Fmoc-Based Synthesis of Peptide-$^{\alpha}$Thioesters: Application to the Total Chemical Synthesis of a Glycoprotein by Native Chemical Ligation

Youngsook Shin,[†] Katharine A. Winans,[†] Bradley J. Backes,[†] Stephen B. H. Kent,[‡] Jonathan A. Ellman,[†] and Carolyn R. Bertozzi*,[†]

Contribution from the Department of Chemistry, University of California, Berkeley, California 94720, and Gryphon Sciences, 250 East Grand Avenue, Suite 90, South San Francisco, California 94080

Received August 9, 1999

Abstract: The technique of native chemical ligation has enabled the total chemical synthesis of proteins with molecular weights far in excess of those achievable by conventional stepwise solid-phase peptide synthesis. The method involves the condensation of two unprotected peptide segments, one bearing a C-terminal $^{\alpha}$thioester and the other an N-terminal cysteine residue, to afford a protein with a native amide linkage at the site of ligation. Here we report an extension of the native chemical ligation method to the total synthesis of a glycosylated protein, the antimicrobial *O*-linked glycoprotein diptericin. The major challenge in our synthesis was preparation of a 24-residue glycopeptide-$^{\alpha}$thioester segment, which was complicated by the incompatibility of glycosidic linkages with Boc chemistry and by the incompatibility of thioesters with Fmoc chemistry. The use of an alkanesulfonamide "safety-catch" linker circumvented this problem and permitted the solid-phase synthesis of the glycopeptide-$^{\alpha}$thioester using standard Fmoc chemistry protocols. Ligation of this thioester with a 58-residue glycopeptide bearing an N-terminal cysteine residue yielded the full-length glycoprotein with two sites of glycosylation. The fully deprotected diptericin glycoform was active in antimicrobial assays.

Introduction

Oligosaccharides covalently bound to proteins can profoundly affect protein folding, stability, immunogenicity, and biological activity.[1] To understand the relationship between oligosaccharide structure and glycoprotein function, and to facilitate the development of glycoprotein pharmaceutical agents, homogeneous glycoproteins bearing well-defined oligosaccharide structures must be obtained. Recombinant expression is a major source of glycoproteins for fundamental research and pharmaceutical applications, but suffers from complications intrinsic to oligosaccharide biosynthesis.[2] For example, differences in the efficiencies of various biosynthetic steps may result in a mixture of oligosaccharides of different length and composition at each glycosylation site.[3] Therefore a glycoprotein often exists as a heterogeneous mixture of "glycoforms", protein molecules which differ only in the structures of the bound oligosaccharides. Such glycoform mixtures are not ideal for structural and functional analysis.

Homogeneous and structurally defined glycoproteins can in principle be obtained by total chemical synthesis. Facilitated by advances in oligosaccharide chemistry, the synthesis of small glycopeptides bearing reasonably complex glycans has now perfected by many research groups.[4,5] The extension of these methods to larger glycoproteins remains encumbered by the limitations of traditional stepwise solid-phase peptide synthesis (SPPS); the linear assembly of peptides larger than about 60 amino acids is compromised by the accumulation of resin-bound byproducts.[6] Most full-length proteins and glycoproteins of biological interest exceed this length and are therefore not accessible by conventional methods.

To overcome this limitation of peptide synthesis, selective ligation strategies have been developed for the assembly of proteins with native amide bonds from unprotected peptide building blocks.[7] For example, native chemical ligation, a method developed by Kent and co-workers,[7b] makes use of the remarkable intramolecular acylating power of the thioester functionality. The first step of the process involves the chemose

* Corresponding author. Telephone: (510) 643−1682. Fax: (510) 643-2628. E-mail: bertozzi@cchem.berkeley.edu.

[†] University of California.

[‡] Gryphon Sciences.

(1) (a) Varki, A. *Glycobiology* **1993**, *3*, 97. (b) Bill, R. M.; Flitsch, S. L. *Chem. Biol.* **1996**, *3*, 145. (c) Lis, H.; Sharon, N. *Eur. J. Biochem.* **1993**, *218*, 1.

(2) Cumming, D. A. *Glycobiology* **1991**, *1*, 115.

(3) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683.

(4) Reviewed in Taylor, C. M. *Tetrahedron* **1998**, *54*, 11317.

(5) (a) Danishefsky, S. J.; Hu, S.; Cirillo, P. F.; Eckhardt, M.; Seeberger, P. H. *Chem. Eur. J.* **1997**, *3*, 1617. (b) Jansson, A. M.; Jensen, K. J.; Meldal, M.; Lomako, J.; Wieslawa, M.; Olsen, C. E.; Bock, K. *J. Chem. Soc., Perkin Trans. 1* **1996**, 1001. (c) Nakahara, Y.; Nakahara, Y.; Ogawa, T. *Carbohydr. Res.* **1996**, *292*, 71. (d) Hietter, H.; Schultz, M.; Kunz, H. *Synlett* **1996**, *12*, 1219. (e) Rio-Anneheim, S.; Paulsen, H.; Meldal, M.; Bock, K. *J. Chem. Soc., Perkin Trans 1* **1995**, 1071. (f) Rademann, J.; Schmidt, R. R. *Carbohydr. Res.* **1995**, *269*, 217. (g) Paulsen, H.; Bielfeldt, T.; Peters, S.; Meldal, M.; Bock, K. *Liebigs Ann. Chem.* **1994**, 369. (h) Seitz, O.; Wong, C.-H. *J. Am. Chem. Soc.* **1997**, *119*, 8766. (i) Gijsen, H. J. M.; Qiao, L.; Fitz, W.; Wong, C.-H. *Chem. Rev.* **1996**, *96*, 443. (j) Schwarz, J. B.; Kuduk, S. D.; Chen, X.-T.; Sames, D.; Glunz, P. W.; Danishefsky, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 2662.

(6) (a) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149. (b) Kent, S. B. H. *Annu. Rev. Biochem.* **1988**, *57*, 957. (c) Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180.

(7) (a) Kemp, D. S.; Carey, R. I. *J. Org. Chem.* **1993**, *58*, 2216. (b) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776. (c) Tam, J. P.; Lu, Y.-A.; Liu, C.-F.; Shao, J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 12485. (d) Muir, T. W.; Dawson, P. E.; Kent, S. B. H. *Methods Enzymol.* **1997**, *289*, 266. (e) Dawson, P. E. *Methods Enzymol.* **1997**, *287*, 34.

10.1021/ja992881j CCC: $18.00
Published on Web 12/07/1999

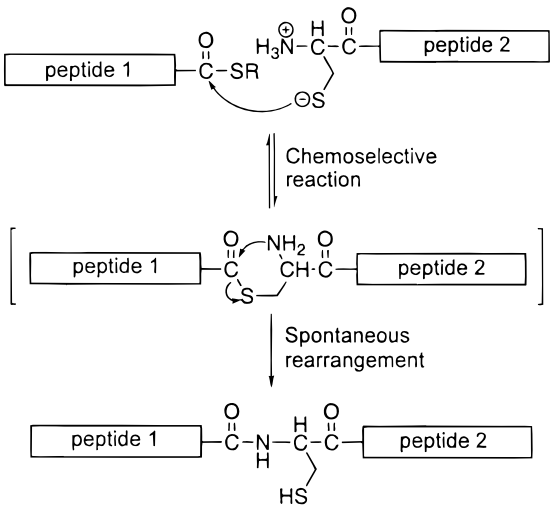

Figure 1. Synthesis of proteins by native chemical ligation of unprotected peptide segments in aqueous solution at neutral pH.

lective trans-thioesterification of an unprotected peptide-α-thioester with the sulfhydryl group of a second unprotected peptide segment containing an amino-terminal cysteine residue; this yields a thioester-linked intermediate as the initial covalent product (Figure 1). The intermediate undergoes spontaneous, rapid intramolecular reaction to form a native peptide bond at the ligation site. In this fashion, synthetically tractable peptides of 60 residues or less can be ligated to form large functional proteins.

The reactions that comprise native chemical ligation are mild, selective and compatible with the presence of glycans, suggesting that extension to glycoprotein synthesis would be feasible. However, the preparation of an unprotected glycopeptide-αthioester presents a significant challenge. Previously described syntheses of peptide-αthioesters utilize *tert*-butoxycarbonyl (Boc)-based protocols[8] involving acidic procedures such as repeated exposure of the peptide to TFA and cleavage/deprotection with HF.[6b] Such conditions damage glycosidic linkages and therefore preclude the application of Boc chemistry to glycopeptide synthesis. For this reason, glycopeptides are constructed primarily using 9-fluorenylmethoxycarbonyl (Fmoc)-based chemistry[4,5] which requires only a single cleavage/deprotection step with TFA.[9] Unfortunately, the repeated exposure of the resin-bound peptide to base during Fmoc deprotection at each step of the synthesis results in cleavage of thioesters.[10] Consequently, the major obstacle to glycoprotein synthesis by native chemical ligation is a lack of methods for the synthesis of glycopeptide-αthioesters.

We report here a new synthetic route for unprotected peptide-αthioesters using Fmoc-based SPPS and demonstrate its utility in the synthesis of a glycoprotein by native chemical ligation.

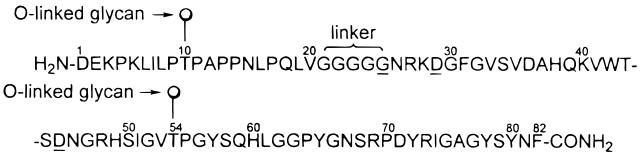

Figure 2. Amino acid sequence of diptericin, an 82-residue antibacterial glycoprotein. The two shaded circles (residues 10 and 54) in the sequence indicate the presence of an O-linked glycan. The C-terminus is posttranslationally modified by amidation. Sites of amino acid substitution in the synthetic target are underlined.

The method capitalizes on a recent modification of Kenner's sulfonamide "safety-catch" linker, reported by Backes and Ellman,[11] which we have adapted for Fmoc-based glycopeptide synthesis. The C-terminal residue of the peptide is attached to resin via an acid- and base-stable *N*-acyl sulfonamide linkage. After peptide synthesis, the sulfonamide is activated by cyanomethylation and then cleaved with thiol nucleophiles. As a target for chemical synthesis, we chose a variant of diptericin, an 82-residue antibacterial glycoprotein produced by insects in response to immunological challenge.[12] The glycoprotein is large enough to present a significant synthetic challenge, and furthermore, newly characterized antimicrobial agents such as diptericin are important therapeutic targets and potential leads for drug design. Native diptericin exists as a mixture of *O*-linked glycoforms; one of the simplest of these possesses single GalNAc residues at the two glycosylation sites Thr^{10} and Thr^{54} (Figure 2). We synthesized this diptericin glycoform by native chemical ligation of two glycopeptide fragments, each of which was generated by Fmoc-based SPPS. The glycoprotein obtained in this fashion was active in bacterial growth inhibition assays.

Results and Discussion

Synthetic Strategy. Diptericin consists of an N-terminal proline-rich domain (residues 1−20) and a C-terminal glycine-rich domain (residues 26−82) separated by five glycine residues (Figure 2). In previous work, we found that the main determinant of antibacterial activity lies in the C-terminal region, and residues Gly^{21}-Gly^{25} might therefore function as a flexible linker between the two domains.[13] We introduced three conservative amino acid substitutions into the diptericin sequence to facilitate chemical synthesis. First and foremost, native diptericin is devoid of cysteine residues which are obligate for native chemical ligation.[7b] Thus, we replaced Gly^{25} within the putative interdomain linker with Cys to afford a ligation site. Given the propensity of glycine residues to form flexible loops, our expectation was that this modification would not affect the global structure or function of the molecule. Moreover, the introduction of a ligation site between individual domains will allow the convergent synthesis of domain-specific variants for future structure/activity studies.

The second two modifications were motivated by two potentially problematic dipeptide motifs, Asp^{29}-Gly^{30} and Asp^{45}-Asn^{46}, which are prone to aspartimide formation via intramolecular cyclization during the synthesis.[14] This was avoided by

(8) (a) Hojo, H.; Aimoto, S. *Bull. Chem. Soc. Jpn.* **1991**, *64*, 111. (b) Dawson, P. E.; Churchill, M. J.; Ghadiri, M. R.; Kent, S. B. H. *J. Am. Chem. Soc.* **1997**, *119*, 4325. (c) Hackeng, T. M.; Mounier, C. M.; Bon, C.; Dawson, P. E.; Griffin, J. H.; Kent, S. B. H. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 7845. (d) Camarero, J. A.; Cotton, G. J.; Adeva, A.; Muir, T. W. *J. Pept. Res.* **1998**, *51*, 303.

(9) (a) Fields, G. B.; Noble, R. L. *Int. J. Pept. Protein Res.* **1990**, *35*, 161. (b) Wellings, D. A.; Atherton, E. *Methods Enzymol.* **1997**, *289*, 44. (c) Chang, C.-D.; Meienhofer, J. *Int. J. Pept. Protein Res.* **1978**, *11*, 246. (d) Atherton, E.; Fox, H.; Harkiss, D.; Logan, C. J.; Sheppard, R. C.; Williams, B. J. *J. Chem. Soc., Chem. Commun.* **1978**, 537. (e) Atherton, E.; Fox, H.; Harkiss, D.; Sheppard, R. C. *J. Chem. Soc., Chem. Commun.* **1978**, 539. (f) Chang, C.-D.; Felix, A. M.; Jimenez, M. H.; Meienhofer, J. *Int. J. Pept. Protein Res.* **1980**, *15*, 485.

(10) Our unpublished results.

(11) (a) Backes, B. J.; Ellman, J. A. *J. Org. Chem.* **1999**, *64*, 2322. (b) Backes, B. J.; Virgilio, A. A.; Ellman, J. A. *J. Am. Chem. Soc.* **1996**, *118*, 3055.

(12) Bulet, P.; Hegy, G.; Lambert, J.; Dorsselaer, A. V.; Hoffmann, J. A.; Hetru, C. *Biochemistry* **1995**, *34*, 7394.

(13) Winans, K. A.; King, D. S.; Rao, V. R.; Bertozzi, C. R. *Biochemistry* **1999**, *38*, 11700.

(14) (a) Quibell, M.; Owen, D.; Packman, L. C.; Johnson, T. *J. Chem. Soc., Chem. Commun.* **1994**, 2343. (b) Johnson, T.; Quibell, M.; Owen, D.; Sheppard, R. C. *J. Chem. Soc., Chem. Commun.* **1993**, *4*, 369.

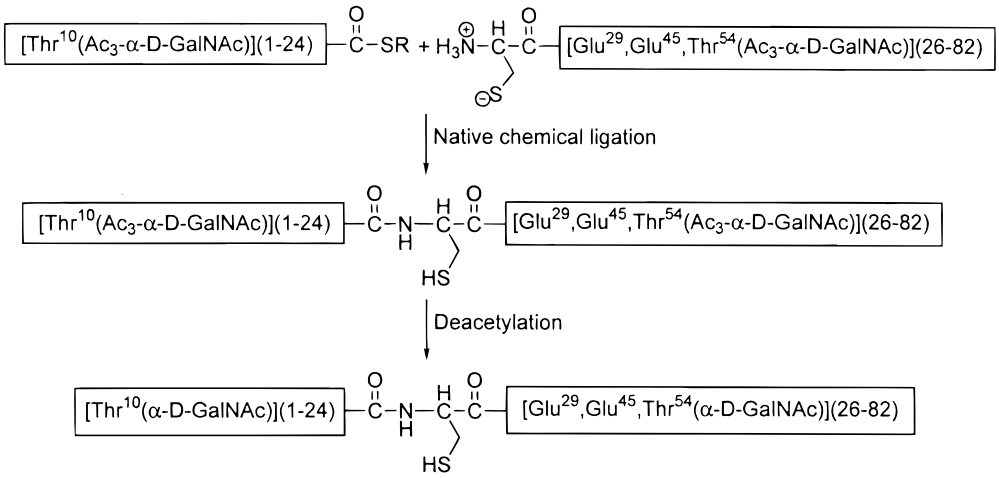

Figure 3. Strategy for native chemical ligation in the synthesis of a diptericin variant.

substituting Asp[29] and Asp[45] with glutamic acid residues, a modification that we believe will not adversely affect bioactivity.[13]

The GalNAc residues at Thr[10] and Thr[54] were installed by incorporating the glycosylated amino acid N^{α}-Fmoc-Thr(Ac_3-α-D-GalNAc)[15] during solid-phase synthesis. The acetyl-protecting groups on the GalNAc residues were removed after native chemical ligation by treatment with aqueous hydrazine. The two segments for native chemical ligation were (1) the glycopeptide-$^{\alpha}$thioester comprising residues 1−24 bearing Ac_3-α-D-GalNAc at Thr[10] and (2) the glycopeptide comprising residues 25−82 bearing Ac_3-α-D-GalNAc at Thr[54] and the substituted residues Cys[25], Glu[29], and Glu[45]. The general approach is illustrated schematically in Figure 3.

Synthesis of [Thr[10](Ac_3-α-D-GalNAc)]diptericin(1−24)-$^{\alpha}$COSBn. We synthesized the benzyl glycopeptide-$^{\alpha}$thioester ((1−24)-$^{\alpha}$COSBn) using the alkanesulfonamide "safety-catch" linker[11] as depicted in Figure 4. Although benzyl thioesters are relatively unreactive in transthioesterification reactions compared to more commonly used phenyl thioesters, their resistance to hydrolysis promotes easier purification and handling.[7e,8b] The benzyl thioester can be converted to a more reactive phenyl thioester in situ during native chemical ligation by simply including excess thiophenol in the reaction mixture.[8b]

The first amino acid, Fmoc-Gly[24]-OH, was loaded onto commercially available 4-sulfamylbutyryl AM resin (**1**, Figure 4) by two successive coupling reactions using PyBOP and DIEA.[16] A loading yield of 55% was obtained after only one coupling reaction, and this was increased to 97% after the second. All subsequent amino acids were coupled by stepwise solid-phase methods using automated DCC-mediated HOBt ester activation protocols[9] except for the final amino acid residue (Asp), which was introduced as Boc-Asp(tBu)-OH. The Boc group protected the N-terminal amine from unwanted nucleophilic attack on the activated sulfonamide linker; unlike an Fmoc group, the Boc group could be removed concurrently with amino acid side-chain-protecting groups at the end of the synthesis, without harming the $^{\alpha}$thioester. During SPPS, the glycosylated amino acid building block N^{α}-Fmoc-Thr(Ac_3-α-D-GalNAc) (Figure 4) was manually coupled using DIC in the presence of HOBt.

Alkylation of the sulfonamide in support-bound peptide **3** using iodoacetonitrile was performed according to the procedure of Backes et al.[11] to afford the support-bound *N*,*N*-cyanomethylacylalkanesulfonamide active ester **4**. The peptide was liberated from the activated linker with benzyl mercaptan (10% v/v in THF) to provide the peptide-$^{\alpha}$thioester **5**. The addition of a base such as triethylamine to the cleavage reaction was not necessary and did not affect the final yield of thioester.

The amino acid side-chain-protecting groups and N-terminal Boc group in thioester **5** were removed using Reagent K[17] to give unprotected crude glycopeptide-$^{\alpha}$thioester **6**. After the glycopeptide was purified by reversed-phase HPLC, the mass of the final product was confirmed by ESI-MS (2843 Da; 2844.15 Da calculated, average isotope composition). The purified glycopeptide-$^{\alpha}$thioester was obtained in a final isolated yield of 21% on a 0.1 mmol scale.

Synthesis of [Cys[25],Glu[29],Glu[45],Thr[54](Ac_3-α-D-GalNAc)]-Diptericin(25−82). The C-terminal segment of diptericin was synthesized on Rink amide resin using similar stepwise solid-phase methods. After deprotection and cleavage from the solid support using Reagent K, the crude unprotected glycopeptide bearing a C-terminal amide was obtained. The glycopeptide was purified by reversed-phase HPLC, and the mass was confirmed by ESI-MS (6703 Da; 6704.28 Da calculated, average isotope composition). A major byproduct of the synthesis was a glycopeptide missing the three N-terminal residues (Cys[25]Asn[26]-Arg[27]). This "CNR-deletion peptide" copurified with the full-length glycopeptide and was inseparable by HPLC. Since the deletion peptide lacks an N-terminal Cys residue, it was expected to be inert to native chemical ligation. Thus, the mixture was utilized in the subsequent reaction and the unreacted deletion peptide was separated from the final product after the ligation step.

Synthesis of [Thr[10](α-D-GalNAc),Cys[25],Glu[29],Glu[45],Thr[54]-(α-D-GalNAc)]Diptericin(1−82) by native chemical ligation. The ligation of [Thr[10](Ac_3-α-D-GalNAc)]diptericin(1−24)-$^{\alpha}$COSBn (0.5 μmol) and [Cys[25],Glu[29],Glu[45],Thr[54](Ac_3-α-D-GalNAc)]diptericin(25−82) (0.75 μmol) was carried out in the presence of 6 M Gn·HCl, 0.1 M sodium phosphate (pH 7.5) containing 4% thiophenol. Excess thiophenol was used to promote the formation of the highly reactive phenyl thioester glycopeptide and to prevent the cysteine residues from forming disulfide bonds during ligation.[8b,7d] The reaction was monitored by analytical reversed-phase HPLC (Figure 5). After the reaction

(15) N^{α}-Fmoc-Thr(Ac_3-α-D-GalNAc) is available from commercial sources but can also be readily prepared on a multigram scale (ref 13 and citations therein).

(16) The loading yields were determined by piperidine treatment of small aliquots of resin followed by spectrophotometric analysis of the piperidine−dibenzylfulvene adduct in solution (Bunin, B. A. *The Combinatorial Index*; Academic Press: San Diego, 1998; p 213).

(17) King, D. S.; Fields, C. G.; Fields, G. B. *Int. J. Pept. Protein Res.* **1990**, *36*, 255.

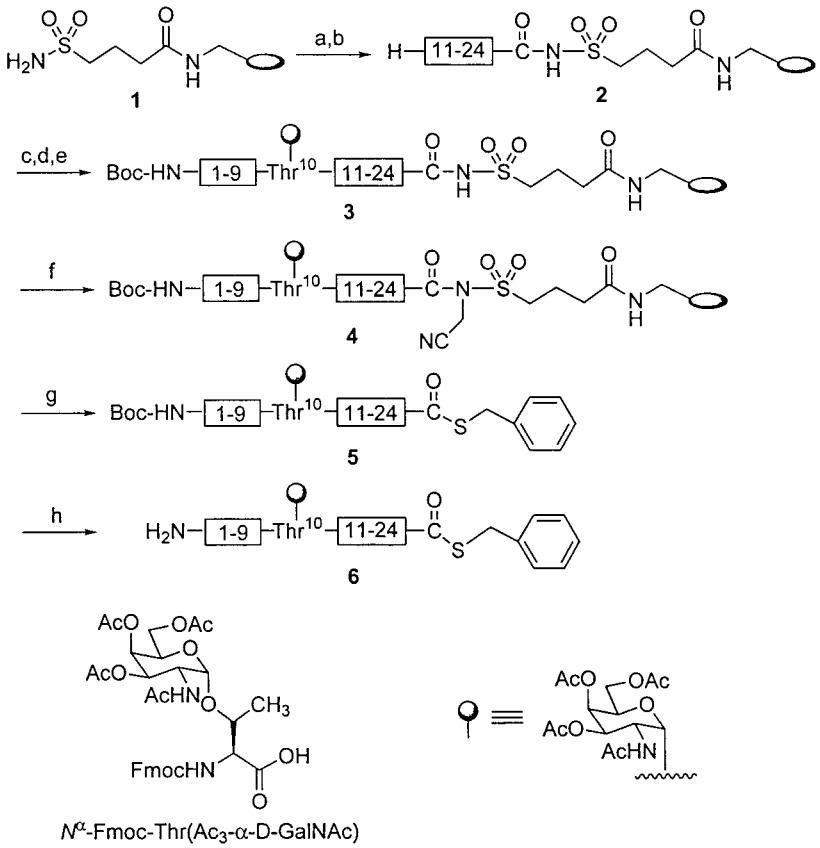

Figure 4. Synthesis of the glycopeptide-αthioester using the alkanesulfonamide "safety-catch" linker. (a) Fmoc-Gly-OH (4 equiv), PyBOP (3 equiv), DIEA (9 equiv), DMF, −20 °C, 8 h followed by room temperature overnight, and then repetition of the procedure; (b) SPPS using N^{α}-Fmoc-amino acids (sequence *11*−*23)*, coupling with DCC/HOBt in NMP; (c) N^{α}-Fmoc-Thr(Ac_3-α-D-GalNAc) (5 equiv), DIC (10 equiv), HOBt (10 equiv), DMF (30 min premix), 30 min; (d) SPPS using N^{α}-Fmoc-amino acids (sequence 2−9), coupling with DCC/HOBt in NMP; (e) N^{α}-Boc-Asp(OtBu)-OH (5 equiv), DIC (10 equiv), HOBt (10 equiv), DMF, 30 min; (f) ICH_2CN, DIEA, NMP, 24 h; (g) BnSH, THF, 24 h; (h) Reagent K (TFA (82.5%), phenol (5%), H_2O (5%), thioanisole (5%), ethanedithiol (2.5%)), 4 h.

proceeded for 1 h at room temperature, a significant amount of ligation product was observed, and the reaction was essentially complete after 18 h. The ligation product was purified by semipreparative reversed-phase HPLC, and the mass was confirmed by ESI-MS (9424 Da; 9424.22 Da calculated, average isotope composition) (Figure 6). The purified glycoprotein was obtained in a yield of 55% based on the limiting glycopeptide-αthioester component.

Finally, the acetate-protecting groups on the carbohydrate substituents were removed by treatment with 5% aqueous hydrazine containing excess DTT.[18] The fully deprotected diptericin variant was purified in an isolated yield of 53%, and the mass was confirmed by ESI-MS (9172 Da; 9171.99 Da calculated, average isotope composition).

Bioactivity of Synthetic Diptericin Variant. To assess the antibacterial potency of the cysteine-containing diptericin variant, we monitored its ability to inhibit the growth of *Escherichia coli* D22 cells using protocol we have recently reported.[13] The synthetic glycoprotein blocked bacterial growth with an IC_{50} of 2.70 ± 0.30 μM, a value within 10-fold of that measured previously by Winans et al. for a more native version of diptericin containing a solitary Asp→Glu mutation at position 45.[13] Therefore, introduction of the Cys residue at the ligation site within the putative interdomain linker and addition of a second Asp→Glu mutation do not drastically affect the bioactivity of diptericin. This convergent synthetic strategy for domain assembly promises to be useful for detailed functional analysis of the glycoprotein.

(18) DTT was essential for preventing disulfide-bound dimer formation.

In summary, we have demonstrated the total chemical synthesis of a biologically active glycoprotein by native chemical ligation. This general synthetic approach should allow access to unprecedented quantities of homogeneous glycoproteins and thereby facilitate biological investigations and pharmaceutical development. Ongoing work will further define the scope of the method as applied to glycoproteins from the both the *O*-linked and *N*-linked families. The construction of a large variety of proteins by means of this[19] and related[20] protein ligation strategies suggests the potential for widespread implementation. The technique of native chemical ligation therefore augments the growing arsenal of tools for glycoprotein research.

Experimental Section

Materials and Methods. All Fmoc-amino acids, Boc-Asp(OtBu)-OH, 4-sulfamylbutyryl AM resin, Rink amide resin, 2-(1*H*-benzotriazol-1-yl)-1,1,3,3,-tetramethyluronium hexafluorophosphate (HBTU), 1-hydroxybenzotriazole (HOBt), and *N*,*N*′-dicyclohexylcarboimide (DCC) were purchased from Novabiochem. (1*H*-Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP), piperidine, *N*,*N*-diisopropylethylamine (DIEA), 1,3-diisopropylcarbodiimide (DIC), benzyl mercaptan, thiophenol, iodoacetonitrile, *tert*-butyl methyl ether, phenol, thioanisole, and ethanedithiol (EDT) were purchased from Aldrich Chemical Co. Dichloromethane (DCM), methanol, *N*,*N*-dimethylformamide (DMF), and trifluoroacetic acid (TFA) were purchased from J. T. Baker. 1-Methyl-2-pyrrolidinone (NMP) and HPLC-grade acetonitrile were purchased from EM Science. All other

(19) Wilken, J.; Kent, S. B. H. *Curr. Opin. Biotechnol.* **1998**, *9*, 412.
(20) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705.

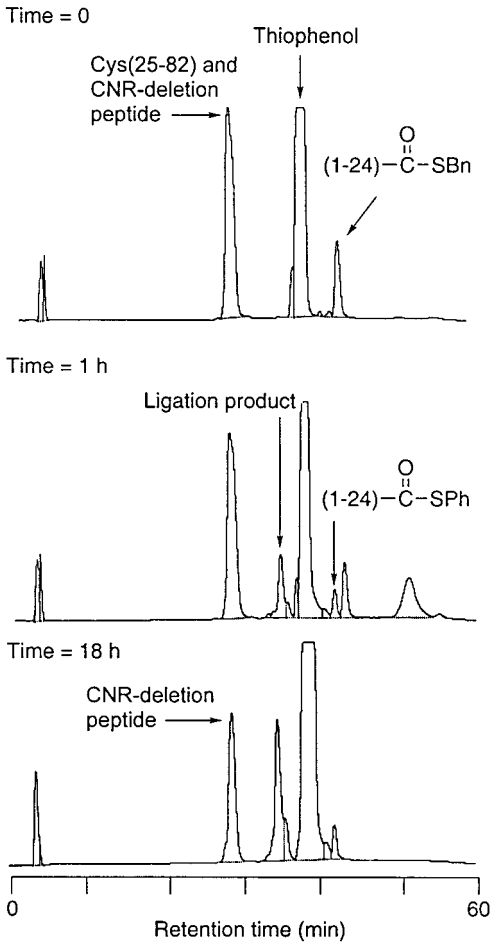

Figure 5. Time course of the ligation reaction. The starting materials were dissolved in 0.1 M sodium phosphate buffer containing 6 M Gn·HCl, pH 7.5, to which 4% thiophenol had been added. The ligation reaction was monitored by reversed-phase HPLC eluting with a gradient of 10−50% B over 50 min, at a flow rate of 1 mL/min. After 18 h, the glycopeptide-$^{\alpha}$thioester was mostly consumed. Conversion of the C-terminal segment to product revealed an underlying peak corresponding to the CNR-deletion peptide which was characterized by ESI-MS.

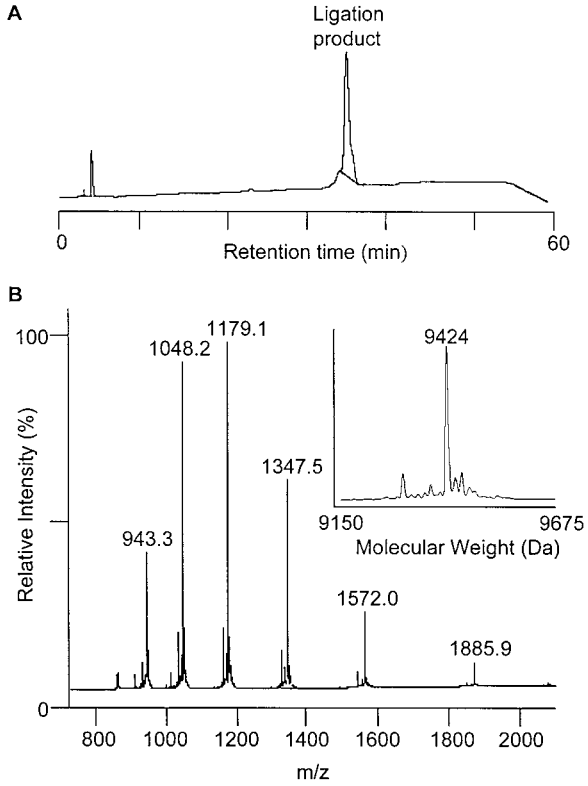

Figure 6. (A) Reversed-phase HPLC analysis of the purified ligation product. (B) ESI-MS of the purified ligation product.

chemical reagents were obtained from commercial suppliers and used without further purification.

Reversed-phase HPLC was performed on a Rainin Dynamax SD-200 HPLC system with 214-nm UV detection, using a Microsorb C-18 analytical column (4.6 × 250 mm) at a flow rate of 1 mL/min, a sempreparative column (10 × 250 mm) at a flow rate of 4 mL/min, or a preparative column (25 × 250 mm) at a flow rate 20 mL/min. All runs used linear gradients of 10−50% buffer B in A (A = water containing 0.1% TFA, B = acetonitrile containing 0.1% TFA) over 50 min. Electrospray ionization mass spectrometry (ESI-MS) was performed at the University of California, Berkeley, on a Micromass triple-quadrupole electrospray mass spectrometer (Quattro). Observed masses were derived from the experimental *m*/*z* values for all observed protonation states of a molecular species, using the software MassLynx (Micromass). ESI-MS values are accurate within ±1 Da for species with molecular masses up to 6 kDa, and ±2 Da for species with molecular masses above 9 kDa. Calculated masses were based on average isotope composition and were derived using the program SynthAssist Software (Applied Biosystems).

General Methods for Solid-Phase Peptide Synthesis (SPPS). Peptides were synthesized by established automated protocols on a Perkin-Elmer ABI 431A peptide synthesizer (user-devised cycles) using N^{α}-Fmoc-protected amino acids and DCC-mediated HOBt ester activation in NMP, except for the glycosylated amino acid N^{α}-Fmoc-Thr-(Ac_3-α-D-GalNAc) and Boc-Asp(OtBu)-OH, both of which were manually coupled. Five equivalents of *N*-protected amino acid was activated for at least 30 min using 10 equiv each of DIC and HOBt in DMF. This solution was added to resin and shaken for 30 min. Samples (∼5 mg) of resin-bound peptide were removed after each coupling step for determination of residual free α-amino groups by the quantitative ninhydrin method or, for proline residues, the chloranil test. Peptide cleavage/deprotection was accomplished with Reagent K (82.5% TFA: 5% phenol: 5% H_2O: 5% thioanisole: 2.5% EDT) for 4 h at room temperature. Soluble crude peptide products were precipitated, washed with *tert*-butyl methyl ether, and then dissolved in 50% aqueous acetonitrile with 0.1% TFA and lyophilized.

[Thr^{10}(Ac_3-α-D-GalNAc)]Diptericin(1−24)-$^{\alpha}$COSBn. Resin Loading and SPPS. To a 50 mL round-bottom flask were added 4-sulfamylbutyryl AM resin **1** (110 mg, 0.10 mmol), DMF (2 mL), DIEA (157 μL, 0.9 mmol), and Fmoc-Gly-OH (119 mg, 0.400 mmol). The reaction mixture was stirred for 10 min followed by cooling to −20 °C. After 20 min, PyBOP (156 mg, 0.300 mmol) was added to the reaction mixture as a solid. The reaction mixture was stirred at −20 °C for 8 h and then warmed to room temperature overnight. The resin was filtered, washed with DMF and DCM, and resubjected to the same coupling conditions. Samples (∼5 mg) of Fmoc-Gly sulfonamide resin were removed after each loading step for Fmoc quantitation. The glycopeptide was elongated on Fmoc-Gly sulfonamide resin as described in the general methods section above.

Activation and Cleavage. The resin-bound peptide was prepared for activation and cleavage by the initial addition of NMP (4 mL) and DIEA (200 μL, 1.10 mmol). Iodoacetonitrile (180 μL, 2.5 mmol), prefiltered through a plug of basic alumina, was added to the reaction mixture with the exclusion of light. The resin was agitated for 24 h, filtered, washed with NMP (5 × 5 mL, 10 min/wash), DCM (3 × 5 mL), and THF (3 × 5 mL), and then transferred to a 50 mL round-bottom flask. To the resin-containing flask was added THF (3.6 mL) and benzyl mercaptan (400 μL, 3.4 mmol), and the reaction mixture was stirred for 24 h. The resin was separated from the solution by filtration and washed three times with 5 mL portions of DCM. The combined filtrates were collected and concentrated.

Side-Chain Deprotection. The side-chain protecting groups on the crude peptide were removed with 4 mL of Reagent K (82.5% TFA:

5% phenol: 5% H_2O: 5% thioanisole: 2.5% EDT) for 4 h at room temperature. The unprotected crude peptide was precipitated and washed with *tert*-butyl methyl ether, and then dissolved in 50% aqueous acetonitrile with 0.1% TFA and lyophilized. The crude peptide was purified by preparative C-18 reversed-phase HPLC and lyophilized to afford 59 mg of the desired product (21% yield based on a resin loading capacity of 0.1 mmol). The purified peptide was characterized by analytical C-18 reversed-phase HPLC and ESI-MS: observed, 2843 Da; calculated average isotope composition for $C_{131}H_{207}N_{29}O_{39}S$, 2844.15 Da.

[Cys^{25},Glu^{29},Glu^{45},Thr^{54}(Ac$_3$-α-D-GalNAc)]diptericin(25−82). SPPS was performed on Rink Amide resin (0.1 mmol scale) as described above. The resulting crude glycopeptide was purified by preparative C-18 reversed-phase HPLC and lyophilized. Analysis by ESI-MS revealed the presence of two glycopeptides, the desired product and a deletion peptide lacking the three C-terminal amino acids (CNR). These were inseparable by HPLC, and thus the mixture was carried on to the following reaction without further attempts at purification. ESI-MS for the desired product: observed, 6703 Da; calculated average isotope composition for $C_{294}H_{431}N_{87}O_{93}S$, 6704.28 Da.

[Thr^{10}(α-D-GalNAc),Cys^{25},Glu^{29},Glu^{45},Thr^{54}(α-D-GalNAc)] Diptericin(1−82). The glycopeptide−$^{\alpha}$thioester (1.4 mg, 0.50 μmol) and C-terminal segment (5.0 mg, 0.75 μmol) were dissolved in 6 M Gn·HCl and 100 mM sodium phosphate (pH 7.5) to give a final concentration of 1 mM and 1.5 mM, respectively. Thiophenol (4% final v/v) was added, and the mixture was briefly stirred to saturate the ligation buffer with thiophenol. The reaction was incubated for 24 h and monitored by analytical C-18 reversed-phase HPLC. Following ligation, the product was purified by semipreparative reversed-phase HPLC and lyophilized to afford 2.6 mg (55% yield based on limiting glycopeptide−$^{\alpha}$thioester) of the full-length glycoprotein. The product was characterized by analytical C-18 reversed-phase HPLC and ESI-MS: observed, 9424 Da; calculated average isotope composition for $C_{418}H_{630}N_{116}O_{132}S$, 9424.22 Da.

Carbohydrate Deacetylation. The purified glycoprotein from the previous reaction (2.6 mg, 0.27 μmol) was treated with 1.5 mL of 5% aqueous hydrazine containing an excess of DTT and the reaction was monitored by HPLC. After standing for 30 min at room temperature, the crude product was purified on an analytical C-18 reversed-phase column, to afford after lyophilization 1.33 mg (53%) of deprotected diptericin. ESI-MS: observed, 9172 Da; calculated average isotope composition for $C_{406}H_{618}N_{116}O_{126}S$, 9171.99 Da.

Bacterial Growth Inhibition Assay. Growth inhibition assays were performed in 96-well microtiter plates (Corning Costar, half area wells) with a final volume of 55 μL per well; bacterial growth was monitored by the change in $OD_{415\ nm}$, which we have found is directly correlated with the density of viable cells.[13] The 55-μL volume comprised 50 μL of bacterial culture (described below) added to 5 μL of serially diluted diptericin variant. Glycoprotein concentrations were determined by taking dry weight measurements and assuming that 80% of that weight was peptide, and the remainder residual water and TFA salt; stock solutions were made by dissolving the peptides in water. The bacterial culture was prepared by growing *E. coli* D22 cells in Luria-Bertani (LB) media containing streptomycin (50 μg/mL) to mid-logarithmic phase, then diluting with LB/streptomycin to $OD_{415\ nm} = 0.01$ (absorbance of a 50-μL sample in one well of a 96-well microtiter plate). Plates were incubated for 24 h at 25 °C with periodic shaking. Growth was monitored by measuring $OD_{415\ nm}$ on a BioRad 550 microtiterplate reader. ΔOD values were calculated by subtracting $OD_{415\ nm}$ at $t = 0$ h from $OD_{415\ nm}$ at $t = 24$ h. Curves were fit using a logistic equation as described in Delean et al.,[21] and IC_{50} values were determined from inhibition curves. Four replicate experiments were performed, and error margins represent the standard deviation.

Acknowledgment. The authors thank Dr. Paolo Botti and Vikram Rao for numerous helpful discussions. The authors acknowledge funding from the Arnold and Mabel Beckman Foundation, the W. M. Keck Foundation, the Camille and Henry Dreyfus Foundation, and the Burroughs Wellcome Fund. K.A.W. was supported by a graduate fellowship from Glaxo Wellcome. This research was supported by grants to C.R.B from the National Science Foundation (CAREER Award CHE-9734439) and the University of California BioSTAR Project (S97-73) in partnership with Gryphon Sciences.

JA992881J

(21) DeLean, A.; Munson, P. J.; Rodbard, D. *Am. J. Physiol.* **1978**, *235*, E97.