J|A|C|S
ARTICLES
Published on Web 03/06/2007

Site-Specific Incorporation of Glycosylated Serine and Tyrosine Derivatives into Proteins

Nour Eddine Fahmi, Larisa Dedkova, Bixun Wang, Serguei Golovine, and Sidney M. Hecht*

Contribution from the Departments of Chemistry and Biology, University of Virginia, Charlottesville, Virginia 22901

Received October 18, 2006; E-mail: sidhecht@virginia.edu

Abstract: Glycosylation of proteins can have a dramatic effect on their physical, chemical, and biological properties. Analogues of dihydrofolate reductase and firefly luciferase containing glycosylated amino acids at single, predetermined sites have been elaborated. Misacylated suppressor tRNAs activated with glycosylated serine and tyrosine derivatives were used for suppression of the nonsense codons in a cell-free protein biosynthesizing system, thereby permitting the preparation of the desired glycosylated proteins. In this fashion, it was possible to obtain proteins containing both mono- and diglycosylated amino acids, including glycosylated serine and tyrosine moieties. For the modified firefly luciferases, the effect of these substitutions on the wavelength of the light emitted by firefly luciferase was investigated. The maximum wavelength for mutants containing peracetylated glycosylated serine derivatives at position 284 showed a red shift in the emission spectra. For mutants containing glycosylated tyrosines, the red shift was observed only when the carbohydrate moiety was fully deacetylated.

Introduction

The post-translational modification of proteins by the attachment of carbohydrates to individual amino acids constitutes one of the most important structural modifications of proteins. In the majority of natural glycoproteins, the sugar residues are covalently attached to the protein backbone through either an N-linked asparagine residue or an O-linked serine or threonine.[1] Although less common, *O*-glycosylation of a tyrosine residue and *C*-glycosylation of a tryptophan residue have also been reported.[2,3] In recent years, it has become evident that protein-linked oligosaccharides can have critical effects on the structure and biological function of the glycoproteins.[4] For instance, glycosylated proteins are involved in cell−cell recognition events[5] and can have greater resistance toward thermolysis, proteolysis, and enzymatic degradation.[6,7] Glycosylation can also affect peptide backbone conformation and folding[8] and the catalytic activity of the enzyme.[9]

Investigation of the structure and function of individual glycoproteins has been complicated by the fact that naturally occurring glycoproteins exist as a number of glycoforms, all of which share the same backbone but differ either in the structure or position of the attached carbohydrates.[1a] Only in a few studies have the isolation of pure glycoforms been successful from natural glycoproteins, and then only by extensive chromato-

(1) (a) Rademacher, T. W.; Parekh, R. B.; Dwek, R. A. *Annu. Rev. Biochem.* **1988**, *57*, 785. (b) Kobata, A. *Eur. J. Biochem.* **1992**, *209*, 483. (c) Lis, H.; Sharon, N. *Eur. J. Biochem.* **1993**, *218*, 1. (d) Bill, R. M.; Flitsch, S. L. *Chem. Biol.* **1996**, *3*, 145. (e) Bill, R. M.; Revers, L.; Wilson, I. B. H. In *Protein Glycosylation*; Kluwer Academic: Boston, 1998. (f) Sears, P.; Wong, C.-H. *Cell. Mol. Life Sci.* **1998**, *54*, 223.

(2) (a) Chen, P. S.; Mitchell, H. K.; Neuweg, M. *Insect Biochem.* **1978**, *8*, 279. (b) Rodriguez, I. R.; Whelan, W. J. *Biochem. Biophys. Res. Commun.* **1985**, *132*, 829.

(3) (a) Löffler, A.; Doucey, M.-A.; Jansson, A. M.; Müller, D. R.; de Beer, T.; Hess, D.; Meldal, M.; Richter, W. J.; Vlieghenthart, J. F. G.; Hofsteenge, J. *Biochemistry* **1996**, *35*, 12005. (b) Doucey, M.-A.; Hess, D.; Blommers, M. J. J.; Hofsteenge, J. *Glycobiology* **1999**, *9*, 435.

(4) For reviews, see: (a) Varki, A. *Glycobiology* **1993**, *3*, 97. (b) Wyss, D. F.; Wagner, G. *Curr. Opin. Biotechnol.* **1996**, *7*, 409. (c) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683. (d) Seitz, O. *ChemBioChem.* **2000**, *1*, 214.

(5) (a) Kojima, N.; Shiota, M.; Sadahira, Y.; Kazuko, H.; Hakomori, S. *J. Biol. Chem.* **1992**, *267*, 17264. (b) Sharon, N.; Lis, H. *Sci. Am.* **1993**, 82. (c) Lasky, L. A. *Annu. Rev. Biochem.* **1995**, *64*, 113. (d) Snell, W. J.; White, J. M. *Cell* **1996**, *85*, 629. (e) Pratt, M. R.; Bertozzi, C. R. *Chem. Soc. Rev.* **2005**, *34*, 58.

(6) (a) Gibson, R.; Schlesinger, S.; Kornfeld, S. *J. Biol. Chem.* **1979**, *254*, 3600. (b) Opdenakker, G.; Rudd, P. M.; Ponting, C. P.; Dwek, R. A. *FASEB J.* **1993**, *7*, 1330. (c) Wang, C.; Eufemi, M.; Turano, C.; Giartosio, A. *Biochemistry* **1996**, *35*, 7299. (d) Sola, R. J.; Griebenow, K. *FEBS Lett.* **2006**, *580*, 1685.

(7) (a) Bernard, B. A.; Yamada, K. M.; Olden, K. *J. Biol. Chem.* **1982**, *257*, 8549. (b) Bernard, B. A.; Newton, S. A.; Olden, K. *J. Biol. Chem.* **1983**, *258*, 12198. (c) Hill, T. G.; Wang, P.; Huston, M. E.; Wartchow, C. A.; Oehler, L. M.; Smith, M. B.; Bednarski, M. D.; Callstrom, M. R. *Tetrahedron Lett.* **1991**, *32*, 6823. (d) Wang, P.; Hill, T. G.; Wartchow, C. A.; Huston, M. E.; Oehler, L. M.; Smith, M. B.; Bednarski, M. D.; Callstrom, M. R. *J. Am. Chem. Soc.* **1992**, *114*, 378. (e) Sola, R. J.; Al-Azzam, W.; Griebenow, K. *Biotechnol. Bioeng.* **2006**, *94*, 1072.

(8) (a) Gerken, T. A.; Butenhof, K. J.; Shogren, R. *Biochemistry* **1989**, *28*, 5536. (b) Joao, H. C.; Scragg, I. G.; Dwek, R. A. *FEBS Lett.* **1992**, *307*, 343. (c) Andreotti, A. H.; Kahne, D. *J. Am. Chem. Soc.* **1993**, *115*, 3352. (d) Davis, J. T.; Hirani, S.; Bartlett, C.; Reid, B. R. *J. Biol. Chem.* **1994**, *269*, 3331. (e) Liang, R.; Andreotti, A. H.; Kahne, D. *J. Am. Chem. Soc.* **1995**, *117*, 10395. (f) Imperiali, B.; Rickert, K. W. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 97. (g) Live, D. H.; Kumar, R. A.; Beebe, X.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12759. (h) Huang, X.; Smith, M. C.; Berzofski, J. A.; Barchi, J. J., Jr. *FEBS Lett.* **1996**, *393*, 280. (i) Tagashira, M.; Toma, K. *Trends Glycosci. Glycotechnol.* **2005**, *17*, 145.

(9) (a) Semenkovich, C. F.; Luo, C. C.; Nakanishi, M. K.; Chen, S. H.; Smith, L. C.; Chan, L. *J. Biol. Chem.* **1990**, *265*, 5429. (b) Levashov, A. V.; Rariy, R. V.; Martinek, K.; Klyachko, N. L. *FEBS Lett.* **1993**, *336*, 385. (c) Longo, M. A.; Combes, D. *FEBS Lett.* **1995**, *375*, 63. (d) Kohen, A.; Jonsson, T.; Klinman, J. P. *Biochemistry* **1997**, *36*, 2603. (e) Longo, M. A.; Combes, D. *J. Mol. Catal. B: Enzymology* **1997**, *2*, 281. (f) Yáñez, E.; Carmona, T. A.; Tiemblo, M.; Jiménez, A.; Fernández-Lobato, M. *Biochem. J.* **1998**, *329*, 65. (g) Wu, Y. I.; Munshi, H. G.; Sen, R.; Snipas, S. J.; Salvesen, G. S.; Fridman, R.; Stack, M. S. *J. Biol. Chem.* **2004**, *279*, 8278. (h) Srimathi, S.; Jayaraman, G. *Protein J.* **2005**, *24*, 79.

10.1021/ja067466n CCC: $37.00

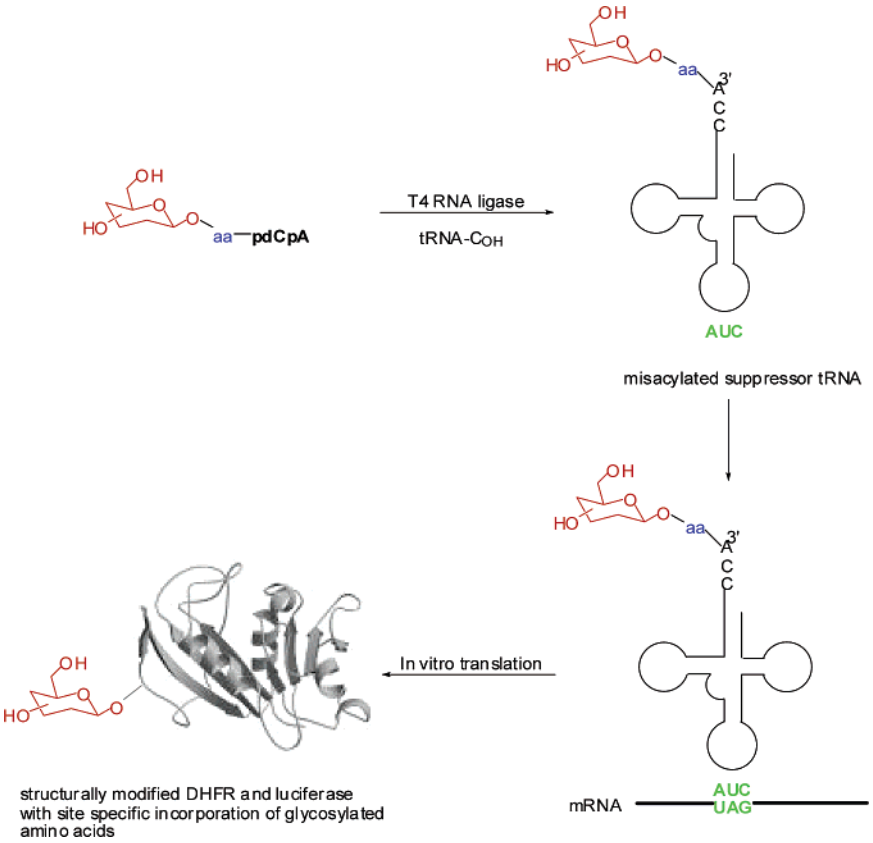


Figure 1. Preparation of glycoproteins by *in vitro* suppression of a nonsense codon with a glycosylated aminoacyl-tRNA$_{CUA}$. This is effected by enzymatic ligation of a glycosylated aminoacyldinucleotide (pdCpA) to an abbreviated tRNA lacking that dinucleotide.

graphic separations.[10] In order to study, and potentially modify, the biological and physical processes mediated by the carbohydrate constituents of glycoproteins, it is imperative to develop methods to access glycoproteins with uniform and well-defined carbohydrate constituents.

The introduction of noncoded amino acids into proteins at predetermined sites using nonsense codon suppression by misacylated tRNAs is a powerful technique that can afford important insights into protein structure and function.[11] Although glycosylation of proteins is a post-translational process in nature, recent experiments have shown that it is possible to obtain glycoproteins with monosaccharides attached at specific sites using nonsense codon suppression (Figure 1).[12] We previously reported a general strategy for the synthesis of *O*-glycosylated seryl-pdCpAs and their ligation to truncated tRNAs,[13] and an alternative strategy has recently been described.[14] Further, the elaboration of modified glycoproteins in intact bacterial cells has also been described.[15] In the present report, we describe the extension of this methodology to the synthesis of *O*-glycosylated tyrosyl-pdCpAs and the corresponding misacylated tRNAs. By the use of an *in vitro* translation system, the molecules obtained (**1**−**12**) (Figure 2) were incorporated site specifically into two different proteins, namely dihydrofolate reductase (DHFR) and firefly luciferase.

The present results include the first report of the incorporation of a glycosylated tyrosine into a protein by *in vitro* protein synthesis. The effect of this modification on the properties of luciferase, especially the wavelength of the emitted light, is of particular interest due to the numerous biochemical and structural biology studies of this protein in recent years. The

(10) Rudd, P. M.; Joao, H. C.; Coghill, E.; Fiten, P.; Saunders, M. R.; Opdenakker, G.; Dwek, R. A. *Biochemistry* **1994**, *33*, 17.

(11) (a) Pezzuto, J. M.; Hecht, S. M. *J. Biol. Chem.* **1980**, *255*, 865. (b) Heckler, T. G.; Chang, L. -H.; Zama, Y.; Naka, T.; Hecht, S. M. *Tetrahedron* **1984**, *40*, 87. (c) Heckler, T. G.; Chang, L. -H.; Zama, Y.; Naka, T.; Chorghade, M. S.; Hecht, S. M. *Biochemistry* **1984**, *23*, 1468. (d) Roesser, J. R.; Chorghade, M. S.; Hecht, S. M. *Biochemistry* **1986**, *25*, 6361. (e) Payne, R. C.; Nichols, B. P.; Hecht, S. M. *Biochemistry* **1987**, *26*, 3197. (f) Heckler, T. G.; Roesser, J. R.; Cheng, X.; Chang, P. -I.; Hecht, S. M. *Biochemistry* **1988**, *27*, 7254. (g) Baldini, G.; Martoglio, B.; Schachenmann, A.; Zugliani, C.; Brunner, J. *Biochemistry* **1988**, *27*, 7951. (h) Roesser, J. R.; Xu, C.; Payne, R. C.; Surratt, C. K.; Hecht, S. M. *Biochemistry* **1989**, *28*, 5185. (i) Robertson, S. A.; Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Nucleic Acids Res.* **1989**, *17*, 9649. (j) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182. (k) Noren, C. J.; Anthony-Cahill, S. J.; Suich, D. J.; Noren, K. A.; Griffith, M. C.; Schultz, P. G. *Nucleic Acids Res.* **1990**, *18*, 83. (l) Mendel, D.; Ellman, J. A.; Schultz, P. G. *J. Am. Chem. Soc.* **1991**, *113*, 2758. (m) Hecht, S. M. *Acc. Chem. Res.* **1992**, *25*, 545. (n) Ellman, J. A.; Mendel, D.; Schultz, P. G. *Science* **1992**, *255*, 197. (o) Mendel, D.; Ellman, J. A.; Chang, Z.; Veenstra, D. L.; Kollman, P. A.; Schultz, P. G. *Science* **1992**, *256*, 1798. (p) Steward, L. E.; Collins, C. S.; Gilmore, M. A.; Carlson, J. E.; Ross, J. B. A.; Chamberlin, A. R. *J. Am. Chem. Soc.* **1997**, *119*, 6. (q) England, P. M.; Lester, H. A.; Davidson, N.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11025. (r) Murakami, H.; Hohsaka, T.; Ashizuka, Y.; Sisido, M. *J. Am. Chem. Soc.* **1998**, *120*, 7520. (s) Hohsaka, T.; Ashizuka, Y.; Taira, H.; Murakami, H.; Sisido, M. *Biochemistry* **2001**, *40*, 11060. (t) Lu, T.; Ting, A. Y.; Mainland, J.; Jan, I. Y.; Schultz, P. G.; Yang, J. *Nat. Neurosci.* **2001**, *4*, 239. (u) Wang, B.; Brown, K. C.; Lodder, M.; Craik, C. S.; Hecht, S. M. *Biochemistry* **2002**, *41*, 2805.

(12) (a) Mamaev, S. V.; Laikhter, A. L.; Arslan, T.; Hecht, S. M. *J. Am. Chem. Soc.* **1996**, *118*, 7243. (b) Arslan, T.; Mamaev, S. V.; Mamaeva, N. V.; Hecht, S. M. *J. Am. Chem. Soc.* **1997**, *119*, 10877. (c) Schmidt, R. R.; Castro-Palomino, J. C.; Retz, O. *Pure Appl. Chem.* **1999**, *71*, 729.

(13) Fahmi, N.; Golovine, S.; Wang, B.; Hecht, S. M. *Carbohydr. Res.* **2001**, *330*, 149.

(14) Roehrig, C. H.; Retz, O. A.; Hareng, L.; Hartung, T.; Schmidt, R. R. *ChemBioChem.* **2005**, *6*, 1805.

(15) (a) Xu, R.; Hanson, S. R.; Zhang, Z.; Yang, Y. -Y.; Schultz, P. G.; Wong, C.-H. *J. Am. Chem. Soc.* **2004**, *126*, 15654. (b) Hang, Z.; Gildersleeve, J.; Yang, Y.-Y.; Xu, R.; Loo, J. A.; Urya, S.; Wong, C.-H.; Schultz, P. G. *Science* **2004**, *303*, 371. (c) Hanson, S. R.; Xu, R.; Hwang, D. R.; Zhang, Z. W.; Yang, Y. Y.; Galan, C.; Schultz, P. G.; Wong, C. H. *Glycobiology* **2004**, *14*, 1090.

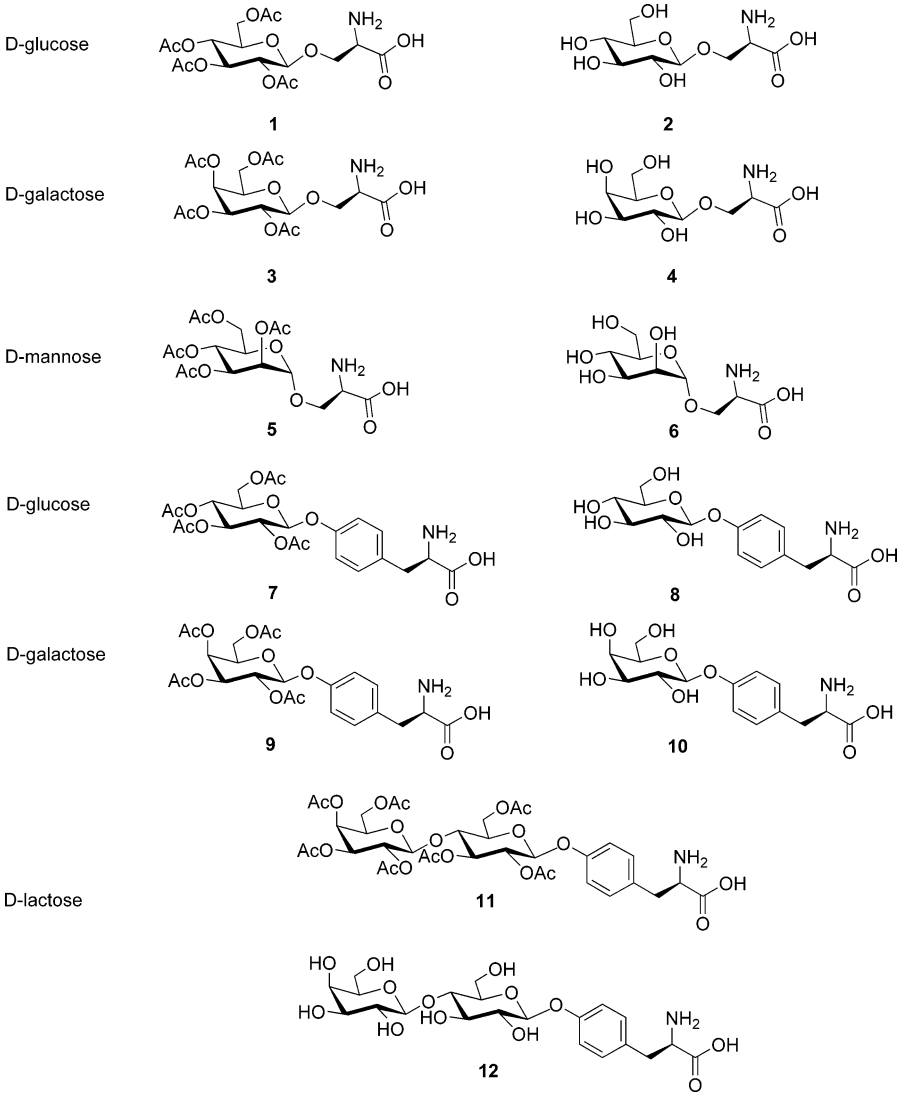
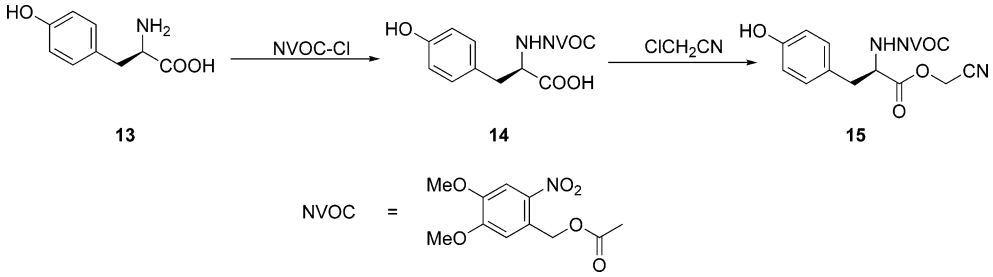


Figure 2. Glycosylated amino acids utilized in this study.

Scheme 1



sugar residues in the elaborated proteins can also be used as starting points for ligation to other saccharides using a combination of enzymatic or chemical techniques to provide more complex glycoproteins.

Also enabled by the present results is the conjugation of the glycosylated proteins at specific, predetermined sites to solid supports containing suitable functional groups such as hydrazine or boronic acid moieties.

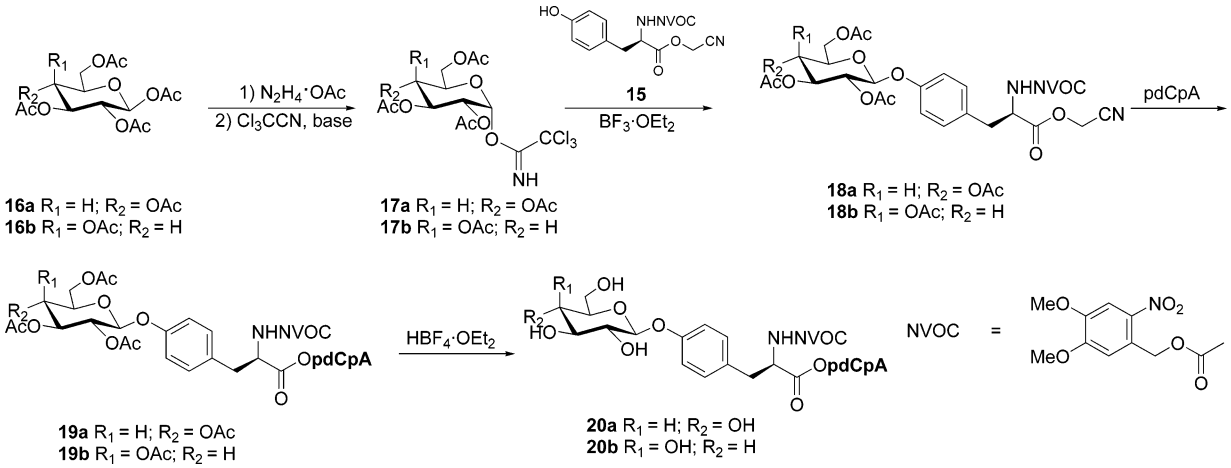
Results

To enable the preparation of the requisite glycosylated aminoacyl-tRNAs, glycosylated seryl- and tyrosyl-pdCpA esters were prepared initially. The tyrosine derivatives had not been synthesized previously, and their synthesis is described below.

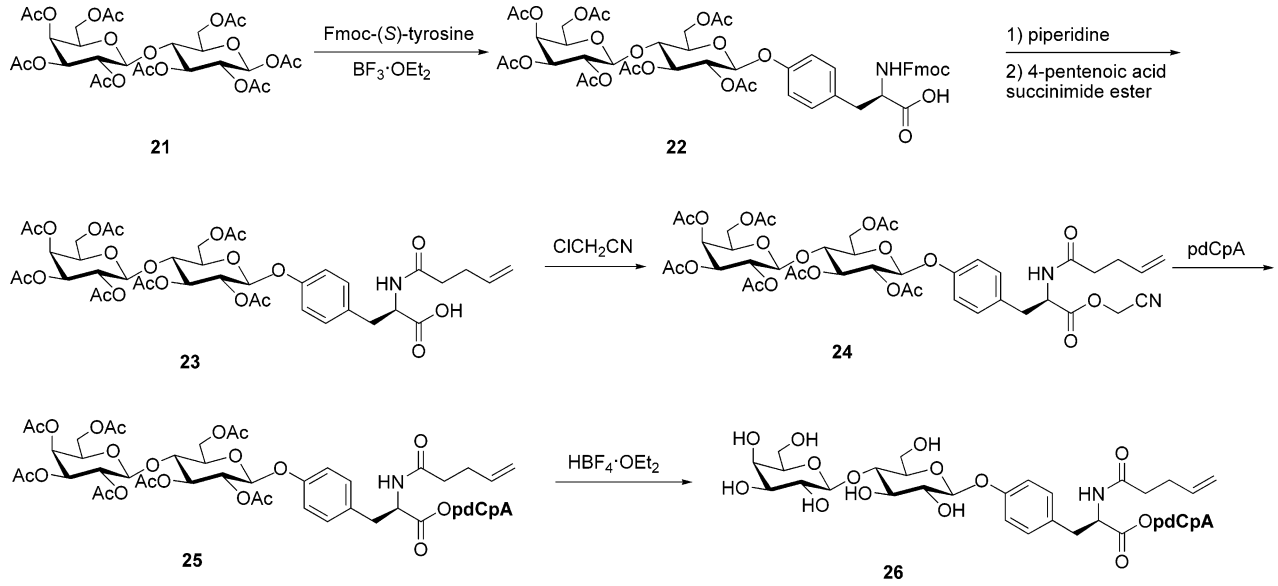
Preparation of Glycosylated Tyrosyl-pdCpA Esters. NVOC-(*S*)-tyrosine cyanomethyl ester **(15)** was prepared in two steps from (*S*)-tyrosine **(13)** as reported previously (Scheme 1).[12b] The amino group was protected with 2-nitroveratryl chloroformate (NVOC-Cl) in the presence of sodium bicarbonate, giving **14** in 57% yield. Conversion to the respective cyanomethyl ester was carried out in 54% yield using chloroacetonitrile and triethylamine.

Although several methods have been reported for the glycosylation of the tyrosine side chain using different glycosyl donors

Scheme 2



Scheme 3



and catalysts, the application of these procedures in the present situation was problematic owing to poor yields. The route actually employed for the synthesis of glycosylated tyrosines **18a** and **18b** is outlined in Scheme 2. Glycosyl donor **17a** was prepared from β-D-glucose peracetate by selective removal of the anomeric ester using hydrazine acetate in dimethylformamide (DMF),[16] followed by treatment with trichloroacetonitrile in the presence of potassium carbonate. After 48 h, the trichloroacetimidate **17a** was obtained as a single anomer (α anomer) in 67% yield.[17] Attempts to selectively deacetylate the anomeric hydroxyl using benzylamine in tetrahydrofuran (THF)[18] failed to give the desired product in pure form.

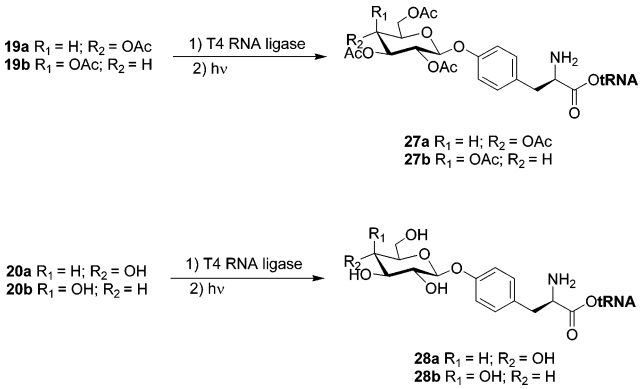
Galactose trichloroacetimidate **17b** was prepared using a published procedure.[19] β-D-Galactose peracetate was selectively deacetylated at the anomeric position using hydrazine acetate in DMF in quantitative yield as judged by thin-layer chromatography (TLC). Treatment of the crude product with trichloroacetonitrile using sodium as a base afforded the α- and β-trichloroacetimidates **17b** in nearly 4:1 ratio; these were easily separated by column chromatography.

For the glycosylation of tyrosine derivative **15**, only the α-trichloroacetimidates were used as glycosyl donors. The coupling reaction between **15** and the imidates **17a** and **17b** was accomplished using BF_3·OEt_2 as promoter in dichloromethane. Glycosylated tyrosines **18a** and **18b** were obtained as pure β anomers in 77 and 70% yields, respectively (Scheme 2). The formation of 1,2-trans isomers undoubtedly resulted from neighboring group participation of the C-2 acetyl donor functionality.

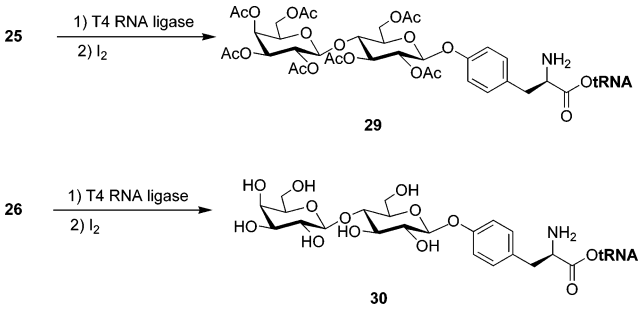
For the glycosylation of tyrosine with lactose, a route similar to the one reported previously for serine has been followed (Scheme 3).[13,20] β-D-Lactose peracetate **(21)** was treated with unprotected Fmoc-(*S*)-tyrosine in the presence of BF_3·OEt_2 to give glycosylated tyrosine **22** in 35% yield. The reaction required a large excess of BF_3·OEt_2 and proceeded slowly compared to glycosylation of serine. *N*-(9-Fluorenylmethoxycarbonyl)-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-tyrosine **(22)** was isolated by careful chromatography, and the product was slightly

(16) (a) Excoffier, G.; Gagnaire, G.; Utille, J.-P. *Carbohydr. Res.* **1975**, *39*, 368. (b) Lassaletta, J. M.; Carlsson, K.; Garegg, P. J.; Schmidt, R. R. *J. Org. Chem.* **1996**, *61*, 6873.
(17) Schmidt, R. R.; Stumpp, M. *Liebigs Ann. Chem.* **1983**, 1249.
(18) Boger, D. L.; Honda, T. *J. Am. Chem. Soc.* **1994**, *116*, 5647.
(19) Amvam-Zollo, P.-H.; Sinay, P. *Carbohydr. Res.* **1996**, *150*, 199.
(20) Salvador, L. A.; Elofsson, M.; Kihlberg, J. *Tetrahedron* **1995**, *51*, 5643.

Scheme 4



Scheme 5



contaminated with unreacted Fmoc-(*S*)-tyrosine. Removal of the Fmoc protecting group using a 4:1 mixture of dichloromethane−piperidine afforded the free amino acid, which was directly N-protected by acylation with 4-pentenoic acid succinimide ester to give amide **23** in pure form in 34% yield after chromatography on silica gel. Activation of the free acid as a cyanomethyl ester was carried out as previously using chloroacetonitrile and sodium carbonate in DMF. The ester **24** was obtained in 53% yield (Scheme 3).

The cyanomethyl esters **18a**, **18b**, and **24** were coupled with the tris(tetrabutylammonium) salt of pdCpA[11i] in acetonitrile at 25 °C to give the pdCpA esters **19a**, **19b**, and **25** in 58, 44, and 48% yields, respectively. The progress of the reactions was monitored by HPLC, and all reactions were complete within 24 h. The products were purified by reversed-phase high-performance liquid chromatography (HPLC).

The acid-catalyzed removal of the acetyl groups from the sugar moieties has been reported previously for glycosyl serine pdCpA esters.[13] The application of the same procedure in the case of glycosyl tyrosine pdCpA esters **19a**, **19b**, and **25** proved to be successful in the complete deacetylation of the glucose, galactose, and lactose moieties, respectively, without affecting the glycosidic or the aminoacyl bonds. Thus, treatment of compounds **19a**, **19b**, and **25** in methanol with a solution of tetrafluoroboric acid in diethyl ether for 3 h at room temperature gave the target pdCpA esters **20a**, **20b**, and **26** in 45, 40, and 35% yields, respectively. In each case, some partially deacetylated derivatives were detected by HPLC but not isolated. Longer reaction times increased the complexity of the reaction mixtures, and thus extending the reaction time was not a useful strategy. The overall yields of **20a** and **20b** from the trichloroacetimidates were 20 and 12%, respectively. The overall yield of **26** was only about 1%, reflecting the longer synthetic sequence employed for its preparation, but was still more than adequate to support the preparation of the glycosylated tyrosyl-$tRNA_{CUA}$.

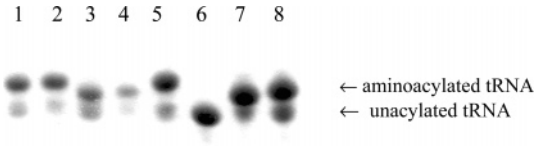


Figure 3. Analysis of protected aminoacylated $tRNA^{Phe}_{CUA}$ by electrophoresis on an acidic 8% polyacrylamide−7 M urea gel. Lane 1, 2,3,4,6-tetra-*O*-acetyl-*β*-D-galactosylseryl-tRNA; lane 2, 2,3,4,6-tetra-*O*-acetyl-*β*-D-glucosylseryl-tRNA; lane 3, *β*-D-galactosylseryl-tRNA; lane 4, tyrosyl-tRNA; lane 5, 4-*O*-(*β*-D-galactopyranosyl)-*β*-D-glucopyranosyltyrosyl-tRNA; lane 6, unacylated tRNA; lane 7, *β*-D-glucosyltyrosyl-tRNA; lane 8, *β*-D-galactosyltyrosyl-tRNA.

Table 1. Suppression Efficiencies of Aminoacyl-$tRNA_{CUA}$'s at Positions 10 and 27 of Dihydrofolate Reductase

	suppression efficiency (%)[a]	
amino acid	position 10	position 27
unacylated $tRNA_{CUA}$	1	
β-D-glucosyl-(*S*)-serine (**2**)	23	43
2,3,4,6-tetra-*O*-acetyl-*β*-D-galactosyl-(*S*)-serine (**3**)	15	10
β-D-galactosyl-(*S*)-serine (**4**)	18	16
α-D-mannosyl-(*S*)-serine (**6**)	34	48
[2,3,4,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-*β*-D-galactosyl)-*β*-D-glucosyl]-(*S*)-tyrosine (**11**)	3	4
4-*O*-(*β*-D-galactosyl)- *β*-D-glucosyl)-(*S*)-tyrosine (**12**)	11	12

[a] Suppression efficiency is defined as the yield of full-length protein obtained by readthrough of a UAG codon with a (glycosylated) aminoacyl-tRNA relative to the yield of unmodified protein obtained using the wild-type mRNA.

Preparation of Misacylated tRNAs. The synthesis of suppressor tRNAs containing glycosylated serines **1**−**6** was reported previously.[13] The preparation of misacylated tRNAs containing glycosylated tyrosine derivatives **7**−**12** was accomplished as shown in Schemes 4 and 5, by T4 RNA ligase-mediated coupling of N-protected glycosyl tyrosyl-pdCpAs with $tRNA^{Phe}_{CUA}$ transcript lacking the 3′-terminal cytidine and adenosine moieties.[11,21] The yield of misacylated tRNAs was determined following polyacrylamide gel electrophoresis (PAGE) under acidic conditions (Figure 3).[22] The ligation reactions typically proceeded largely to completion, but no purification was required in any case as the $tRNA_{CUA}$ transcript does not participate in protein synthesis.

***In Vitro* Synthesis of Proteins Containing Glycosylated Serine and Tyrosine Derivatives.** Two model proteins were used to study the incorporation of the glycosylated amino acids into proteins using misacylated suppressor tRNAs, and the successful incorporation of compounds **2**−**12** is reported. These proteins, namely *Photinus pyralis* luciferase and *Escherichia*

(21) (a) Lodder, M.; Golovine, S.; Hecht, S. M. *J. Org. Chem.* **1997**, *62*, 778. (b) Lodder, M.; Golovine, S.; Laikhter, A. L.; Karginov, V. A.; Hecht, S. M. *J. Org. Chem.* **1998**, *63*, 794.
(22) Varshney, U.; Lee, C.-P.; RajBhandary, U. L. *J. Biol. Chem.* **1991**, *266*, 24712.

Table 2. Suppression Efficiencies of Aminoacyl-$tRNA_{CUA}$'s at Position 284 of *P. pyralis* Luciferase

amino acid	suppression efficiency (%)[a]
unacylated $tRNA_{CUA}$	0.5
β-D-glucosyl-(*S*)-serine (**2**)	10
2,3,4,6-tetra-*O*-acetyl-β-D-galactosyl-(*S*)-serine (**3**)	9
α-D-mannosyl-(*S*)-serine (**6**)	13
β-D-glucosyl-(*S*)-tyrosine (**8**)	13
β-D-galactosyl-(*S*)-tyrosine (**10**)	19
4-*O*-(β-D-galactosyl)-β-D-glucosyl)-(*S*)-tyrosine (**12**)	30
serine	24

[a] Suppression efficiency is defined as the yield of full-length protein obtained by readthrough of a UAG codon with a (glycosylated) aminoacyl-tRNA relative to the yield of unmodified protein obtained using the wild-type mRNA.

Table 3. Mass Spectrometric Determination of Glycosylated Analogues of *E. coli* DHFR

		MW	
position	amino acid	calcd	found
10	valine (wild-type)	19 182	19 185
	β-D-galactosyl-(*S*)-serine (**4**)	19 334	19 303
22	tryptophan (wild-type)	19 182	19 185
	β-D-galactosyl-(*S*)-serine (**4**)	19 244	19 198

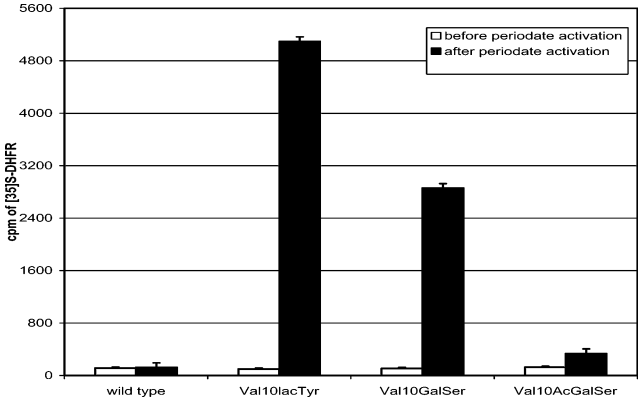


Figure 4. Binding of elaborated glycoproteins to a hydrazine surface on a microtiter plate following treatment with sodium periodate.

coli DHFR, have been employed in previous studies.[12a,b,23] In the present study, position 284 of luciferase and positions 10, 22, and 27 of DHFR were chosen for the incorporation of glycosylated amino acids. The preparation of both proteins was carried out in an *in vitro* protein biosynthesizing system in the presence of mRNA containing a stop codon at the position to be modified and a suppressor tRNA activated with the glycosylated amino acid of choice. Both prokaryotic (*E. coli*) and eukaryotic (rabbit reticulocyte) systems were evaluated for their ability to incorporate the glycosylated serine and tyrosine derivatives. The level of protein synthesis was estimated by phosphorimager analysis of the samples after electrophoresis on sodium dodecyl sulfate (SDS) polyacrylamide gels. As shown in Figures 1 and 2 in the Supporting Information, the translation technique allows the synthesis of full-length proteins from the modified mRNA only in the presence of misacylated tRNA; with a single exception (involving **11**, Table 1), the suppression efficiencies for all glycosylated serine and tyrosine derivatives varied from 9 to 48% (Tables 1 and 2). Readthrough with a natural amino acid in the absence of misacylated $tRNA_{CUA}$ was minimal (0.5−1.0%, Tables 1 and 2) and is likely significantly lower in the presence of the misacylated $tRNA_{CUA}$.

(23) (a) Karginov, V. A.; Mamaev, S. V.; Hecht, S. M. *Nucleic Acids Res.* **1997**, *25*, 3912. (b) Short, G. F., III; Golovine, S. Y.; Hecht, S. M. *Biochemistry* **1999**, *38*, 8808. (c) Karginov, A. V.; Lodder, M.; Hecht, S. M. *Nucleic Acids Res.* **1999**, *27*, 3283. (d) Karginov, V. A.; Mamaev, S. V.; An, H.; Van Cleve, M. D.; Hecht, S. M.; Komatsoulis, G. A.; Abelson, J. N. *J. Am. Chem. Soc.* **1997**, *119*, 8166. (e) Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M. *J. Am. Chem. Soc.* **2003**, *125*, 6616.

Table 4. Emission Characteristics of Glycosylated Analogues of *P. pyralis* Luciferase

amino acid in position 284	specific activity (%)	emission wavelength (nm)
serine (wild-type)	103 ± 2	561
β-D-glucosyl-(*S*)-serine (**2**)	10 ± 1	561
β-D-galactosyl-(*S*)-serine (**4**)	17 ± 2	561
β-D-mannosyl-(*S*)-serine (**6**)	12 ± 2	561
2,3,4,6-tetra-*O*-acetyl-β-D-galactosyl-(*S*)-serine (**3**)	3 ± 1	585
2,3,4,6-tetra-*O*-acetyl-α-D-mannosyl-(*S*)-serine (**5**)	4 ± 1	587
tyrosine	6 ±1	585
phenylalanine	14 ± 2	587
2-naphthylalanine	15 ± 2	603
β-D-glucosyl-(*S*)-tyrosine (**8**)	3 ± 1	601
β-D-galactosyl-(*S*)-tyrosine (**10**)	2 ± 0.2	603
β-D-lactosyl-(*S*)-tyrosine (**12**)	2 ± 0.2	599
2,3,4,6-tetra-*O*-acetyl-β-D-glucosyl-(*S*)-tyrosine (**7**)	5 ± 1	585
2,3,4,6-tetra-*O*-acetyl-β-D-galactosyl-(*S*)-tyrosine (**9**)	4 ± 1	586

To confirm the presence of the sugar moieties in the modified proteins, two glycosylated DHFR analogues were synthesized on a preparative scale (∼5−20 μg) and were subjected to mass spectrometric analysis. The MALDI spectra were helpful and provided masses within 31−46 Da of the calculated molecular weights for the glycosylated proteins (Table 3). The agreement between calculated and found masses was typical of values obtained previously for glycosylated proteins and undoubtedly reflects sample size and preparation.[24] To further confirm the presence of the carbohydrate moiety in the synthesized proteins, three DHFR analogues (modified at position 10) were prepared in the presence of galactosylseryl-, lactosyltyrosyl-, and tetra-*O*-acetylgalactosylseryl-tRNAs and were then assayed for their ability to bind to a hydrazine surface after oxidative degradation of the vicinal diols of the sugar moiety. The ^{35}S-labeled DHFRs were synthesized *in vitro* and purified by (diethylamino)ethyl (DEAE) Sepharose and Ni−nitrilotriacetic acid (NTA) agarose chromatography. Following treatment with sodium periodate to effect oxidative cleavage of the vicinal diols, the proteins were incubated with a hydrazine-containing surface. After the non-covalently bound proteins were washed away, the amount of protein specifically retained on the surface was determined by radioactivity. As expected, the modified DHFRs containing lactosyltyrosine (**12**) and galactosylserine (**4**) at position 10 bound to the hydrazine surface, unlike the non-glycosylated wild-type DHFR or the peracetylated galactosylserine (**3**) (Figure 4). This result confirms unequivocally the presence of the sugar moiety on the protein, i.e., that the glycosylated misacylated

(24) See, e.g.: (a) Witte, K.; Sears, P.; Martin, R.; Wong, C.-H. *J. Am. Chem. Soc.* **1997**, *119*, 2114. (b) Macmillan, D.; Bill, R. M.; Sage, K. A.; Fern, D.; Flitsch, S. L. *Chem. Biol.* **2001**, *8*, 133.

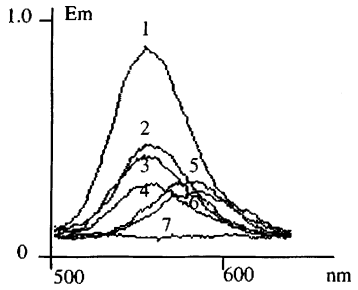


Figure 5. Light emission spectra of different samples of *P. pyralis* luciferase synthesized with different glycosylated serine derivatives at position 284. Curve 1, serine; curve 2, β-D-glucosylserine (**2**); curve 3, α-D-mannosylserine (**6**); curve 4, β-D-galactosylserine (**4**); curve 5, 2,3,4,6-tetra-*O*-acetyl-β-D-galactosylserine (**3**); curve 6, 2,3,4,6-tetra-*O*-acetyl-α-D-mannosylserine (**5**); curve 7, background.

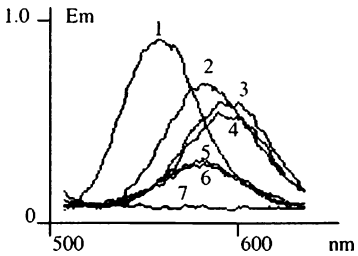


Figure 6. Light emission spectra of samples of *P. pyralis* luciferase synthesized with different glycosylated tyrosine derivatives at position 284. Curve 1, serine; curve 2, tyrosine; curve 3, β-D-glucosyltyrosine (**8**); curve 4, β-D-galactosyltyrosine (**10**); curve 5, 2,3,4,6-tetra-*O*-acetyl-β-D-glucosyltyrosine (**7**); curve 6, 2,3,4,6-tetra-*O*-acetyl-β-D-galactosyltyrosine (**9**); curve 7, background.

tRNAs mediate the synthesis of modified proteins containing glycosylated amino acids.

Effect of Glycosylation at Position 284 of *Photinus pyralis* Luciferase on the Light Emission Spectrum. Using *P. pyralis* luciferase as a model, we studied the effect of glycosylation on the light emission spectrum of this enzyme. Ten analogues containing different glycosylated derivatives of serine and tyrosine were synthesized by readthrough of a UAG codon with misacylated $tRNA_{CUA}$'s. Analogues of luciferase containing non-glycosylated natural and unnatural amino acids (serine, tyrosine, phenylalanine, and 2-naphthylalanine) were also prepared. Luciferase synthesized from the wild-type gene was used as a control in all experiments. The light emission spectra of all luciferases were determined and compared using two parameters: maximum wavelength (λ_{max}) and relative specific activity (the intensity of light emission produced by equal amounts of enzyme, expressed as a percent of light emission relative to the wild-type enzyme). Table 4 shows that luciferase synthesized from a modified mRNA in the presence of seryl-tRNA has the same specific activity and wavelength characteristics as the authentic wild-type enzyme. A modified enzyme with tyrosine at position 284 gave about a 20-fold decrease in the specific activity and a red shift of 24 nm. Modified enzymes with glycosylated serines **2**−**6** gave a red shift of about 24−26 nm in emission spectra only when the sugar moiety was fully acetylated (Figure 5). However, an opposite effect was observed in the case of modified luciferases containing glycosylated tyrosines **7**−**10** and **12**. Only mutants incorporating unprotected glycosylated tyrosines **8**, **10**, and **12** gave a red shift of about 14−18 nm compared to luciferase with tyrosine at position 284 (Figure 6).

Discussion

Recent advances in carbohydrate chemistry and solid-phase synthesis have made possible the synthesis of glycopeptides and their mimics bearing reasonably complex glycans,[25] and some impressive structures have been reported.[26] However, efficient access to full-length glycoproteins is still a challenging task. The classical methods for protein glycosylation take advantage of the inherent reactivity of the lysine or cysteine side chain.[27,28] However, these methods are not residue specific and result in multiple glycosylations when more than one lysine or cysteine residue are present in the protein. This problem has been circumvented by incorporating a single cysteine residue using site-directed mutagenesis, followed by selective glycosylation.[24b,29] Although a small library of mono- and diglycosylated proteins has been prepared, the resulting linkages are different structurally from the natural ones. This may be problematic, given that some reports have shown that the first and second sugar residues directly linked to the protein can play a major role in modulating local peptide conformation.[8c,e] Perturbation of this linkage might result in loss of native conformation; therefore, protein function could be altered. Another approach for the synthesis of modified glycoproteins using chemoselective ligation has been reported by Bertozzi and others.[30] Native chemical ligation[31] has also been used for the total chemical synthesis of complex glycoproteins.[32] The preparation of glycoproteins using expressed

(25) For reviews, see: (a) Kunz, H. *Angew. Chem., Int. Ed. Engl.* **1987**, *26*, 294. (b) Kunz, H. *Pure Appl. Chem.* **1993**, *65*, 1223. (c) Garg, H. G.; von dem Burgh, K.; Kunz, H. *Adv. Carbohydr. Chem. Biochem.* **1994**, *50*, 277. (d) Meldal, M. *Curr. Opin. Struct. Biol.* **1994**, *4*, 710. (e) Arsequell, G.; Valencia, G. *Tetrahedron: Asymmetry* **1997**, *8*, 2839. (f) Taylor, C. M. *Tetrahedron* **1998**, *54*, 11317. (g) Arsequell, G.; Valencia, G. *Tetrahedron: Asymmetry* **1999**, *10*, 3045. (h) Seitz, O. *ChemBioChem.* **2000**, *1*, 214.

(26) (a) Paulsen, H.; Bielfeldt, T.; Peters, S.; Meldal, M.; Bock, K. *Liebigs Ann. Chem*. **1994**, 369. (b) Rio-Anneheim, S.; Paulsen, H.; Meldal, M.; Bock, K. *J. Chem. Soc., Perkin Trans. 1* **1995**, 1071. (c) Rademann, J.; Schmidt, R. R. *Carbohydr. Res.* **1995**, *269*, 217. (d) Jansson, A. M.; Jensen, K. J.; Meldal, M.; Lamako, J.; Wieslawa, M.; Olsen, C. E.; Bock, K. *J. Chem. Soc., Perkin Trans. 1* **1996**, 1001. (e) Nakahara, Y.; Nakahara, Y.; Ogawa, T. *Carbohydr. Res.* **1996**, *292*, 71. (f) Danishefsky, S. J.; Hu, S.; Cirillo, P. F.; Eckhardt, M.; Seeberger, P. *Chem. Eur. J.* **1997**, *3*, 1617. (g) Seitz, O.; Wong, C.-H. *J. Am. Chem. Soc.* **1997**, *119*, 8766. (h) Schwartz, J. B.; Kuduk, S. D.; Chen, X.-T.; Sames, D.; Glunz, P. W.; Danishefsky, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 2662.

(27) (a) Buss, D. H.; Goldstein, I. J. *J. Chem. Soc. C* **1968**, 1457. (b) Gray, G. R. *Arch. Biochem. Biophys.* **1974**, *163*, 426. (c) Lemieux, R. U.; Bundle, D. R.; Backer, D. A. *J. Am. Chem. Soc.* **1975**, *97*, 4076. (d) Lee, Y. C.; Stowell, C. P.; Krantz, M. J. *Biochemistry* **1976**, *15*, 3956. (e) Quétard, C.; Bourgerie, S.; Normand-Sdiqui, N.; Mayer, R.; Strecker, G.; Midoux, P.; Roche, A.-C.; Monsigny, M. *Bioconjugate Chem.* **1998**, *9*, 268.

(28) (a) Davis, N. J.; Flitsch, S. L. *Tetrahedron Lett.* **1991**, *32*, 6793. (b) Wong, S. Y. C.; Guile, G. R.; Dwek, R. A.; Arsequell, G. *Biochem. J.* **1994**, *300*, 843. (c) Macindoe, W. M.; Van Oigen, A. H.; Boons, G.-J. *J. Chem. Soc., Chem. Commun.* **1998**, 847. (d) Shin, I.; Jung, H.-J.; Lee, M.-R. *Tetrahedron Lett.* **2001**, *42*, 1325.

(29) (a) Davis, B. G.; Lloyd, R. C.; Jones, J. B. *J. Org. Chem.* **1998**, *63*, 9614. (b) Davis, B. G.; Jones, J. B. *Synlett* **1999**, 1495−1507. (c) Davis, B. G.; Lloyd, R. C.; Jones, J. B. *Bioorg. Med. Chem.* **2000**, *8*, 1527. (d) Davis, B. G.; Jones, J. B. *Bioorg. Med. Chem.* **2000**, *8*, 1537. (e) Davis, B. G.; Maughan, M. A. T.; Green, M. P.; Ullman, A. *Tetrahedron: Asymmetry* **2000**, *11*, 245.

(30) (a) Cervigni, S. E.; Dumy, P.; Mutter, M. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1230. (b) Rodriguez, E. C.; Winans, K. A.; King, D. S.; Bertozzi, C. *R. J. Am. Chem. Soc.* **1997**, *119*, 9905. (c) Zhao, Y.; Kent, S. B. H.; Chait, B. T. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 1629. (d) Lemieux, G. A.; Bertozzi, C. R. *TIBTECH* **1998**, *16*, 506. (e) Marcaurelle, L. A.; Bertozzi, C. R. *Chem. Eur. J.* **1999**, *5*, 1384. (f) Peluso, S.; Imperiali, B. *Tetrahedron Lett.* **2001**, *42*, 2085. (g) Durieux, P.; Fernandez-Carneado, J.; Tuchscherer, G. *Tetrahedron Lett.* **2001**, *42*, 2297. (h) Muir, T. W. *Annu. Rev. Biochem.* **2003**, *72*, 249.

(31) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776.

(32) (a) Shin, Y.; Winans, K. A.; Bakes, B. J.; Kent, S. B. H.; Ellman, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1999**, *121*, 11684. (b) Marcaurelle, L. A.; Mizoue, L. S.; Wilken, J.; Oldham, L.; Kent, S. B. H.; Handel, T. M.; Bertozzi, C. R. *Chem. Eur. J.* **2001**, *7*, 1129. (c) Miller, J. S.; Dudkin, V. Y.; Lyon, G. J.; Muir, T. W.; Danishefsky, S. J. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 431.

protein ligation has also been studied.[33,34] A new methodology for the enzymatic and chemical synthesis of pure glycoform proteins has been developed by Wong and co-workers.[24a,35] By the use of glycosidases, the natural *N*-glycan structure from ribonuclease B was truncated, leaving only the *β*-*N*-GlcNAc residue. The protein thus obtained was subjected to a series of glycotransferase-catalyzed reactions to give a new glycoform bearing the SLe^x tetrasaccharide. Mizuno and co-workers developed an enzyme-catalyzed transglycosylation reaction to transfer an entire oligosaccharide onto a GlcNAc-containing peptide.[36] In a similar manner, endoglycosidase-catalyzed transglycosylation has been applied to the synthesis of a Man_6-$GlcNAc_2$ glycoprotein.[37]

Through the use of misacylated tRNAs that have been activated with noncognate amino acids, many peptides and proteins containing unnatural amino acids have been elaborated.[11] To produce proteins containing modified amino acids at a single site, a unique codon is employed in combination with a misacylated tRNA bearing the corresponding anticodon. The most widely used combination involves the use of the stop codon UAG placed at the site of interest and a chemically misacylated suppressor tRNA capable of reading through that particular nonsense codon (Figure 1). Recently, a few studies have demonstrated the feasibility of preparing glycoproteins with monosaccharides attached at specific sites using this method.[12] A key step in this strategy involves the synthesis of tRNAs acylated with glycosylated amino acids. In this fashion, glucosyl serine **2** has been incorporated in place of serine at position 286 of *Luciola mingrelica* luciferase and also at position 10 of DHFR.[12a,b] Schmidt and co-workers have also described the preparation of hARF protein altered to contain *N*-acetyl-3,4,6-tri-*O*-acetyl-D-glucosylserine instead of lysine and several glycosylated amino acids at position 133 of human interleukin granulocyte-colony stimulating factor.[14] An alternative translation system using a four-base codon for incorporation of glycosylated amino acids is under development by Sisido and co-workers.[38] An important characteristic of this approach is the ability to prepare numerous analogues of a specific protein to permit definition of the effects of such changes.

Finally, Wong and Schultz have recently described the selection of novel aminoacyl-tRNA synthetases that enabled the incorporation of *N*-acetylglucosaminyl-*β*-*O*-serine and *N*-acetylgalactosylamino-α-*O*-threonine into proteins such as myoglobin interleukin and RNase A in an intact bacterial system.[15] The use of whole cell systems for the preparation of (glyco)proteins has the important advantage of providing access to larger amounts of protein, thereby enabling detailed studies of the derived proteins. It may be noted, however, that the whole cell expression of glycoproteins potentially requires the selection of a novel aminoacyl-tRNA synthetase for each glycoform. Thus, the approach described in the present article may be considered more convenient when only limited amounts of a larger variety of glycoforms are of interest.

From a synthetic point of view, in addition to the problems related to the stability of the tRNA molecule and the amino acid protecting groups that are used, the sugar moiety presents its own technical problems. Especially in the case of *O*-glycosylated serines, the acid and base lability of the glycosidic bond complicates the choice of sugar OH protecting groups. The α-amino groups of serine and tyrosine were protected with the 4-pentenoyl group or the 2-nitroveratryl (NVOC) group. These protecting groups can be removed easily, at the tRNA level, under mild conditions without affecting the integrity of the tRNA molecule or the aminoacyl bond.[39] The sugar OH groups were protected with acetyl groups during the first steps of the synthesis. The deacetylation of the sugar moiety at the pdCpA level was performed successfully under acidic conditions without disrupting the sensitive aminoacyl bond to the tRNA.

The incorporation of glycosylated serine and tyrosine derivatives into positions 10 and 27 of DHFR was demonstrated. Although individual glycosylated amino acids **2**−**4**, **6**, **11**, and **12** were generally incorporated with quite similar efficiencies at both positions, the level of incorporation varied significantly from one glycosylated amino acid to another (3−43% suppression efficiency) (Table 1). The electrophoretic analysis of the DHFR analogues containing glycosylated amino acids at position 10 is shown (Figure 1 in the Supporting Information). The factors that control the efficiency of ribosomal incorporation of individual amino acids are not well understood at present, although it has been shown that analogues of polar and charged amino acids tend to be incorporated with less efficiency.

Another enzyme chosen for study was firefly luciferase. The crystal structure of firefly luciferase from *P. pyralis*, which has been reported,[40] shows that Ser 284 is situated far from the putative active site of this enzyme (Figure 7). Nonetheless, we reported previously that substitution of Ser 284 by other amino acids can lead to significant changes in the emission spectrum.[41] As noted previously,[12a,b,41] the introduction of lipophilic amino acids into this position (or the analogous position in *L. mingrelica* luciferase) resulted in a shift to longer wavelength of the light emitted. This was interpreted in terms of the presence of several lipophilic amino acids on the loops adjacent to the amino acid at position 284, which presumably resulted in the aggregation of hydrophobic residues from those loops with the lipophilic amino acid incorporated at position 284. The resulting change in protein conformation in this region could plausibly alter the adenoside triphosphate (ATP) and luciferin binding domains sufficiently to produce the observed alterations in the wavelength of emitted light. This interpretation is fully consistent with the earlier finding that point mutations in many regions of *L. cruciata* luciferase resulted in a shift in the wavelength of emitted light.[42]

In comparison, the introduction of large or polar amino acids into position 286 of *L. mingrelica* luciferase had no effect on

(33) Tolbert, T. J.; Wong, C.-H *J. Am. Chem. Soc.* **2000**, *122*, 5421.
(34) Macmillan, D.; Bertozzi, C. R. *Tetrahedron* **2000**, *56*, 9515.
(35) Liu, H.; Wang, L.; Brock, A.; Wong, C.-H.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 1702.
(36) Mizuno, M.; Haneda, K.; Iguchi, R.; Muramoto, I.; Kawakami, T.; Aimoto, S.; Yamamoto, K.; Inazu, T. *J. Am. Chem. Soc.* **1999**, *121*, 284.
(37) Takegawa, K.; Tabuchi, M.; Yamaguchi, S.; Kondo, A.; Kato, I.; Iwahara, S. *J. Biol. Chem.* **1995**, *270*, 3094.
(38) (a) Manabe, S.; Sakamoto, K.; Nakahara, Y.; Sisido, M.; Hohsaka, T.; Ito, Y. *Bioorg. Med. Chem.* **2002**, *10*, 573. (b) Matsubara, T. H.; Oiwa, K.; Hohsaka, T.; Sadamoto, R.; Niikura, K.; Fukuhara, N.; Takimoto, A.; Kondo, H.; Nishimura, S. *Chem.-Eur. J.* **2005**, *11*, 6974.
(39) Hecht, S. M. Synthesis of Modified Proteins Using Misacylated tRNAs. In *Protein Engineering*; RajBhandary, U. L., Koehrer, C., Eds.; Springer: Berlin, 2007, in press.
(40) Conti, E.; Franks, N. P.; Brick, P. *Structure* **1996**, *4*, 287.
(41) Hecht, S. M.; Wang, B.; Mamaev, S. V.; Arslan, T.; Short, G. F., III; Lodder, M.; Golovine, S. *Nucleic Acids Res. Symp. Ser.* **1998**, *39*, 15.
(42) Kajiyama, N.; Nakano, E. *Protein Eng.* **1991**, *4*, 691.

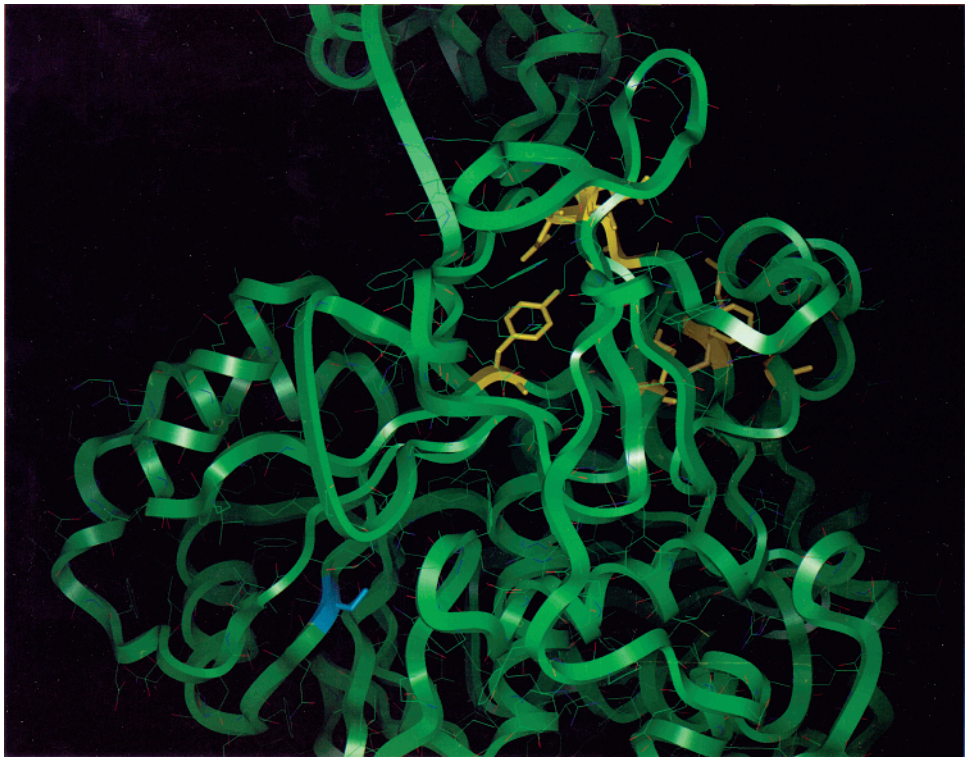


Figure 7. Structure of *P. pyralis* firefly luciferase illustrating Ser 284 (blue) and key residues that form the putative ATP binding pocket (yellow).

the wavelength of emitted light. Since *L. mingrelica* and *P. pyralis* luciferases exhibit significant homology, it seemed reasonable to interpret this in terms of the structure for the *P. pyralis* luciferase, which places the side chain of Ser 284 in a solvent-exposed region. However, until the present study, no amino acids containing large, polar side chains had actually been introduced into position 284 of *P. pyralis* luciferase. As noted in Table 4, the introduction of three monoglycosylated serine derivatives into position 284 had no effect on the wavelength of emitted light, consistent with our earlier study using *L. mingrelica* luciferase. The more lipophilic acetylated analogues **3** and **5** afforded luciferases that emitted at significantly longer wavelength when introduced into position 284. Interestingly, the effects of sugar acetylation were reversed for the *P. pyralis* luciferases containing glycosylated tyrosine derivatives (Table 4). Presumably, the aromatic moiety in tyrosine provides sufficient lipophilic character to produce the effects noted previously;[12a,b] the sugar moieties, which are farther from the protein backbone than those in glycosylated serine derivatives, must participate in novel interactions not observed with other substitutions at position 284 (286). The same effect was also observed in the case of the luciferase from *L. mingrelica*.[12a,b] It is likely that this residue may well play an important role in maintaining the native structure of luciferase, which forms a pocket important for luciferin binding and light emission. A further observation worthy of note is that the introduction of acetyl groups on the sugar moieties of the glycosylated tryosines resulted in a shift to shorter wavelength and perhaps a small increase in the efficiency of light emission. Thus, unlike the observations made for serine derivatives at position 284, the hydrophilicity of tyrosine analogues in this position did not produce a uniform response in terms of the resulting wavelength of emitted light.

The red shift in emission spectra is usually linked to a change in the polarity of the substrate binding center,[43] which is a function of two main parameters: hydrophilicity and mobility of the molecules. They determine the ability of the molecules surrounding the chromophore to support the optimal excitation−emission events. The results obtained demonstrate the subtle role of Ser 284 in this event. Previously, we reported that the introduction of charged or polar amino acids at position 284 had little effect on the wavelength of emitted light, while lipophilic amino acids resulted in a shift to longer wavelength.[12a,b,41] The difference between the emission spectra of analogues containing peracetylated and deacetylated sugars attached to serine is entirely consistent with the earlier observation (Table 4). As noted previously,[12b] Ser 284 is within reasonable proximity of a number of lipophilic amino acids on the surface of *P. pyralis*[40] and may contribute to the formation or stability of a hydrophobic pocket on the surface of the enzyme. The size of the side chain in this position is also important for controlling the wavelength of emitted light. Analogues with tyrosine and phenylalanine, which have side chains of practically the same size, exhibited the same emission spectra. The same red shift was observed in the case of larger analogues with a glycosylated tyrosine or 2-naphthylalanine side-chain group. It is interesting to note that the specific activity of the analogues having hydrophobic residues such as phenylalanine and 2-naphthylalanine was greater than that of those containing the more hydrophilic tyrosine and its glycosylated analogues.

The results presented here clearly demonstrate that a number of glycosylated serine and tyrosine derivatives have been

(43) (a) White, E. H.; Rapaport, E.; Seliger, H. H.; McElroy, W. D. *Bioorg. Chem.* **1971**, *1*, 92. (b) Wood, K. V. *Photochem. Photobiol.* **1995**, *62*, 662. (c) Branchini, B. R.; Magyar, R. A.; Murtiashaw, M. H.; Anderson, S. M.; Helgerson, L. C.; Zimmer, M. *Biochemistry* **1999**, *38*, 13223. (d) McCapra, F. *Methods Enzymol.* **2000**, *305*, 3.

incorporated into DHFR and firefly luciferase. This is the first study in which glycosylated tyrosine analogues have been incorporated into proteins. While glycosylated tyrosine derivatives are less abundant in nature than those of serine and asparagine, the present findings nonetheless significantly extend the repertoire of carbohydrate modifications available for elaborating glycoproteins. They also show the significant possibilities offered by unnatural amino acids in the study of interactions involving functional sites of complex proteins.

Experimental Section

General Methods. ^{1}H and ^{13}C NMR spectra were recorded on a General Electric QE-300 spectrometer. Moisture-sensitive reactions were conducted under argon in oven-dried glassware. All chemical reagents were purchased from Aldrich Chemical Co. (Milwaukee, WI) or Sigma Chemical Co. (St. Louis, MO) and used without further purification. The amino acids were purchased from Novabiochem. Acetonitrile and dichloromethane were distilled from CaH_2. DMF was distilled from CaH_2 under diminished pressure. Triethylamine was distilled from P_2O_5. Analytical thin-layer chromatography was performed on 60 F_{254} (E. Merck) plates and visualized using iodine or sulfuric acid. Flash chromatography was performed using 230−400 mesh silica gel. High-resolution mass spectra were recorded at the Michigan State University-NIH Mass Spectrometry Facility or at the Nebraska Center for Mass Spectrometry. [^{35}S]Methionine (1000 Ci/mmol, 10 μCi/μL) was purchased from Amersham Corp. AmpliScribe translation kits were from Epicentre Technologies (Madison, WI). Ni-NTA resin was obtained from Qiagen (Valencia, CA). Nuclease-treated rabbit reticulocyte lysate system, amino acid mixtures, and the luciferase assay system were purchased from Promega (Madison, WI). T4 DNA ligase, purified acylated bovine serum albumin (BSA), T4 RNA ligase, and *Fok*I endonuclease were obtained from New England Biolabs (Beverly, MA). DEAE Sepharose CL-6B was from Sigma Chemical Co. Fluorescence spectral data were obtained using a Hitachi F2000 fluorescence spectrophotometer. Ultraviolet spectral measurements were made using a Perkin-Elmer Lambda 20 UV/vis spectrometer. Radioactivity measurements were performed with a Beckman LS-100C liquid scintillation counter. SDS−PAGE was performed according to the Laemmli procedure.[44] Gels were visualized and quantified by phosphorimager analysis using a Molecular Dynamics 400E Phosphorimager equipped with ImageQuant version 5.0.

Synthesis of Aminoacylated pdCpA Derivatives. *N*-(6-Nitroveratryloxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-*β*-D-glucopyranosyl)-(*S*)-tyrosine Cyanomethyl Ester (18a). To a solution containing 90 mg (0.18 mmol) of the α-trichloroacetimidate **17a**[17] and 65 mg (0.14 mmol) of NVOC-(*S*)-tyrosine cyanomethyl ester (**15**)[12b] in 1.5 mL of CH_2Cl_2 was added 18 μL (0.14 mmol) of BF_3·OEt_2. After the solution was stirred at room temperature for 3 h, another 49 mg (98 μmol) of the trichloroacetimidate was added, and stirring was continued for 2 h. The reaction was quenched by addition of 1 mL of 1 N $NaHSO_4$ and extracted with three 5-mL portions of CH_2Cl_2. The organic phase was dried ($MgSO_4$) and concentrated under diminished pressure. The crude product was applied to a silica gel column (30 × 1 cm); elution with 60:40 ethyl acetate−hexanes provided *N*-(6-nitroveratryloxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-*β*-D-glucopyranosyl)-(*S*)-tyrosine cyanomethyl ester (**18a**) as a yellow foam: yield 86 mg (77%); ^{1}H NMR ($CDCl_3$) δ 2.00, 2.02, 2.03, 2.05 (4s, 12H), 3.10 (d, 2H, J = 5.9 Hz), 3.84 (m, 1H), 3.88 (s, 6H), 4.14 (d, 1H, J = 2.1 H), 4.26 (dd, 1H, J = 5.0 Hz, J = 12.2 Hz), 4.73 (m, 2H), 5.05 (d, 1H, J = 7.0 Hz), 5.15 (q, 1H, J = 8.1 Hz), 5.25 (m, 2H), 5.47 (q, 2H, J = 14.9 Hz), 6.89 (s, 1H), 6.92 (d, 2H, J = 8.5 Hz), 7.05 (d, 2H, J = 8.5 Hz), 7.67 (s, 1H); ^{13}C NMR ($CDCl_3$) δ 20.54, 36.80, 48.90, 54.59, 56.35, 61.79, 64.15, 68.10, 71.00, 71.95, 72.59, 98.86, 108.15, 110.11, 113.64, 117.35, 129.48, 130.29, 155.17, 156.16, 169.35, 170.16, 170.55; mass spectrum (FAB) *m*/*z* 790.2338 $(M + H)^+$ ($C_{35}H_{40}N_3O_{18}$ requires *m*/*z* 790.2307).

***N*-(6-Nitroveratryloxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-*β*-D-galactopyranosyl)-(*S*)-tyrosine Cyanomethyl Ester (18b).** To a solution containing 139 mg (0.28 mmol) of α-trichloroacetimidate **17b**[19] and 65 mg (0.14 mmol) of NVOC-(*S*)-tyrosine cyanomethyl ester (**15**)[12b] in 2.5 mL of CH_2Cl_2 was added 18 μL (0.14 mmol) of BF_3·OEt_2. After the solution was stirred at room temperature for 2 h, the reaction was quenched by the addition of 2 mL of 1 N $NaHSO_4$ and extracted with three 5-mL portions of CH_2Cl_2. The organic phase was dried ($MgSO_4$) and concentrated under diminished pressure. The crude product was applied to a silica gel column (28 × 2 cm); elution with 50:50 and then 60:40 ethyl acetate−hexanes provided *N*-(6-nitroveratryloxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-*β*-D-galactopyranosyl)-(*S*)-tyrosine cyanomethyl ester (**18b**) as a yellow foam: yield 79 mg (70%); $[\alpha]_D$ +6 (*c* 1.1, $CHCl_3$); ^{1}H NMR ($CDCl_3$) δ 2.00, 2.05, 2.06, 2.17 (4s, 12H), 3.11 (d, 2H, J = 5.8 Hz), 3.94 (s, 6H), 4.07 (m, 2H), 4.18 (m, 2H), 5.03 (d, 1H, J = 8 Hz), 5.10 (dd, 1H, J = 3.4 Hz, J = 1.03 Hz), 5.26 (d, 1H, J = 7.5 Hz), 5.46 (m, 4H), 6.93 (d, 2H, J = 8.8 Hz), 6.97 (s, 1H), 7.08 (d, 2H, J = 8.08 Hz), 7.69 (s, 1H); mass spectrum (FAB) *m*/*z* 790.2323 $(M + H)^+$ ($C_{35}H_{40}N_3O_{18}$ requires *m*/*z* 790.2307).

***N*-(6-Nitroveratryloxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-*β*-D-glucopyranosyl)-(*S*)-tyrosine pdCpA Ester (19a).** To a conical vial containing 21 mg (25.7 μmol) of *N*-(6-nitroveratryloxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-*β*-D-glucopyranosyl)-(*S*)-tyrosine cyanomethyl ester (**18a**) was added a solution of 7 mg (5.15 μmol) of the tris(tetrabutylammonium) salt of pdCpA[11i] in 90 μL of freshly distilled acetonitrile. The reaction mixture was stirred at room temperature for 24 h. A 2-μL aliquot of the reaction mixture was diluted with 58 μL of 1:1 acetonitrile−50 mM NH_4OAc, pH 4.5, and was analyzed by HPLC on a C_{18} reversed-phase column (250 × 10 mm). The column was washed with 1% → 63% acetonitrile in 50 mM NH_4OAc, pH 4.5, over a period of 45 min at a flow rate of 3.5 mL/min (monitoring at 260 nm). The remaining reaction mixture was diluted to a total volume of 800 μL of 1:1 acetonitrile−50 mM NH_4OAc, pH 4.5, and purified using the same C_{18} reversed-phase column. Dinucleotide derivative **19a** (retention time 27.3 min) was recovered from the appropriate fractions by lyophilization as a colorless solid: yield 4.1 mg (58%); mass spectrum (ESI) *m*/*z* 1367.3 $(M - H)^-$ (theoretical *m*/*z* 1367.3).

***N*-(6-Nitroveratryloxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-*β*-D-galactopyranosyl)-(*S*)-tyrosine pdCpA Ester (19b).** To a conical vial containing 15 mg (18.7 μmol) of *N*-(6-nitroveratryloxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-*β*-D-galactopyranosyl)-(*S*)-tyrosine cyanomethyl ester (**18b**) was added a solution of 5.1 mg (3.75 μmol) of the tris(tetrabutylammonium) salt of pdCpA[11i] in 60 μL of freshly distilled acetonitrile, followed by 5 μL of triethylamine. The reaction mixture was stirred at room temperature for 20 h. A 2-μL aliquot of the reaction mixture was diluted with 58 μL of 1:1 acetonitrile−50 mM NH_4OAc, pH 4.5, and was analyzed by HPLC on a C_{18} reversed-phase column (250 × 10 mm). The column was washed with 1% → 63% acetonitrile in 50 mM NH_4OAc, pH 4.5, over a period of 45 min at a flow rate of 3.5 mL/min (monitoring at 260 nm). The remaining reaction mixture was diluted to a total volume of 600 μL of 1:1 acetonitrile−50 mM NH_4OAc, pH 4.5, and purified using the same C_{18} reversed-phase column. Dinucleotide derivative **19b** (retention time 28.6 min) was recovered from the appropriate fractions by lyophilization as a colorless solid: yield 2.3 mg (44%); mass spectrum (ESI) *m*/*z* 1367.6 $(M - H)^-$ (theoretical *m*/*z* 1367.3).

***N*-(6-Nitroveratryloxycarbonyl)-*O*-*β*-D-glucopyranosyl-(*S*)-tyrosine pdCpA Ester (20a).** To a conical vial containing 2 mg (1.46 μmol) of *N*-(6-nitroveratryloxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-*β*-D-glucopyranosyl)-(*S*)-tyrosine pdCpA ester (**19a**) in 300 μL of dry methanol was added 75 μL of a 54% solution of HBF_4 in diethyl ether. The reaction mixture was stirred at room temperature for 3 h. A 15-μL aliquot of the reaction mixture was diluted with 35 μL of 1:1 acetonitrile−50 mM NH_4OAc, pH 4.5, and was analyzed by HPLC on

(44) Laemmli, V. K. *Nature* **1970**, *227*, 680.

a C_{18} reversed-phase column (250 × 10 mm). The column was washed with 1% → 63% acetonitrile in 50 mM NH_4OAc, pH 4.5, over a period of 45 min at a flow rate of 3.5 mL/min (monitoring at 260 nm). The remaining reaction mixture was diluted to a total volume of 600 μL of 1:1 acetonitrile−50 mM NH_4OAc, pH 4.5, and purified using the same C_{18} reversed-phase column. Dinucleotide derivative **20a** (retention times 17.5 and 18.2 min) was recovered from the appropriate fractions by lyophilization as a colorless solid: yield 0.8 mg (45%); mass spectrum (ESI) *m*/*z* 1199.4 $(M - H)^-$ (theoretical *m*/*z* 1199.3).

***N*-(6-Nitroveratryloxycarbonyl)-*O*-β-D-galactopyranosyl-(*S*)-tyrosine pdCpA Ester (20b).** To a conical vial containing 2.3 mg (1.68 μmol) of *N*-(6-nitroveratryloxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-(*S*)-tyrosine pdCpA ester (**19b**) in 350 μL of dry methanol was added 90 μL of a 54% solution of HBF_4 in diethyl ether. The reaction mixture was stirred at room temperature for 3 h. A 20-μL aliquot of the reaction mixture was diluted with 30 μL of 1:1 acetonitrile−50 mM NH_4OAc, pH 4.5, and was analyzed by HPLC on a C_{18} reversed-phase column (250 × 10 mm). The column was washed with 1% → 63% acetonitrile in 50 mM NH_4OAc, pH 4.5, over a period of 45 min at a flow rate of 3.5 mL/min (monitoring at 260 nm). The remaining reaction mixture was diluted to a total volume of 700 μL of 1:1 acetonitrile−50 mM NH_4OAc, pH 4.5, and purified using the same C_{18} reversed-phase column. Dinucleotide derivative **20b** (retention times 17.5 and 18.2 min) was recovered from the appropriate fractions by lyophilization as a colorless solid: yield 0.8 mg (40%); mass spectrum (ESI) *m*/*z* 1201.1 $(M + H)^+$ (theoretical *m*/*z* 1201.2), 1199.3 $(M - H)^-$ (theoretical *m*/*z* 1199.3).

***N*-(9-Fluorenylmethoxycarbonyl)-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-tyrosine (22).** To a solution of 1.35 g (2.0 mmol) of β-D-lactose octaacetate (**21**) and 0.97 g (2.4 mmol) of *N*-Fmoc-(*S*)-tyrosine in 40 mL of dry dichloromethane was added 0.76 mL (0.85 g, 6.0 mmol) of $BF_3{\cdot}OEt_2$. The reaction mixture was stirred at room temperature under argon for 24 h, after which another 0.76 mL (0.85 g, 6.0 mmol) of $BF_3{\cdot}OEt_2$ was added and the reaction mixture was stirred for an additional 16 h. The reaction was quenched by addition of 50 mL of 0.5 N aqueous $NaHSO_4$ and extracted with two 50-mL portions of dichloromethane. The combined organic phase was dried ($MgSO_4$), filtered, and concentrated under diminished pressure. The residue was purified by chromatography on a silica gel column (30 × 3 cm); elution with a step gradient of methanol in dichloromethane (2−16%) afforded *N*-(9-fluorenylmethoxycarbonyl)-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-tyrosine (**22**), slightly contaminated with unreacted Fmoc-(*S*)-tyrosine as a colorless foam: yield 720 mg (35%); silica gel TLC R_f 0.76 (7:3 chloroform−methanol); $[\alpha]^{22}_D$ +30 (*c* 1.2, $CHCl_3$); ^{13}C NMR ($CDCl_3$) δ 20.15, 20.25, 20.33, 20.37, 20.43, 46.67, 60.48, 61.62, 66.30, 68.73, 70.28, 70.47, 70.60, 71.07, 72.28, 72.44, 75.81, 79.84, 95.16, 98.40, 100.72, 101.32, 116.17, 116.87, 119.70, 124.63, 126.69, 127.44, 130.09, 140.86, 143.23, 143.28, 155.48, 155.53, 168.82, 169.35, 169.49, 169.71, 169.79, 169.85, 169.89, 170.11, 170.28; mass spectrum (CI) *m*/*z* 1021.7 $(M)^+$ (theoretical *m*/*z* 1021.3).

***N*-(4-Pentenoyl)-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-tyrosine (23).** To 0.6 g (0.58 mmol) of *N*-(9-fluorenylmethoxycarbonyl)-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-tyrosine (**22**) in 4.8 mL of dichloromethane was added 1.2 mL of piperidine (1.03 g, 12.1 mmol) at 25 °C for 30 min. The reaction mixture was diluted with 20 mL of toluene, and the solvent was concentrated under diminished pressure (T < 30 °C). The residue was dissolved in 8 mL of DMF, and 174 mg (0.88 mmol) of 4-pentenoic acid succinimide ester was added, followed by a solution of 74 mg (0.88 mmol) of $NaHCO_3$ in 3 mL of water. The reaction mixture was stirred vigorously at room temperature for 24 h. The mixture was poured into 12 g of ice, acidified with 12 mL of 1 N $NaHSO_4$, and extracted with three 15-mL portions of dichloromethane. The combined organic phase was dried ($MgSO_4$), filtered, and concentrated under diminished pressure. The residue was purified by chromatography on a silica gel column (27 × 3 cm); elution with a gradient of methanol in dichloromethane (10−30%) afforded *N*-(4-pentenoyl)-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-tyrosine (**23**) as a colorless foam: yield 175 mg (34%); silica gel TLC R_f 0.6 (7:3 chloroform−methanol); $[\alpha]^{22}_D$ +13 (*c* 1.4, $CHCl_3$); ^1H NMR ($CDCl_3$) δ 1.96 (s, 3H), 2.04 (s, 3H), 2.06 (s, 6H), 2.08 (s, 6H), 2.15 (s, 3H), 2.29 (m, 4H), 3.04 (m, 1H), 3.12 (m, 1H), 3.75 (m, 1H), 3.90 (t, 2H, J = 9.0 Hz), 4.12 (m, 3H), 4.51 (m, 2H), 4.81 (br d, 1H), 4.93−5.18 (m, 6H), 5.27 (t, 1H, J = 8.9 Hz), 5.65 (m, 1H), 6.53 (br d, 1H), 6.87 (d, 2H, J = 8.5 Hz), 7.05 (d, 2H, J = 8.1 Hz); mass spectrum (FAB) *m*/*z* 904.2889 $(M + Na)^+$ ($C_{40}H_{51}NNaO_{21}$ requires *m*/*z* 904.2851).

***N*-(4-Pentenoyl)-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-tyrosine Cyanomethyl Ester (24).** To a solution containing 115 mg (0.13 mmol) of *N*-(4-pentenoyl)-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-tyrosine (**23**) and 56 mg (0.52 mmol) of anhydrous Na_2CO_3 in 4 mL of freshly distilled DMF was added 66 μL (78 μg, 1.04 mmol) of chloroacetonitrile. The reaction mixture was stirred at room temperature under argon for 16 h. The reaction mixture was poured into 4 g of ice, acidified with 4 mL of 1 N aqueous $NaHSO_4$, and extracted with three 10-mL portions of dichloromethane. The combined organic phase was dried ($MgSO_4$), filtered, and concentrated under diminished pressure. The residue was purified by chromatography on a silica gel column (20 × 3 cm); elution with a gradient of ethyl acetate in hexane (50−100%) afforded *N*-(4-pentenoyl)-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-tyrosine cyanomethyl ester (**24**) as a colorless foam: yield 28 mg (53%); silica gel TLC R_f 0.69 (ethyl acetate); $[\alpha]^{23}_D$ +5 (*c* 1.3, $CHCl_3$); ^1H NMR ($CDCl_3$) δ 1.94, 2.03, 2.04, 2.05, 2.06, 2.13 (6s, 21H), 2.27 (m, 4H), 3.06 (d, 1H, J = 6.2 Hz), 3.77 (m, 1H), 3.87 (m, 2H), 4.09 (m, 3H), 4.48 (m, 2H), 4.64 (d, 1H, J = 5.8 Hz), 4.80 (d, 1H, J = 5.8 Hz), 4.86 (d, 1H, J = 7.3 Hz), 4.92−5.16 (m, 8H), 5.25 (t, 1H, J = 8.8 Hz), 5.33 (d, 1H, J = 3.1 Hz), 5.67 (m, 1H), 5.93 (d, 1H, J = 7.7 Hz); ^{13}C NMR ($CDCl_3$) δ 19.76, 20.14, 20.27, 20.42, 28.80, 29.29, 34.92, 36.43, 48.50, 52.44, 60.42, 61.56, 66.19, 68.65, 70.30, 70.54, 70.99, 72.35, 72.41, 75.73, 98.26, 100.72, 113.42, 115.47, 116.90, 129.97, 136.24, 168.71, 169.24, 169.37, 169.70, 169.77, 169.95, 170.01, 171.72; mass spectrum (FAB) *m*/*z* 921.3195 $(M + H)^+$ ($C_{42}H_{53}N_2O_{21}$ requires *m*/*z* 921.3141).

***N*-(4-Pentenoyl)-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-tyrosine pdCpA Ester (25).** To a conical vial containing 17.2 mg (18.7 μmol) of *N*-(4-pentenoyl)-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-tyrosine cyanomethyl ester (**24**) was added a solution of 5.1 mg (3.7 μmol) of the tris-(tetrabutylammonium) salt of pdCpA[11i] in 68 μL of freshly distilled acetonitrile. The reaction mixture was stirred at room temperature for 24 h. A 2-μL aliquot of the reaction mixture was diluted with 58 μL of 1:1 acetonitrile−50 mM NH_4OAc, pH 4.5, and was analyzed by HPLC on a C_{18} reversed-phase column (250 × 10 mm). The column was washed with 1% → 63% acetonitrile in 50 mM NH_4OAc, pH 4.5, over a period of 45 min at a flow rate of 3.5 mL/min (monitoring at 260 nm). The remaining reaction mixture was diluted to a total volume of 600 μL of 1:1 acetonitrile−50 mM NH_4OAc, pH 4.5, and purified using the same C_{18} reversed-phase column. Dinucleotide derivative **25** (retention time 27−28 min) was recovered from the appropriate fractions by lyophilization as a colorless solid: yield 2.7 mg (48%); mass spectrum (FAB) *m*/*z* 1500.4049 $(M + H)^+$ ($C_{59}H_{76}N_9O_{33}P_2$ requires *m*/*z* 1500.4020).

***N*-(4-Pentenoyl)-*O*-[4-*O*-(β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-tyrosine pdCpA Ester (26).** To a conical vial containing 1.4 mg (0.9 μmol) of *N*-(4-pentenoyl)-*O*-[2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-β-D-glucopyranosyl]-(*S*)-ty-

rosine pdCpA ester **(25)** in 400 μL of dry methanol was added 100 μL of a 54% solution of HBF_4 in diethyl ether. The reaction mixture was stirred at room temperature for 3 h. A 10-μL aliquot of the reaction mixture was diluted with 40 μL of 1:1 acetonitrile−50 mM NH_4OAc, pH 4.5, and was analyzed by HPLC on a C_{18} reversed-phase column (250 × 10 mm). The column was washed with 1% → 63% acetonitrile in 50 mM NH_4OAc, pH 4.5, over a period of 45 min at a flow rate of 3.5 mL/min (monitoring at 260 nm). The remaining reaction mixture was diluted to a total volume of 300 μL of 1:1 acetonitrile−50 mM NH_4OAc, pH 4.5, and purified using the same C_{18} reversed-phase column. Dinucleotide derivative **26** (retention time 11−12 min) was recovered from the appropriate fractions by lyophilization as a colorless solid: yield 0.4 mg (35%); mass spectrum (FAB) *m*/*z* 1206.3260 ($M + H)^+$ ($C_{45}H_{62}N_9O_{26}P_2$ requires *m*/*z* 1206.3280).

Preparation of Glycosylated Aminoacyl-tRNA. The reactions were carried out in 100 μL (total volume) of 50 mM Na−Hepes buffer, pH 7.5, containing 0.5 mM ATP, 15 mM $MgCl_2$, 100 μg $tRNA_{CUA}$-C_{OH}, 1.0 A_{260} unit of N-protected (glycosylated) aminoacyl-pdCpA, 20−25% dimethyl sulfoxide, and 200 units of T4 RNA ligase. The reaction mixture was incubated at 37 °C for 1 h and then quenched by the addition of 10 μL of 3 M NaOAc, pH 5.0. The N-protected misacylated $tRNA_{CUA}$ was recovered by treatment with 4 vol of EtOH. Deprotection of the *N*-pentenoyl protecting group was carried out in a 100-μL reaction mixture containing 10 μg/μL of *N*-4-pentenoylaminoacyl-tRNA, 15% THF, and 8 mM iodine. After incubation at 25 °C for 50 min, the deprotected samples were precipitated with 0.3 M NaOAc, pH 5.2, and 2.5 vol of cold EtOH. The deprotected aminoacyl-tRNAs were dissolved in 1 mM KOAc to a final concentration of 3−5 μg/μL and stored in aliquots at −80 °C. Analysis of the prepared aminoacyl-tRNAs was carried out by polyacrylamide−urea gel electrophoresis at pH 5.0.[22]

***In Vitro* Protein Translation Reactions.** Incorporation of glycosylated amino acids into position 10 of DHFR was carried out using a rabbit reticulocyte lysate system, while incorporation into position 22 of DHFR and position 284 of firefly luciferase was performed in an *E. coli* S30 system. In the *in vitro* translation system using rabbit reticulocyte lysate, proteins were synthesized in 10−1000 μL of reaction mixture that contained the following per 100 μL: 70 μL of methionine-depleted, nuclease-treated rabbit reticulocyte lysate, 80 μCi of [^{35}S]-methionine (1000Ci/mmol), 2 μL of 1 mM amino acids mixture *without* methionine, 10 μg of the appropriate mRNA, and 35 μg (∼1.4 nmol) of deprotected aminoacyl-tRNAs, or unacylated tRNA as a control. The reaction mixture was incubated at 30 °C for 2 h. Translation reactions using an *E. coli* S-30 translation system were carried out in 10−1000 μL of reaction mixture that contained 40% S-30 extract, 40−100 μg/mL of plasmid DNA, 35 mM Tris−acetate, pH 7.0, 190 mM potassium glutamate, 30 mM ammonium acetate, 2 mM dithiothreitol, 10 mM magnesium acetate, 20 mM phosphoenolpyruvate, 3.5% PEG6000, 20 mM ATP and GTP, 5 mM CTP and UTP, 100 μM amino acids mixture *without* methionine, 50 μM methionine, 10 μCi/mL of [^{35}S]methionine, and 300 μg/mL of deprotected aminoacyl-tRNA. Translation was carried out for 60−90 min at 37 °C for DHFR and at 30 °C for luciferase synthesis. Aliquots from *in vitro* translation were analyzed by SDS−PAGE, and quantification of the band corresponding to the synthesized proteins was carried out by phosphorimager analysis. Suppression efficiency was calculated as the percent protein produced via nonsense suppression relative to protein production from the wild-type mRNA.

Purification of *in Vitro* Synthesized Proteins. *In vitro* translation mixtures (200 μL) containing ^{35}S-labeled protein were incubated with 1 μg of RNase A for 30 min at room temperature and then applied to a 200-μL DEAE Sepharose CL-6B column that had been equilibrated with 5 mM potassium phosphate buffer, pH 7.0. The column was washed sequentially with 5 vol of 5 mM potassium phosphate buffer, pH 7.0, containing 10 mM β-mercaptoethanol and 50 mM KCl. The remaining proteins were eluted from the column with 3 vol of 1 M KCl in 5 mM potassium phosphate buffer, pH 7.0. The fractions containing radioactivity were combined and applied to a 100-μL Ni-NTA agarose column equilibrated with 300 mM NaCl, 5 mM imidazole, and 100 μg/mL BSA in 50 mM sodium phosphate, pH 8.0. After the column was washed with 5 vol of the above equilibration buffer, the protein was eluted with 8 vol of 300 mM NaCl, 250 mM imidazole, and 10 μM BSA in 50 mM sodium phosphate, pH 8.0. The radioactive fractions were combined and dialyzed (Spectra/Por 2, MW cutoff 12−14 kDa) against 20 mM NaCl in 10 mM sodium phosphate, 10% glycerol, pH 7.0.

Luciferase Activity Assay. Luciferase activity was determined using a commercial luciferase assay system (Promega). After *in vitro* translation, 5-μL aliquots of the translation reaction mixture were used in 12.5% SDS−PAGE analysis to determine the yield and suppression efficiencies of the reaction, and 25-μL aliquots were added to 100 μL of the luciferase assay reagent to measure light emission using a Hitachi F2000 fluorescence spectrophotometer.

Binding of *in Vitro* Synthesized Glycoproteins to a Hydrazine Surface. Immediately before experiments, the buffer of purified proteins was changed to 0.1 M NaOAc, pH 5.5, by using a microconcentrator (Microcon YM-10, Millipore Corp.). Sodium periodate was then added slowly to the protein samples to a final concentration of 10 mM. After incubation at 25 °C for 4 h in a light-proof container, the reaction was quenched with 1/20 vol of a 1:100 dilution of ethylene glycol in deionized water. The activated protein samples were transferred to dialysis tubing (Spectra/Por 2, MW cutoff 12−14 kDa) and dialyzed against 1000 mL of 0.1 M sodium acetate, pH 5.5, for 6 h at 25 °C with a change of buffer every hour. The samples were concentrated with a microconcentrator (Microcon YM-10) until the radioactivity was >500 cpm/μL, and then 56 000 cpm of each protein sample was diluted in 0.1 M NaOAc, pH 5.5, to a final volume of 200 μL and applied to each well on a polystyrene carbohydrate binding plate having a hydrazine-containing surface (Corning Carbo-BIND, available from Sigma). The plate was incubated at 25 °C for 4 h, then the solution was decanted. The wells were washed sequentially with five 300-μL portions of the following: deionized water; 300 mM NaCl in 10 mM sodium phosphate, pH 7.0; 10% Tween 20 in 10 mM sodium phosphate, pH 7.0; 300 mM NaCl; and 10 μg/mL BSA in 10 mM sodium phosphate, pH 7.0. The washing was continued until the amount of ^{35}S washed away from the wells was less than 30 cpm/300 μL. The radioactivity of the ^{35}S-DHFR remaining in the well was determined using a Beckman LS-100C liquid scintillation counter.

Acknowledgment. This work was supported by NIH Research Grant CA77359, awarded by the National Cancer Institute.

Supporting Information Available: Figures illustrating the gel electrophoretic behavior of proteins containing glycosylated amino acids. This material is available free of charge via the Internet at http://pubs.acs.org.

JA067466N