CHEMISTRY
A EUROPEAN JOURNAL

DOI: 10.1002/chem.200500531

Molecular Design of Glycoprotein Mimetics: Glycoblotting by Engineered Proteins with an Oxylamino-Functionalized Amino Acid Residue

Naoki Matsubara,[a] Kei Oiwa,[a] Takahiro Hohsaka,[b] Reiko Sadamoto,[a] Kenichi Niikura,[a] Norio Fukuhara,[a] Akio Takimoto,[c] Hirosato Kondo,[c] and Shin-Ichiro Nishimura*[a]

Abstract: The general and efficient method for the site-directed glycosylation of proteins is a key step in order to understand the biological importance of the carbohydrate chains of proteins and to control functional roles of the engineered glycoproteins in terms of the development of improved glycoprotein therapeutics. We have developed a novel method for site-directed glycosylation of proteins based on chemoselective blotting of common reducing sugars by genetically encoded proteins. The oxylamino-functionalized L-homoserine residues, 2-amino-4-*O*-(*N*-methylaminooxy) butanoic acid and 2-amino-4-aminooxy butanoic acid, were efficiently incorporated into proteins by using the four-base codon/anticodon pair strategy in *Escherichia coli* in vitro translation. Direct and chemoselective coupling between unmodified simple sugars and N-methylaminooxy group displayed on the engineered streptavidin allowed for the combinatorial synthesis of novel glycoprotein mimetics.

Keywords: carbohydrates · glycoprotein · glycosylation · mass spectrometry

Introduction

Glycosylation is one of the most important processes for post-translational modifications in proteins, because this modification step affects protein folding and stability, modifies the intrinsic biological activity of proteins, and regulates their target molecules/cells to be recognized in the biological systems.[1] However, it is well known that naturally occurring glycoproteins are often present as a population of many different and heterogeneous glycoforms that makes the investigation of glycosylation effects on protein structure and function difficult. Therefore, the general and efficient method for the site-directed glycosylation of proteins is a key step both to understand biological importance of the carbohydrate chains of proteins and to control functional roles of the engineered glycoproteins in terms of the development of improved glycoprotein therapeutics.[2] Recently, a pioneering work by Schultz and co-workers provided a general strategy for the cotranslational synthesis of selectively glycosylated proteins in which the modified amino acids carrying β-GlcNAc and α-GalNAc residues are genetically encoded in *Escherichia coli*.[3,4] This method requires the evolutional procedure for the isolation of an orthogonal *Methanococcus jannaschii* tyrosyl tRNA synthetase (MjTyrRS) that specifically charges the corresponding *M. jannaschii* suppressor ($mutRNA^{Tyr}_{CUA}$) with these sugar amino acids in response to the amber codon (TAG) in *E. coli*.

Our interest has been focused on the high-throughput generation of glycoprotein and glycopeptide mimetics in order to investigate the effect of the site-specific glycosylation on the protein structure and functions. This will become a key and rate-limiting step for the discovery of the new generation glycoprotein drugs. We have established a facile method for the site specific introduction of the ω-aminoalkyl glycosides based on the transamination reaction by transglutaminases at designated glutamine residues of the synthetic peptides[5] or the mutant proteins.[6] A keto-containing

[a] N. Matsubara, K. Oiwa, Dr. R. Sadamoto, Dr. K. Niikura, Prof. Dr. N. Fukuhara, Prof. Dr. S.-I. Nishimura
Laboratory for Bio-Macromolecular Chemistry
Division of Biological Sciences, Graduate School of Science
Frontier Research Center for Post-Genome Science and Technology
Hokkaido University, Sapporo 001-0021 (Japan)
Fax: (+81) 11-706-9042
E-mail: shin@glyco.sci.hokudai.ac.jp

[b] Dr. T. Hohsaka
School of Material Science
Japan Advanced Institute of Science and Technology
1-1 Asahidai, Tatsunokuchi Ishikawa 923-1292 (Japan)

[c] Dr. A. Takimoto, Dr. H. Kondo
Discovery Research Laboratories
Shionogi & Co. Ltd.
Osaka 553-0002 (Japan)

amino acid, *p*-acetyl-L-phenylalanine, has been successfully incorporated in response to the amber nonsense codon[7] and this unnatural amino acid can be used for the subsequent coupling with aminooxy saccharide derivatives[8] to generate glycoprotein mimetics. Although these approaches made the incorporation of carbohydrates into the engineered proteins possible, both strategies definitely need chemically modified sugar derivatives having an appropriately functionalized group at each anomeric position. In the present study, we report the feasibility of the frameshift suppression strategy on the basis of the four-base codon/anticodon system[9] in the production of glycoprotein mimetics. In case of the strategy using expanded genetic codes, it has been suggested that many kinds of unnatural amino acids can be incorporated into proteins by using the chemically aminoacylated frameshift suppressors in in vitro translation systems in response to both various four- and five-base codons.[10] We thought that this approach can be used for the production of the specific precursor proteins having unnatural amino acids that allow the direct glycosylation (glycoblotting)[11] with a variety of unmodified sugars (naturally occurring oligosaccharides) at the desired positions. It should be noted that this strategy permits combinatorial synthesis of glycoprotein mimetics from a single precursor mutant protein. In addition, combined use of this method with the common procedure using the noncoding triplet codons[12] will allow the incorporation of two or more different carbohydrates into distinct sites of a single polypeptide.

Results and Discussion

Synthesis of aminoacyl-tRNA carrying unnatural amino acids: As a versatile nonproteinogenic functional group with a unique reactivity, we selected the oxylamino groups because it has been demonstrated that the synthetic peptides bearing this functional group can be used for generating neoglycopeptides.[13,14] Recently, it was also reported that combined use of the glycoblotting by oxylamino-containing polymers and MALDI-TOF/TOF mass spectrometry allows for both facile purification and precise analysis of common oligosaccharides and glycopeptides from human serum or mouse skin glycoproteins.[11,15] In this study, two unnatural amino acids, 2-amino-4-*O*-(*N*-methylaminooxy) butanoic acid (**1**) and 2-amino-4-aminooxy butanoic acid (**2**), were employed for the synthesis of novel aminoacyl-tRNA (Figure 1). In order to

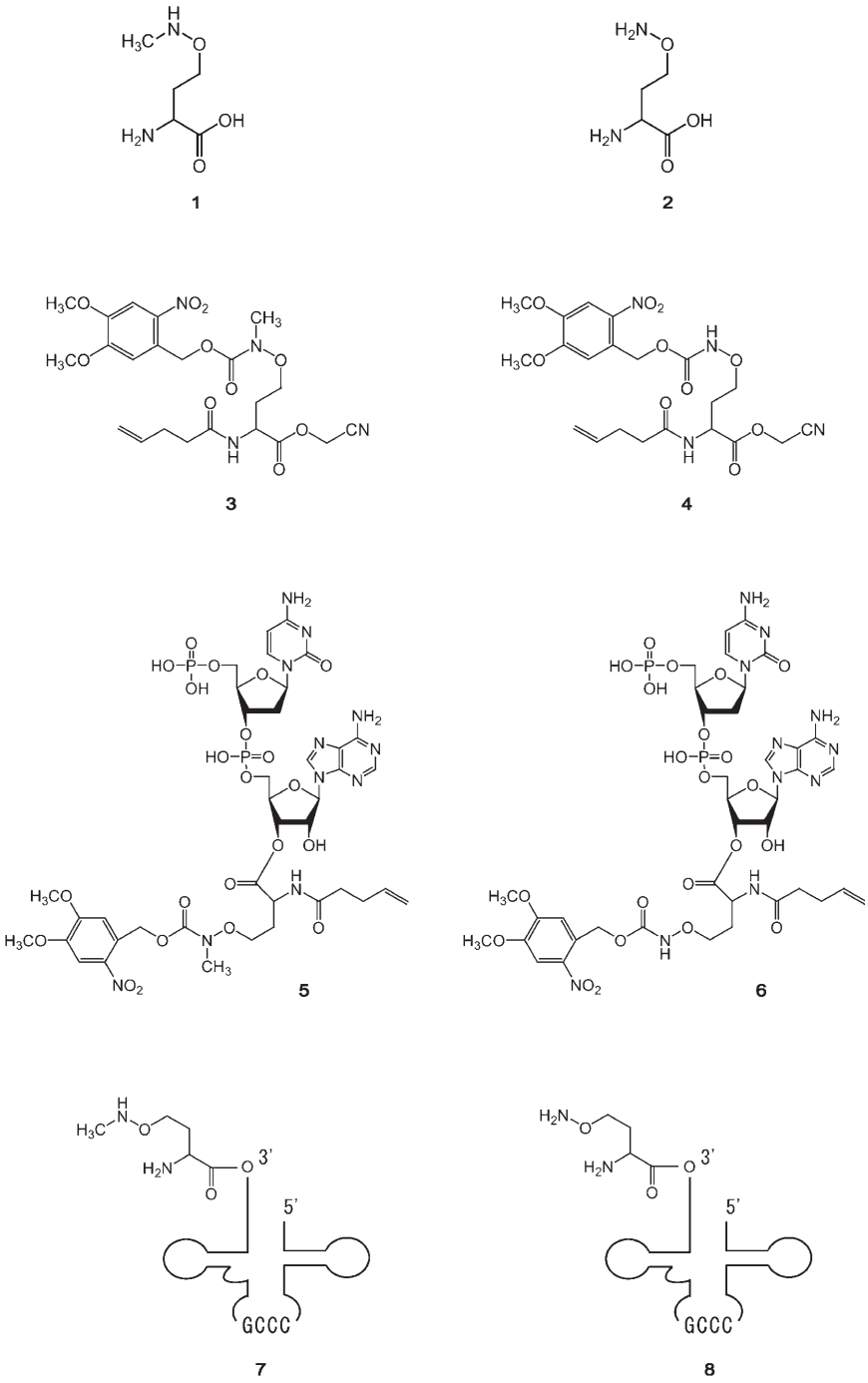
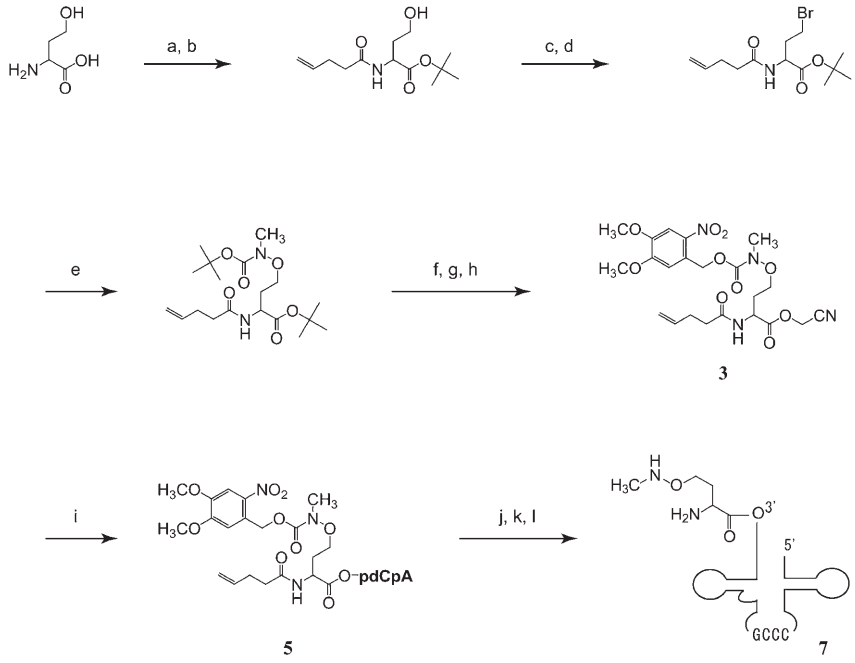


Figure 1. Chemical structures of the key compounds used in this study.

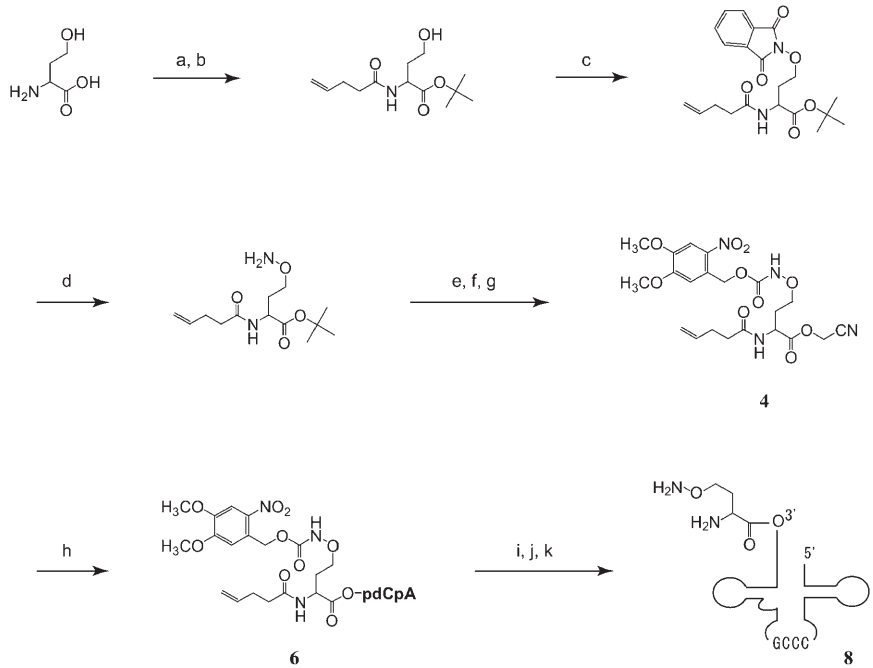
prepare the aminoacyl-tRNA$_{CCCG}$ charged with these unnatural amino acids, key intermediates **3** and **4** were synthesized from L-homoserine according to the synthetic procedures indicated in Schemes 1 and 2. Compounds **3** and **4** were coupled with pdCpA[16] to give the aminoacyl-pdCpA derivatives **5** and **6**. Next, they were employed for the conjugation with tRNA$_{CCCG}$ in the presence of T4 RNA ligase according to the method reported by Hecht[17] and the subsequent deprotection afforded the target unnatural amino acid-tRNA$_{CCCG}$, Hser(ONHMe)-tRNA$_{CCCG}$ (**7**) and Hser-(ONH$_2$)-tRNA$_{CCCG}$ (**8**) in good yield.

In vitro transcription/translation: Position specific incorporation of the unnatural amino acids by in vitro transcription/translation was preliminarily carried out according to the protocols reported previously.[10] As expected, in vitro transcription/translation using *E. coli* S30 extract in the presence of **7** or **8** proceeded smoothly and permitted position specific incorporation of these two unnatural amino acids into T4 phage lysozyme and streptavidin. Figure 2a shows the Western blot analysis of the incorporation of **1** and **2** into the Gln45 position of the T4 phage lysozyme. These results suggested that the efficiency of the incorporation was estimated to be 17% (compound **1**) and 12% (compound **2**), respectively. Similarly, these amino acids were incorporated into the Tyr83 position of the full-length streptavidin with 21% (compound **1**) and 15% (compound **2**) as shown in Figure 2b. Considering the expression levels of wild-type T4 phage lysozyme and streptavidin produced by the present in vitro transcription/translation system (Figure 2c) and the satisfactory purity of the protein checked by a silver staining (Figure 2d), we employed the mutant streptavidin having an unusual amino acid residue **1** for further modification study.

Glycoblotting by engineered streptavidin: The N-methyl oxylamino group incorporated into the mutant streptavidin



Scheme 1. a) 4-Pentenoic anhydride, Et_3N, H_2O/MeOH 3:2; b) BTEAC, K_2CO_3, *t*BuBr, DMF (53% over two steps); c) MsCl, Et_3N, CH_2Cl_2; d) LiBr, acetone (81% over two steps); e) NaH, *N*-methyl-*N*-*tert*-butoxycarbonylhydroxylamine, DMF (69%); f) TFA; g) NVOC-Cl, Et_3N, H_2O/dioxane 1:1; h) $CNCH_2Cl$, Et_3N, CH_3CN (82% over three steps); i) pdCpA, DMF (62%); j) tRNA$_{CCCG}$-CA, T4 RNA ligase; k) UV irradiation; l) I_2.



Scheme 2. a) 4-Pentenoic anhydride, Et_3N, H_2O/MeOH 3:2; b) BTEAC, K_2CO_3, *t*BuBr, DMF (53% over two steps); c) *N*-hydroxyphthalimide, Ph_3P, DEAD, THF (84%); d) $MeNH_2$/MeOH (2:3) (89%); e) TFA; f) NVOC-Cl, Et_3N, H_2O/Dioxane (1:1); g) $CNCH_2Cl$, Et_3N, CH_3CN (11% over three steps); h) pdCpA, DMF (39%); i) tRNA$_{CCCG}$-CA, T4 RNA ligase; j) UV irradiation; k) I_2. BTEAC = benzyltriethylammonium chloride.

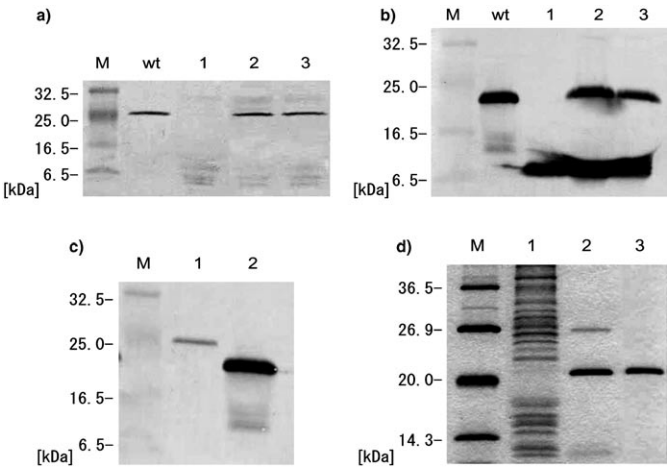


Figure 2. Incorporation of the unnatural amino acids into proteins using in vitro transcription/translation strategy. a) Western blot analysis of the expression of the T4 phage lysozyme containing CGGG at the Gln45: lane M, molecular weight markers; lane wt, T4 phage lysozyme (wild-type); lane 1, in the absence of aminoacyl-$tRNA_{CCCG}$; lane 2, in the presence of compound **7** [Hser(ONHMe)-$tRNA_{CCCG}$]; lane 3, in the presence of compound **8** [Hser(ONH_2)-$tRNA_{CCCG}$]. b) Western blot analysis of the expression of the streptavidin containing CGGG at the Tyr83: lane M, molecular weight marker; lane wt, streptavidin (wild-type); lane 1, in the absence of aminoacyl-$tRNA_{CCCG}$; lane 2, in the presence of compound **7** [Hser(ONHMe)-$tRNA_{CCCG}$]; lane 3, in the presence of compound **8** [Hser(ONH_2)-$tRNA_{CCCG}$]. c) Comparison of the expression level between T4 phage lysozyme and streptavidin. lane M, marker proteins; lane 1, wild-type T4 phage lysozyme; lane 2, wild-type streptavidin. d) Silver staining analysis of the mutant streptavidin containing unnatural amino acid **1**: lane M, molecular weight markers; lane 1, in vitro translation mixture; lane 2, a fraction obtained by T7-tag affinity chromatography; lane 3, a fraction obtained by purification in combination with His-tag affinity chromatography and T7-tag affinity chromatography.

was tested for subsequent glycoblotting (chemoselective ligation with unmodified mono- and oligosaccharides). The general protocols for the glycoblotting experiments are described in the following: Sugar solutions (final concentration; 100–400 mM) were added to a solution of the mutant protein (200 ng) dissolved in 100 mM acetate buffer (pH 4.3, 3.0 μL) and the reaction mixture was incubated at 37 °C for 44 h. As shown in Figure 3, glycoblotting by the engineered streptavidin (Figure 3a) of Glc (D-glucose, Figure 3b), Glcα1,4Glc (maltose, Figure 3c), Glcα1,4Glcα1,4Glc (maltotriose, Figure 3d), GlcNAc (*N*-acetyl-D-glucosamine, Figure 3e), and Galβ1,4Glc (lactose, Figure 3f) were observed by MALDI-TOF MS. In case of the wild-type streptavidin as a negative control, no unfavorable side-reaction such as non-specific glycation at lysine residues was detected. The unmodified protein bearing N-methyl oxylamino group can be removed by trapping with aldehyde-functionalized polymers; this step will make a practical and efficient purification of the target glycoprotein mimetics possible (data not shown).

In addition, it was also revealed that the streptavidin modified with lactose (*m*/*z* 19208.7) can be converted by recombinant α2,3-sialyltransferase into a novel glycoprotein mimetic having GM3 trisaccharide at the position 83 (*m*/*z* 19488.3, Figure 3g). Relatively low yields found for this enzymatic sialylation reaction might be due to the heterogeneous reaction conditions by employing precursor glycoprotein (lactosylated streptavidin) immobilized on the magnetic Ni beads. Table 1 summarizes the molecular weights characterized by MALDI-TOF MS of the glycoprotein mimetics studied in these experiments. Although further optimization study must be carried out for achieving much higher modification yield, the glycoblotting by the engineered proteins having N-methyl oxylamino group will greatly contribute to the construction of the library of artificially glycosylated proteins in terms of the site-directed modifications with a variety of carbohydrates.

Conclusion

We could successfully establish a general method for the synthesis of homogeneous glycoprotein mimetics bearing simple sugar chains at the desired position based on the unique glycoblotting strategy. Since the availability of non-coding triplet codons ultimately limits the number of amino acids encoded by any organism, the present method might become an alternative or a complementary way for the production of a variety of glycoprotein mimetics. Although it remains to be discussed whether this site-specific glycosylation through the non-natural linkage influences the structure and functions of the glycoprotein mimetics, this method will greatly contribute to a more rapid and high-throughput assay system for searching new types of glycosylated protein drugs in combinatorial manner. We are now investigating the effect of this non-natural linkage on the conformation of the glycoprotein by means of NMR/CD spectroscopy. The results of both structural and biological characterization of glycoprotein mimetics will be reported in the near future.

Experimental Section

General methods and materials: T7-Tag antibody, T7-Tag antibody agarose was obtained form Novagen. *E. coli* T7 S30 extract system for circular DNA, ProtoBlot II AP System with stabilized substrate mouse, and MagneHis Ni-particles were purchased from Promega Co. T4 RNA ligase, protein markers, and prestained protein markers were from New England Biolabs. Immun-Blot PVDF membrane for protein blotting was obtained from Bio-Rad. A T4 phage lysozyme gene was purchased from TAKARA Bio Inc. Plasmids (pGSH), encoding a sequences of T7 promoter, a T7 tag at the N-terminal, a $(His)_6$ tag at the C-terminal, and a T7 terminator, using for expression of the streptavidin and T4 phage lysozyme, each of wild-type and mutant including a CGGG codon, and $tRNA_{CCCG}$(-CA) (the insertion of the CCCG anticodon at the anticodon loop and the lack of a CA di-nucleotide unit at the 3′ end) were prepared according to the methods reported previously [T. Hohsaka, D. Kajihara, Y. Ashizuka, H. Murakami, M. Sisido, *J. Am. Chem. Soc.* **1999**, *121*, 34–40]. Chemical reactions were monitored by thin-layer chromatography (TLC) on pre-coated plates of Silica Gel $60F_{254}$ (E. Merck). Visualization was accomplished with UV light (254 nm) and treatment with a solution of $(NH_4)_6Mo_7O_2{\cdot}4H_2O$ (20 g) and $Ce(SO_4)_2$ (0.4 g) in 10% sulfuric acid (400 mL) and heating at 150 °C or ninhydrin reagent (Wako Pure Chemical Industries, Ltd.). Enzymatic reactions in aminoacyl-tRNA synthesis

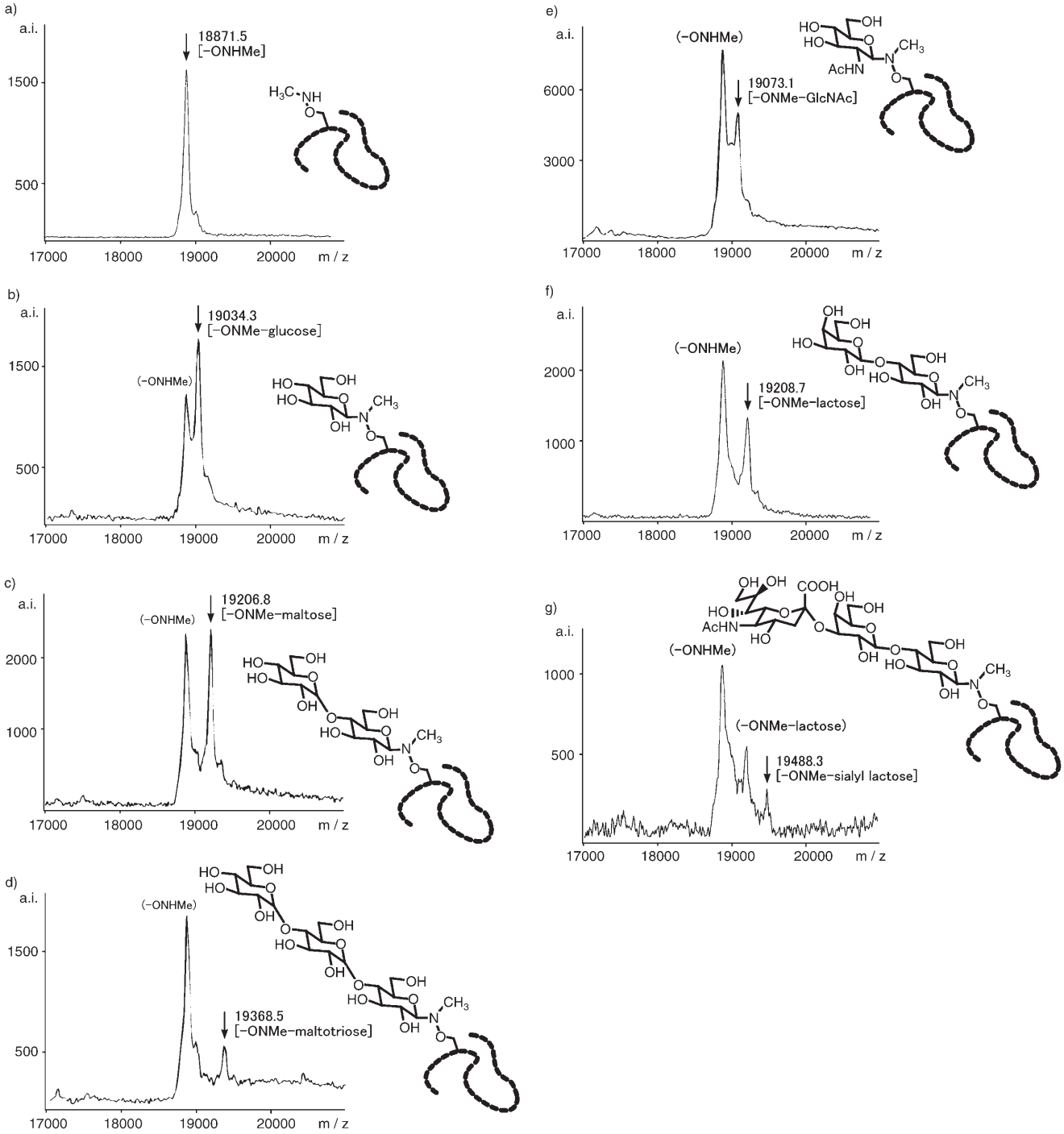


Figure 3. MALDI-TOF MS analysis of the molecular weight of mutant streptavidin and glycoprotein mimetics: a) mutant streptavidin having an unnatural amino acid **1**; product obtained by glycoblotting of mutant streptavidin with b) glucose, c) maltose, d) maltotriose, e) *N*-acetylglucosamine, f) lactose; g) modification of the mutant streptavidin with glycosyltransferase.

were monitored by 15% PAGE with 7 M Urea. Purifications of the synthetic product by column chromatography were carried out by using silica gel 60N (KANTO Co., 63–210 nm mesh). All other reagents were purchased from Sigma-Aldrich Co. Ltd., Wako Pure Chemical Industries, Ltd., Tokyo Kasei Co. Ltd. and Nacalai Tesque, Inc.

NMR spectroscopy: ^{1}H and ^{13}C NMR spectra were recorded and measured at 600 and 120 MHz, respectively, with AVANCE 600 (Bruker) using $CDCl_3$ as a solvent. All signals of the new compounds synthesized in this study were characterized with H,H COSY and HMQC techniques.

Mass spectrometry: FAB mass spectra were obtained with JMS-HX100 (JEOL) using glycerol or *m*-nitrobenzyl alcohol as a matrix. ESI mass spectra were obtained with JMS-700TZ (JEOL). MALDI-TOF MS were measured with ultraflex and biflex (Bruker). Samples were desalted and concentrated using 10 μL C4 ZipTip (Millipore) according to the manu-

Table 1. Theoretical and observed molecular weights of glycosylated streptavidin.

Additional carbohydrate	Calcd mass	Found
none	18875.3	18871.5[a]
glucose	19037.5	19034.3[a]
maltose	19199.6	19206.8[a]
maltotriose	19361.7	19368.5[a]
GlcNAc	19078.5	19073.1[a]
lactose	19199.6	19208.7[a]
sialyllactose	19490.9	19488.3[b]

[a] $[M+H]^+$. [b] $[M-H]^-$.

facturer's instruction. Typically, samples were mixed with the 1 μL of 2,5-dihydroxybenzoic acid (DHB) in 33% acetonitrile containing 0.1% trifluoroacetic acid at the concentration of 10 mg mL^{-1}. The theoretical mass was calculated an average mass of a protein without methionine. The loss of the N-terminal methionine is common in *E. coli* expression system.

Synthesis of aminoacyl pdCpA derivatives

2-(Pent-4-enoyl)amino-4-hydroxybutanoic acid *tert*-butyl ester: 4-Pentenoic anhydride (3.7 mL, 20.2 mmol) and triethylamine (2.8 mL, 20.2 mmol) was added to an ice-cooled solution of L-homoserine (2.0 g, 16.8 mmol) in H_2O (6 mL) and methanol (4 mL). The reaction mixture was stirred at 0°C for 2 h and then concentrated. The crude residue was purified by flash column chromatography (chloroform/methanol/H_2O 65:15:1). The concentrated residue was dissolved in *N*,*N*-dimethylformamide (25 mL) and added to benzyltriethylamine chloride (4.0 g, 17.6 mmol). After being stirred for 5 min, K_2CO_3 (22.1 g, 160 mmol) and *tert*-butyl bromide (26.5 mL, 230 mmol) were added. The reaction mixture was stirred at 55°C for 6 h and then added ice water. The mixture was extracted with ethyl acetate and the organic layer was washed with brine, dried using $MgSO_4$, filtrated, and concentrated. The residual syrup was purified by flash column chromatography (hexane/ethyl acetate 3:2) to give product (2.3 g, 53%). R_f = 0.50 (hexane/ethyl acetate 1:2). ^{1}H NMR (600 MHz, $CDCl_3$): δ = 6.39 (d, J = 7.20 Hz, 1H, NH), 5.85–5.78 (m, 1H, C=CH-), 5.11–5.02 (m, 2H, -C=CH_2), 4.65–4.60 (m, 1H, α-H), 3.80–3.77 (m, 1H, -OH), 3.71–3.65 (m, 1H, γ-H_2-a), 3.56–3.50 (m, 1H, γ-H_2-b), 2.45–2.35 (m, 4H, C=C-CH_2-, C=C-C-CH_2-), 2.21–2.15 (m, 1H, β-H_2-a), 1.54–1.50 (m, 1H, β-H_2-b), 1.48 (s, 9H, $(CH_3)_3$-C-); HR-FAB MS: *m*/*z*: calcd for $C_{13}H_{23}NO_4$: 257.1627; found: 258.1709 $[M+H]^+$.

2-(Pent-4-enoyl)amino-4-bromobutanoic acid *tert*-butyl ester: A solution of 2-(pent-4-enoyl)amino-4-hydroxybutanoic acid *tert*-butyl ester (85 mg, 0.33 mmol) in CH_2Cl_2 (1.5 mL) was cooled at 0°C under nitrogen atmosphere. Triethylamine (55 μL, 0.40 mmol) and methanesulfonyl chloride (27 μL, 0.35 mmol) were added, and the solution was stirred for 1 h. Lithium bromide (207 mg, 2.39 mmol) and acetone (1.0 mL) were added to the reaction mixture, and the mixture was stirred for 19 h. The mixture was concentrated and extracted with ethyl acetate. The organic layer was washed with sat. $NaHCO_3$, brine, dried using $MgSO_4$, filtrated, and concentrated. The residual syrup was subjected to the purification by flash column chromatography (hexane/ethyl/acetate 4:1) to give product (86 mg, 81%). R_f = 0.50 (hexane/ethyl acetate 2:1); ^{1}H NMR (600 MHz, $CDCl_3$): δ = 6.17 (d, J = 7.37 Hz, 1H, NH), 5.85–5.78 (m, 1H, C=CH-), 5.09–5.01 (m, 2H, -C=CH_2), 4.59–4.56 (m, 1H, α-H), 3.41–3.32 (m, 2H, γ-H_2), 2.45–2.31 (m, 5H, C=C-CH_2-, C=C-C-CH_2-, β-H_2-a), 2.22–2.17 (m, 1H, β-H_2-b), 1.47 (s, 9H, $(CH_3)_3$-C-); HR-FAB MS: *m*/*z*: calcd for $C_{13}H_{22}NO_3Br$: 319.0783; found: 320.0862 $[M+H]^+$.

2-(Pent-4-enoyl)amino-4-[*O*-(*N*-methyl-*N*-*tert*-butoxycarbonyl)amino]hydroxybutanoic acid *tert*-butyl ester: NaH (60% dispersion in mineral oil, 13 mg, 0.32 mmol) was added to a solution of 2-(Pent-4-enoyl)amino-4-bromobutanoic acid *tert*-butyl ester (47 mg, 0.32 mmol) in *N*,*N*-dimethylformamide (1 mL) and the mixture was stirred for 5 min under nitrogen atmosphere, then cooled at 0°C. *N*-Methyl-*N*-*tert*-butoxycarbonylhydroxylamine (86 mg, 0.27 mmol) was added to the solution, and the mixture was stirred at 0°C for 1 h. To the mixture was added ice water and the mixture was extracted with ethyl acetate. The organic layer was washed with brine, dried using $MgSO_4$, filtrated, concentrated, and purified by flash column chromatography (hexane/ethyl acetate 4:1) to give product (72 mg, 69%). R_f = 0.28 (hexane/ethyl acetate 3:1); ^{1}H NMR (600 MHz, $CDCl_3$): δ = 7.37 (s, 1H, NH), 5.88–5.81 (m, 1H, C=CH-), 5.08–4.97 (m, 2H, -C=CH_2), 4.63–4.60 (m, 1H, α-H), 3.93–3.86 (m, 2H, γ-H_2), 3.07 (s, 1H, N-CH_3), 2.43–2.33 (m, 4H, C=C-CH_2-, C=C-C-CH_2-), 2.13–2.02 (m, 2H, β-H_2), 1.48, 1.44 (all s, 18H, 2×$(CH_3)_3$-*C*-); ^{13}C NMR (120 MHz, $CDCl_3$): δ = 172.42, 171.09, 157.00 (3C, 3×C=O), 137.27 (C=*C*-), 115.09 (*C*=C-), 81.69, 81.60 (2×$(CH_3)_3$-*C*-), 70.52 (γ-C), 50.58 (α-C), 36.66 (*N*-CH_3), 35.60, 29.75 (2C, C=*C*-*C*-C), 29.60 (β-C), 28.20, 27.89 (2C, 2×$(CH_3)_3$-C-); HR-FAB MS: *m*/*z*: calcd for $C_{19}H_{34}N_2O_6$: 386.2417; found: 387.2487 $[M+H]^+$.

2-(Pent-4-enoyl)amino-4-[*O*-[*N*-methyl-(*N*-6-nitroveratryloxycarbonyl)]amino]hydroxybutanoic acid cyanomethyl ester (3): 2-(Pent-4-enoyl)amino-4-[*O*-(*N*-methyl-*N*-*tert*-butoxycarbonyl)amino]hydroxybutanoic acid *tert*-butyl ester (70 mg, 0.18 mmol) was dissolved in trifluoroacetic acid (1 mL) and stirred for 1 h. The reaction mixture was concentrated, then dissolved in a solution of H_2O (400 μL) and dioxane (400 μL), and cooled at 0°C. To the mixture was added triethylamine (88 μL, 0.63 mmol) and 6-nitroveratryloxycarbonyl chloride (NVOC-Cl; 60 mg, 0.22 μmol). The mixture was stirred at 0°C for 2 h in the dark, then extracted with ethyl acetate. The organic layer was washed with 5% citric acid and brine, dried using $MgSO_4$, filtrated, concentrated, and purified by flash column chromatography (chloroform/methanol 30:1). The concentrated residue was dissolved in acetonitrile (0.5 mL) and cooled at 0°C under nitrogen atmosphere. Triethylamine (90 μL, 0.64 mmol) and chloroacetonitrile (0.1 mL, 1.6 mmol) were added to the solution and the reaction mixture was stirred at room temperature for 24 h. The mixture was extracted with ethyl acetate and the extract was washed with brine, dried using $MgSO_4$, filtrated, concentrated, and purified by flash column chromatography (hexane/ethyl acetate 2:3) to give compound **3** (68 mg, 82%). R_f = 0.40 (hexane/ethyl acetate 1:2); ^{1}H NMR (600 MHz, $CDCl_3$): δ = 7.64 (s, 1H, aromatic), 6.92 (s, 1H, aromatic), 5.79–5.72 (m, 1H, C=CH-), 5.54–5.48 (m, 2H, O-CH_2-Ph), 5.01–4.91 (m, 2H, -C=CH_2), 4.80–4.76 (m, 1H, α-H), 4.72–4.65 (m, 2H, O-CH_2-CN), 3.91, 3.90 (all s, 6H, 2×OCH_3), 3.94–3.89 (m, 2H, γ-H_2), 3.14 (s, 3H, N-CH_3), 2.35–2.29 (m, 4H, C=C-CH_2-, C=C-C-CH_2-), 2.24–2.19 (m, 1H, β-H_2-a), 2.06–2.01 (m, 1H, β-H_2-b); ^{13}C NMR (120 MHz, $CDCl_3$), δ172.84, 170.66, 157.16 (3C, 3×C=O), 153.48, 148.54, 140.19, 126.49, 110.93, 108.32 (6C, aromatic), 137.04 (C=*C*-C), 115.38 (*C*=C-C), 114.03 (-CN), 70.06 (γ-C), 65.12 (O-*C*-Ph), 56.47, 56.45 (2C, 2×OCH_3), 49.50 (α-C), 48.79 (O-*C*-CN), 36.54 (N-CH_3), 35.28, 29.38 (2C, C=*C*-*C*-C), 28.96 (β-C); HR-FAB MS: *m*/*z*: calcd for $C_{22}H_{28}N_4O_{10}$: 508.1805; found: 509.1877 $[M+H]^+$.

2-(Pent-4-enoyl)amino-4-[*O*-[*N*-methyl-(*N*-6-nitroveratryloxycarbonyl)]amino]hydroxybutanoic acid pdCpA ester (5): The aminoacylation of pdCpA was carried out by adding the compound **3** (1.0 mg, 2.0 μmol) to *N*,*N*-dimethylformamide solution of pdCpA *tetra*-*n*-butylammonium salt (10 μL, 0.44 μmol) in a microtube, and the mixture was incubated at 30°C for 2 h. The reaction was monitored by a reverse-phase HPLC (Inertsil ODS-3 150–4.6, 20–100% CH_3OH/0.1 M NH_4OAc, pH 4.5, over 20 min at flow rate of 1.0 mL min^{-1}, detection at 260 nm), the products had the retention time of 14.8 and 15.2 min, for the two positional (2′,3′) isomers. After 2 h, Et_2O (1 mL) was added to the solution and then the precipitation was collected by a centrifugation (20000 g, 4°C, 5 min). The precipitate was dissolved in acetonitrile (20 μL) and the solution was subjected to the re-precipitation by adding Et_2O (1 mL). The crude product was again dissolved in acetonitrile (0.2 mL) and 0.1 M NH_4OAc, pH 4.5 (0.8 mL), and purified by a reverse-phase HPLC (Inertsil ODS-3 50–20, 20% CH_3OH/0.1 M NH_4OAc, pH 4.5, 2 min then 20–100% CH_3OH/0.1 M NH_4OAc, pH 4.5, over 20 min at flow rate of 10 mL min^{-1}, detection at 260 nm), the products with the retention time of 14.2 and 14.6 min were obtained as two positional (2′,3′) isomers. Compound **5** (0.27 μmol, 62%). HR-ESI MS: *m*/*z*: calcd for $C_{39}H_{51}N_{11}O_{22}P_2$: 1087.2685; found: 1086.2632 $[M-H]^-$.

2-(Pent-4-enoyl)amino-4-[*O*-(*N*-phthalimidyl)]hydroxybutanoic acid *tert*-butyl ester: *N*-Hydroxyphthalimide (0.32 g, 2.0 mmol), triphenylphosphine (0.52 g, 2.0 mmol) and diethyl azodicarboxylate (0.31 mL, 2.0 mmol) were added to a solution of 2-(pent-4-enoyl)amino-4-hydroxy-

butanoic acid *tert*-butyl ester (0.42 g, 1.64 mmol) in tetrahydrofuran (10 mL). The reaction mixture was stirred for 2 h at RT under nitrogen atmosphere. The solvent was evaporated, and the residue was purified by flash column chromatography (toluene/ethyl acetate 5:1) to give product (0.56 g, 84%). $R_f = 0.57$ (toluene/ethyl acetate 1:1); ^{1}H NMR (600 MHz, $CDCl_3$): δ = 7.86–7.77 (m, 4H, aromatic), 7.13 (d, $J = 7.13$ Hz, 1H, NH), 5.88–5.84 (m, 1H, C=CH-), 5.10–4.98 (m, 2H, -C=CH_2), 4.73–4.70 (m, 1H, α-H), 4.31–4.29 (m, 2H, γ-H_2), 2.46–2.41 (m, 4H, C=C-CH_2-, C=C-C-CH_2-), 2.33–2.28 (m, 2H, β-H_2), 1.46 (s, 9H, $(CH_3)_3$-C-); ^{13}C NMR (120 MHz, $CDCl_3$): δ = 172.31, 170.64, 163.53 (3C, 3×C=O), 137.16 (C=*C*-), 134.66, 128.81, 123.68 (3C, 3×aromatic), 115.34 (*C*=C-), 82.18 ($(CH_3)_3$-*C*-), 75.30 (γ-C), 50.35 (α-C), 35.61, 29.46 (2C, C=*C*-*C*-C), 29.85 (β-C), 27.91 ($(CH_3)_3$-C-); HR-FAB MS: *m*/*z*: calcd for $C_{21}H_{26}N_2O_6$: 402.1791; found: 403.1879 $[M+H]^+$.

2-(Pent-4-enoyl)amino-4-(*O*-amino)hydroxybutanoic acid *tert*-butyl ester: 2-(Pent-4-enoyl)amino-4-[*O*-(*N*-phthalimidyl)]hydroxybutanoic acid *tert*-butyl ester (0.16 g, 0.39 mmol) was dissolved in 40% methylamine/methanol solution (2 mL). The mixture was stirred for 45 min at RT, and evaporated. The residue was purified by flash column chromatography (toluene/ethyl acetate 1:5) to give the target compound (0.94 g, 89%). $R_f = 0.51$ (ethyl acetate); ^{1}H NMR (600 MHz, $CDCl_3$): δ = 6.31 (d, J = 7.19 Hz, 1H, NH), 5.88–5.80 (m, 1H, C=CH-), 5.44 (s, 2H, O-NH_2), 5.10–5.00 (m, 2H, -C=CH_2), 4.60–4.56 (m, 1H, α-H), 3.76–3.71 (m, 2H, γ-H_2), 2.43–2.31 (m, 4H, C=C-CH_2-, C=C-C-CH_2-), 2.11–1.96 (m, 2H, β-H_2), 1.48 (s, 9H, $(CH_3)_3$-C-); ^{13}C NMR (120 MHz, $CDCl_3$): δ = 171.54, 171.45 (2C, 2×C=O), 136.94 (C=*C*-), 115.56 (*C*=C-), 82.12 ($(CH_3)_3$-*C*-), 72.00 (γ-C), 50.42 (α-C), 35.74, 29.45 (2C, C=*C*-*C*-C), 31.19 (β-C), 27.97 ($(CH_3)_3$-C-); HR-FAB MS: *m*/*z*: calcd for $C_{13}H_{24}N_2O_4$: 272.1736; found: 273.1814 $[M+H]^+$.

2-(Pent-4-enoyl)amino-4-[*O*-(*N*-6-nitroveratryloxycarbonyl)amino]hydroxybutanoic acid cyanomethyl ester (4): 2-(Pent-4-enoyl)amino-4-(*O*-amino)hydroxybutanoic acid *tert*-butyl ester (49 mg, 0.18 mmol) was dissolved in trifluoroacetic acid (1 mL) and the solution was stirred for 1 h. The reaction mixture was concentrated, then the residue was dissolved in H_2O (500 μL) and dioxane (1.0 mL). To this solution was added triethylamine (75 μL, 0.54 mmol) and 6-nitroveratryloxycarbonyl chloride (60 mg, 0.22 μmol). The mixture was stirred at 0°C for 1.5 h in the dark, then extracted with ethyl acetate. The organic layer was washed with 5% citric acid and brine, dried using $MgSO_4$, filtrated, concentrated, and purified by flash column chromatography (chloroform/methanol 25:1). The concentrated residue was dissolved in acetonitrile (0.7 mL) and cooled at 0°C under nitrogen atmosphere. To the solution was added triethylamine (27 μL, 0.20 mmol) and chloroacetonitrile (41 μL, 0.65 mmol) and the mixture was stirred at room temperature for 20 h. The mixture was extracted with ethyl acetate and the extract was washed with brine, dried using $MgSO_4$, filtrated, concentrated, and purified by flash column chromatography (hexane/ethyl acetate 1:5) to give compound **4** (10 mg, 11%). $R_f = 0.50$ (hexane/ethyl acetate 1:5); ^{1}H NMR (600 MHz, $CDCl_3$): δ = 7.75 (s, 1H, NH), 7.73 (1H, aromatic), 6.98 (s, 1H, aromatic), 5.87–5.80 (m, 1H, C=CH-), 5.63–5.57 (m, 2H, O-CH_2-Ph), 5.09–4.99 (m, 2H, -C=CH_2), 4.89–4.86 (m, 1H, α-H), 4.80–4.72 (m, 2H, O-CH_2-CN), 4.05–3.99 (m, 2H, γ-H_2), 3.99, 3.97 (all s, 6H, 2×OCH_3), 2.44–2.37 (m, 4H, C=C-CH_2-, C=C-C-CH_2-), 2.33–2.28 (m, 1H, β-H_2-a), 2.14–2.09 (m, 1H, β-H_2-b); ^{13}C NMR (120 MHz, $CDCl_3$): δ = 1712.81, 170.79, 157.68 (3C, 3×C=O), 153.62, 148.57, 140.03, 126.33, 110.67, 108.31 (6C, aromatic), 136.97 (C=*C*-C), 115.52 (*C*=C-C), 114.10 (-CN), 72.94 (γ-C), 64.82 (O-*C*-Ph), 56.57, 56.46 (2C, 2×OCH_3), 49.50 (α-C), 48.86 (O-*C*-CN), 35.34, 29.38 (2C, C=*C*-*C*-C), 29.19 (β-C); HR-FAB MS: *m*/*z*: calcd for $C_{21}H_{26}N_4O_{10}$: 494.1649; found: 495.1730 $[M+H]^+$.

2-(Pent-4-enoyl)amino-4-[*O*-(*N*-6-nitroveratryloxycarbonyl)amino]hydroxybutanoic acid pdCpA ester (6): The aminoacylation of pdCpA was carried out by adding the compound **4** (0.80 mg, 1.6 μmol) to *N*,*N*-dimethylformamide solution of pdCpA *tetra*-*n*-butylammonium salt (10 μL, 0.44 μmol) in a microtube, and the reaction mixture was incubated at 30°C for 2 h. The reaction was monitored by a reverse-phase HPLC (Inertsil ODS-3 150-4.6, 20–100% CH_3OH/0.1 M NH_4OAc, pH 4.5, over 20 min at flow rate of 1.0 mL min^{-1}, detection at 260 nm), the products with the retention time of 14.2 and 14.5 min were collected as the two positional (2′,3′) isomers. After 2 h, Et_2O (1 mL) was added to the solution and then the precipitation was collected by a centrifugation (20000 g, 4°C, 5 min). The precipitate was dissolved in acetonitrile (20 μL) and the solution was subjected to the re-precipitation by using Et_2O (1 mL) as poor solvent. The crude material was purified by a reverse-phase HPLC to afford compound **6** (0.17 μmol, 39%). HR-ESI MS: *m*/*z*: calcd for $C_{38}H_{49}N_{11}O_{22}P_2$: 1073.2529; found: 1072.2435 $[M-H]^-$.

Synthesis of aminoacyl tRNA: A mixture containing $tRNA_{CCCG}$(-CA) (0.3 nmol), aminoacyl pdCpA (6.0 nmol), 1 mM ATP, 3.3 mM DDT, 15 mM $MgCl_2$, 20 μg mL^{-1} BSA, 15% DMSO and T4 RNA ligase (40 U) in a 20 μL of 55 mM HEPES-Na (pH 7.5) was incubated at 37°C for 20 min. The mixture was treated with 3 M potassium acetate (pH 5.5, 0.3 M potassium acetate as a final concentration) and extracted with phenol/chloroform/isoamylalcohol and chloroform. The solution was irradiated (365 nm) at 0°C for 90 min and then reacted with $^1/_{10}$ volume of 50 mM I_2 (H_2O/THF 1:1) for 20 min at 0°C. The reaction mixture was treated with 3 vol. ethanol to precipitate the aminoacyl-tRNA. The pellet was washed with 70% ethanol, dried, and dissolved in 1 mM potassium acetate (pH 5.5, 3 μL), and the solution was immediately added to the reaction mixture of the in vitro translation system.

In vitro protein biosynthesis: In vitro translation/transcription were performed by using *E. coli* T7 S30 extract systems for circular DNA kit (Promega). For the expression analysis, the reactions were carried out in a 10 μL of a reaction mixture containing plasmid (0.5 μg), 0.1 mM each amino acids except arginine, 0.01 mM arginine, S30 premix without amino acid (4 μL), S30 extract (3 μL). If necessary, 0.01 mM aminoacyl $tRNA_{CCCG}$ was also added. The mixture was incubated at 37°C for 60 min. For the large-scale expression, the reaction volumes were increased up to 200 μL.

Western blotting analysis: In vitro translation mixture (0.25 μL) was applied to 15% polyacrylamide gels followed by SDS-PAGE. All samples were heated prior to application to SDS-PAGE wells. After electroblotting to a PVDF membrane, the membrane was incubated with 3% BSA in TBST (25 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.02% Tween20), T7-tag antibody and anti-mouse IgG (H+L) AP conjugate at 37°C for 30 min. The development with Western Blue was performed at 37°C for 20 min. The efficiency of incorporation of unnatural amino acids was estimated by comparing the band intensity of the full-length polypeptide with its wild-type expressed in vitro translation. The band intensity was evaluated using the Scion Image programs (Scion Corporation).

Silver staining analysis: Each sample was applied to 15% polyacrylamide gels followed by SDS-PAGE. All samples were heated prior to application to SDS-PAGE wells. Silver staining was carried out by using Silver Staining II Kit Wako (Wako Pure Chemical Industries, Ltd.) according to the manufacturer's instructions. Developing time was 5 min.

Purification of proteins by T7-tag affinity chromatography: The volume of the reaction mixture (200 μL) of in vitro translation was adjusted to 0.5 mL with phosphate buffer (4.3 mM Na_2HPO_4, 1.5 mM KH_2PO_4, 2.7 mM KCl, 137 mM NaCl, 0.1% Tween20, 0.002% sodium aside, pH 7.3). The mixture was added to a suspension of T7 tag antibody agarose (50 μL), then the mixture was incubated at RT for 60 min. The agarose gel containing proteins was loaded to an empty spin column, washed with the phosphate buffer (5 mL), and eluted with 100 mM citric acid (pH 2.2, 0.12 mL). The elution was neutralized with 2 M Tris (pH 10.4, 180 μL) and concentrated using Microcon YM-10 (millipore).

Purification of proteins by His-tag affinity chromatography: To a solution containing proteins in 100 mM HEPES-Na (pH 7.5, 50 μL) was added a suspension of MagneHis Ni particles (5 μL). The mixture was incubated at RT for 30 min and washed thoroughly with 100 mM HEPES-Na (pH 7.5, 50 μL). Finally, the target proteins were obtained by desorbing with elution buffer (100 mM HEPES-Na, (pH 7.5) containing 750 mM imidazole, 10 μL).

Glycoblotting by an engineered protein of unmodified optional carbohydrates: Proteins purified by T7 tag affinity chromatography (approx. 200 ng) was dissolved in acetate buffer (pH 4.0, 3 μL) containing carbohydrate to be introduced. The mixture was incubated at 37°C and subjected to the purification by His-tag affinity chromatography and the

product was identified and characterized by MALDI-TOF mass spectrometry.

Sugar elongation (sialylation) reaction of Hser(ONHMe) mutant of streptavidin carrying a lactose residue: Hser(ONHMe) mutant of streptavidin purified by T7 tag affinity chromatography (approx. 200 ng) was dissolved in acetate buffer (pH 4.0, 10 μL) containing 400 mM lactose. The solution was incubated at 37 °C for 44 h. To a solution diluted with 100 mM HEPES-Na (pH 7.5, 20 μL) was added a suspension of MagneHis Ni particles (2 μL). The mixture was incubated at RT for 30 min, washed twice with 100 mM HEPES-Na (pH 7.5, 100 μL) and 50 mM HEPES-Na (pH 7.5, 100 μL), respectively. The magnetic beads bearing lactosylated streptavidin were directly re-suspended by adding to 100 μL 50 mM HEPES-Na (pH 7.5) containing 3 mM CMP-Neu5Ac, 0.01 % BSA, 0.2 % Triton X-100, and 300 mU mL^{-1} α2,3-sialyltransferase. The mixture was incubated at 37 °C for 22 h, then washed twice with 50 mM HEPES-Na (pH 7.5, 100 μL) and 100 mM HEPES-Na (pH 7.5, 100 μL), respectively. The product was obtained by desorbing with elution buffer (100 mM HEPES-Na (pH 7.5) containing 750 mM imidazole, 6 μL) and analyzed by MALDI-TOF mass spectrometry.

Acknowledgements

We thank Dr. M. Kurogochi for his help in measuring MALDI-TOF MS. This work was supported partly by a grant for "Research Project of the Functional Glycomaterials" from the Ministry of Education, Culture, Science, and Technology, Japan.

[1] R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683–720.
[2] S. Elliott, *Nat. Biotechnol.* **2003**, *21*, 414–420.
[3] Z. Zhang, *Science* **2004**, *303*, 371–373.
[4] R. Xu, *J. Am. Chem. Soc.* **2004**, *126*, 15654–15655.
[5] T. Ohta, *Angew. Chem.* **2003**, *115*, 5344–5347; *Angew. Chem. Int. Ed.* **2003**, *42*, 5186–5189.
[6] M. Sato, *J. Am. Chem. Soc.* **2004**, *126*, 14013–14022.
[7] H. Liu, *J. Am. Chem. Soc.* **2003**, *125*, 1702–1703.
[8] E. C. Rodriguez, L. A. Marcaurelle, C. R. Bertozzi, *J. Org. Chem.* **1998**, *63*, 7134–7135.
[9] T. Hohsaka, M. Sisido, *Curr. Opin. Chem. Biol.* **2002**, *6*, 809–815.
[10] T. Hohsaka, *Biochemistry* **2001**, *40*, 11060–11064.
[11] S.-I. Nishimura, *Angew. Chem.* **2005**, *117*, 93–98; *Angew. Chem. Int. Ed.* **2005**, *44*, 91–96.
[12] J. C. Anderson, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7566–7571.
[13] S. E. Cervigni, P. Dumy, M. Mutter, *Angew. Chem.* **996**, *108*, 1325–1328; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1230–1232.
[14] M. R. Carrasco, R. T. Brown, *J. Org. Chem.* **2003**, *68*, 8853–8858.
[15] K. Niikura, R. Kamitani, M. Kuogochi, R. Uematsu, Y. Shinohara, H. Nakagawa, K. Degushi, K. Monde, H. Kondo, S.-I. Nishimura, *Chem. Eur. J.* **2005**, *11*, 3825–3834.
[16] M. Lodder, S. Golovine, S. M. Hecht, *J. Org. Chem.* **1997**, *62*, 778–779.
[17] V. A. Karginov, *J. Am. Chem. Soc.* **1997**, *119*, 8166–8176.

Received: May 13, 2005
Published online: September 6, 2005

