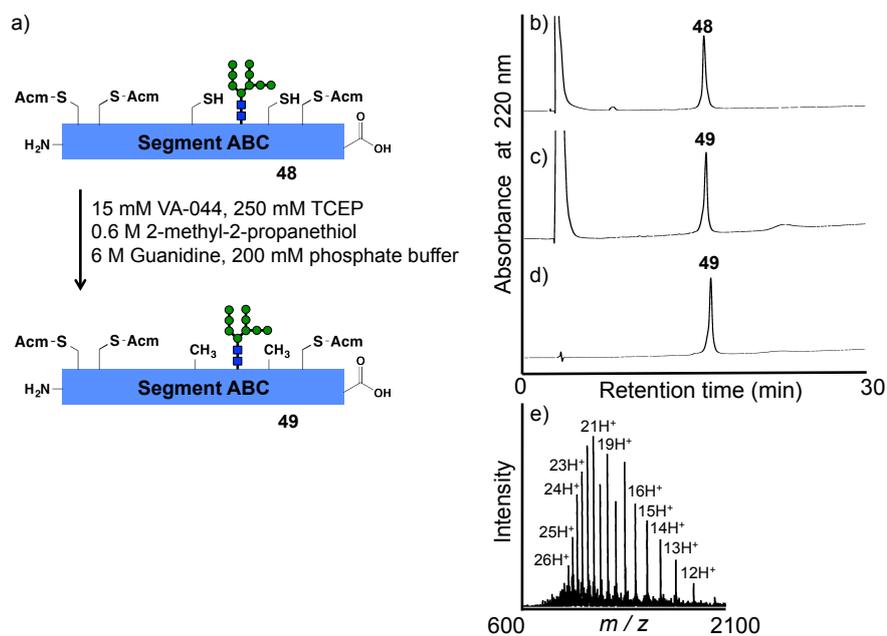


Fig. S68. Desulfurization of glycopeptide segment ABC for the construction of M9-IFN β . a) Desulfurization reaction of compound **48**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after 3 h. d) HPLC profile of purified **49**. e) ESI-MS spectrum of compound **49**. ESI-MS: m/z calcd. for $C_{987}H_{1540}N_{251}O_{310}S_7$: $[M+H]^+$ 22106.8, found 22104.6 (deconvoluted).

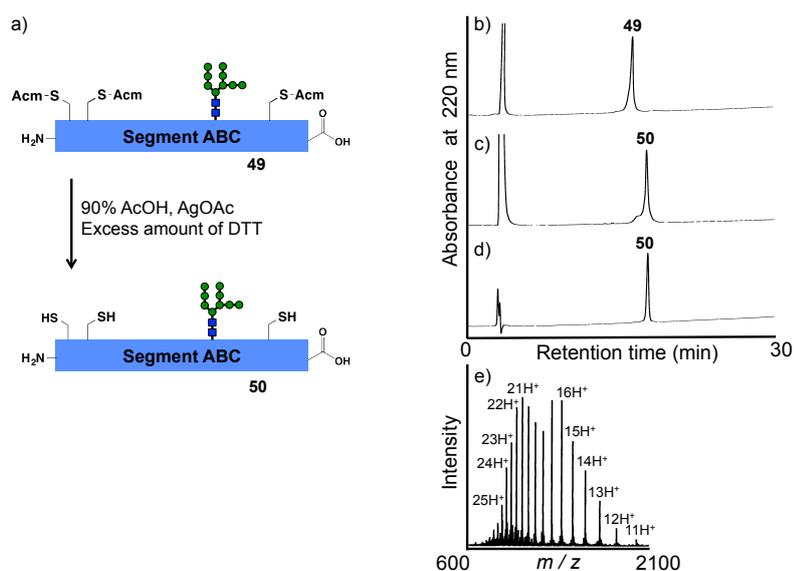


Fig. S69. Deprotection of Acm group of glycopeptide segment ABC for the construction of M9-IFN β . a) Deprotection of Acm groups of compound **49**. b) HPLC profile for starting point ($t < 1$ min). c) HPLC profile after

4 h. d) HPLC profile of purified compound **50**. e) ESI-MS spectrum of compound **50**. ESI-MS: m/z calcd. for $C_{978}H_{1525}N_{248}O_{307}S_7$: $[M+H]^+$ 21893.6, found 21890.1 (deconvoluted).

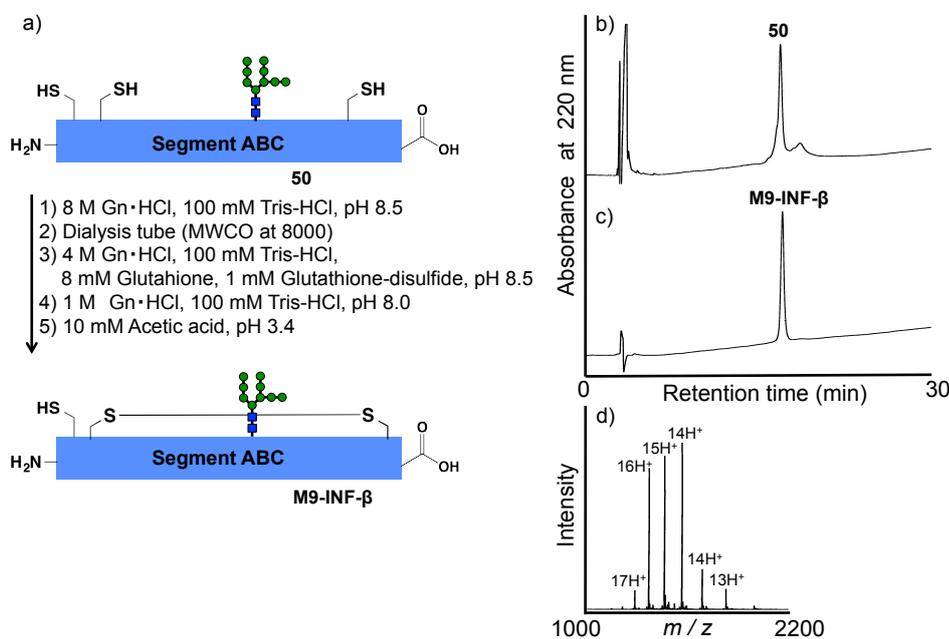


Fig. S70. Folding of segment ABC for the construction of M9-IFN β . a) Folding experiment of compound **50**. b) HPLC profile after folding process. c) HPLC profile of purified native folded M9-IFN- β . d) ESI-MS spectrum of correctly folded M9-IFN- β . FT-ICR-MS: m/z calcd. for $C_{978}H_{1523}N_{248}O_{307}S_7$: $[M+H]^+$ 21889.9489, found 21889.9801 (deconvoluted). This high-resolution mass spectrometry analysis of purified EPOs was performed with direct injections into the mass spectrometer without LC. The mass spectra were not obtained from a single time point of the LC-MS run.

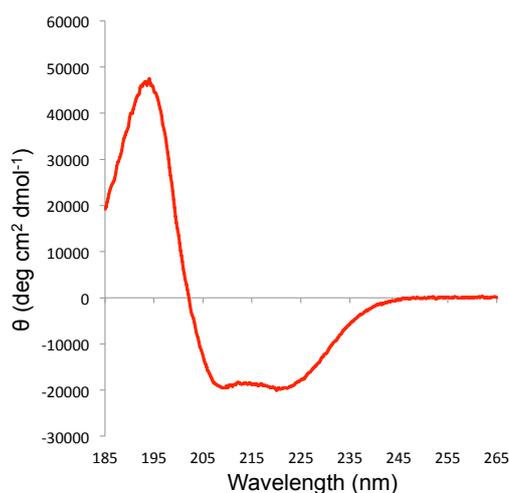


Fig. S71. Circular dichroism (CD) spectrum of synthetic M9-IFN β

References

1. Kitamura, T.; Tojo, A.; Kuwaki, T.; Chiba, S.; Miyazono, K.; Urabe, A.; Takaku, F.; Identification and analysis of human erythropoietin receptors on a factor-dependent cell line, TF-1, *Blood*, **1989**, *73*, 375–380.
2. (a) Izumi, M.; Makimura, Y.; Simone, D.; Seko, A.; Kanamori, A.; Sakono, M.; Ito, Y.; Kajihara, Y. Chemical Synthesis of Intentionally Misfolded Homogeneous Glycoprotein: A Unique Approach for the Study of Glycoprotein Quality Control, *J. Am. Chem. Soc.* **2012**, *134*, 7238–7241. (b) Izumi, M.; Oka, Y.; Okamoto, R.; Seko, A.; Takeda, Y.; Ito, Y.; Kajihara, Y. Synthesis of Glc1Man9-Glycoprotein Probes by a Misfolding/Enzymatic Glucosylation/Misfolding Sequence. *Angew Chem Int Ed* **2016**, *55*, 3968-3971.
3. Cardamone, M.; Puri, N. K. Spectrofluorimetric assessment of the surface hydrophobicity of proteins, *Biochem. J.* **1992**, *282 (Pt 2)*, 589–593.
4. Eriksson, L. C.; Torndal, U. B.; Andersson, G. N. Isolation and characterization of endoplasmic reticulum and Golgi apparatus from hepatocyte nodules in male wistar rats, *Cancer Res.* **1983**, *43*, 3335–3347.
5. (a) Bazzi, M. D.; Rabbani, N.; Duhaiman, A. S. Hydrophobicity of the NADPH binding domain of camel lens zeta-crystallin, *Biochim. Biophys. Acta.* **2001**, *1546*, 71-78. (b) Zhu, C.; Gao, Y.; Li, H.; Meng, S.; Li, L.; Francisco, J. S.; Zeng, X. C., Characterizing hydrophobicity of amino acid side chains in a protein environment via measuring contact angle of a water nanodroplet on planar peptide network. *Proc. Nat. Acad. Sci.* **2016**, *113 (46)*, 12946-12951. (c) Leon, R.; Murray, J.I.; Cragg, G.; Farnell, B.; West, N.R.; Pace, T.C.; Watson, P.H.; Bohne, C.; Boulanger, M.J.; Hof, F. *Biochemistry*, 2009, *48*, 10591-10600.

6. Izumi, M.; Murakami, M.; Okamoto, R.; Kajihara, Y. Safe and efficient Boc-SPPS for the synthesis of glycopeptide- α -thioesters, *J. Pept. Sci.* **2014**, *20*, 98–101.
7. Makimura, Y.; Kiuchi, T.; Izumi, M.; Simone, D.; Ito, Y.; Kajihara, Y. Efficient synthesis of glycopeptide- α -thioesters with a high-mannose type oligosaccharide by means of tert-Boc-solid phase peptide synthesis, *Carbohydr. Res.* **2012**, *364*, 41–48.
8. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B.; Synthesis of proteins by native chemical ligation, *Science* **1994**, *266*, 776–779.
9. Wan, Q.; Danishefsky, S. J.; Free-Radical-Based, Specific Desulfurization of Cysteine: A Powerful Advance in the Synthesis of Polypeptides and Glycopolypeptides, *Angew. Chem. Int. Ed.* **2007**, *119*, 9408–9412.
10. Kochendoerfer, G. G.; Chen, S.-Y.; Mao, F.; Cressman, S.; Traviglia, S.; Shao, H.; Hunter, C. L.; Low, D. W.; Cagle, E. N.; Carnevali, M.; Gueriguian, V.; Keogh, P. J.; Porter, H.; Stratton, S. M.; Wiedeke, M. C.; Wilken, J.; Tang, J.; Levy, J. J.; Miranda, L. P.; Crnogorac, M. M.; Kalbag, S.; Botti, P.; Schindler-Horvat, J.; Savatski, L.; Adamson, J. W.; Kung, A.; Kent, S. B. H.; Bradburne, J. A., Design and Chemical Synthesis of a Homogeneous Polymer-Modified Erythropoiesis Protein, *Science* **2003**, *299*, 884–887.
11. Narhi, L. O.; Arakawa, T.; Aoki, K.; Wen, J.; Elliott, S.; Boone, T.; Cheetham, J. Asn to Lys mutations at three sites which are N-glycosylated in the mammalian protein decrease the aggregation of Escherichia coli-derived erythropoietin, *Protein engineering* **2001**, *14*, 135–140.
12. Liu, S.; Pentelute, B. L.; Kent, S. B. H. Convergent Chemical Synthesis of [Lysine24, 38, 83] Human Erythropoietin, *Angew. Chem. Int. Ed.* **2011**, *51*, 993–999.
13. Murakami, M.; Kiuchi, T.; Nishihara, M.; Tezuka, K.; Okamoto, R.; Izumi, M.; Kajihara, Y. Chemical synthesis of erythropoietin glycoforms for insights into the relationship between glycosylation pattern and bioactivity, *Sci. Adv.* **2016**, *2*, e1500678–e1500678 (2016).
14. Trombetta, S. E.; Parodi, A. J. Purification to apparent homogeneity and partial characterization of rat liver UDP-glucose: glycoprotein glucosyltransferase, *Journal of Biological Chemistry* **1992**, *267*, 9236–9240.

15. Dedola, S.; Izumi, M.; Makimura, Y.; Seko, A.; Kanamori, A.; Sakono, M.; Ito, Y.; Kajihara, Y. Folding of synthetic homogeneous glycoproteins in the presence of a glycoprotein folding sensor enzyme, *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 2883–2887.
16. Sousa, M.; Parodi, A. J.; The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc: glycoprotein glucosyltransferase, *EMBO J.* **1995**, *14*, 4196.
17. Caramelo, J. J.; Castro, O. A.; Alonso, L. G.; de Prat-Gay, G.; Parodi, A. J. UDP-Glc: glycoprotein glucosyltransferase recognizes structured and solvent accessible hydrophobic patches in molten globule-like folding intermediates, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 86–91.
18. Sousa, M. C.; Ferrero-Garcia, M. A.; Parodi, A. J. Recognition of the oligosaccharide and protein moieties of glycoproteins by the UDP-Glc: glycoprotein glucosyltransferase, *Biochemistry* **1992**, *31*, 97–105.
19. Sakamoto, I.; Tezuka, K.; Fukae, K.; Ishii, K.; Taduru, K.; Maeda, M.; Ouchi, M.; Yoshida, K.; Nambu, Y.; Igarashi, J.; Hayashi, N.; Tsuji, T.; Kajihara, Y. Chemical Synthesis of Homogeneous Human Glycosyl-interferon- β That Exhibits Potent Antitumor Activity in Vivo, *J. Am. Chem. Soc.* **2012**, *134*, 5428–5431.