Chemical Synthesis of Lymphotactin: A Glycosylated Chemokine with a C-Terminal Mucin-Like Domain

Lisa A. Marcaurelle,[a] Laura S. Mizoue,[b] Jill Wilken,[c] Lisa Oldham,[d] Stephen B. H. Kent,[c] Tracy M. Handel,[b] and Carolyn R. Bertozzi*[a]

Abstract: The synthesis of a 93-residue chemokine, lymphotactin, containing eight sites of *O*-linked glycosylation, was achieved using the technique of native chemical ligation. A single GalNAc residue was incorporated at each glycosylation site using standard Fmoc-chemistry to achieve the first total synthesis of a mucin-type glycoprotein. Using this approach quantities of homogeneous material were obtained for structural and functional analysis.

Keywords: chemical ligation · chemokines · glycoproteins · mucin · solid-phase synthesis

Introduction

Mucin-type glycosylation, in which oligosaccharides are linked via an α-GalNAc residue to either serine or threonine (Figure 1), is the most common type of *O*-linked glycosylation found in mammals and other eukaryotes. Typically, this type of glycosylation is associated with glycoproteins known as "mucins" which contain dense clusters of *O*-linked glycans and can be up to 80% carbohydrate by weight.[1] However, mucin-type oligosaccharides are not restricted to mucins. They can also be found in small regions of proteins which are otherwise not heavily glycosylated such as CD45, the low density lipoprotein (LDL) receptor and the decay accelerating factor.[2–4] Although the precise roles of these glycosylated regions are unclear, they have now been found in a number of other proteins, including certain members of the chemokine family.

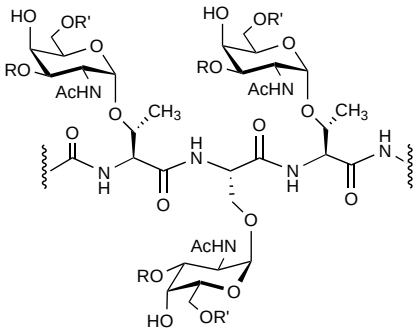


Figure 1. General structure of a mucin-type glycoprotein: R, R' = carbohydrate residues.

Chemokines are a large family of ≈8–12 kD proteins that mediate key steps in infection and inflammation by binding to seven-transmembrane receptors on target cells. Lymphotactin (Lptn) is a 93-residue chemokine that acts as a chemoattractant for both T-cells and natural killer cells.[5] Examination of the amino acid sequence of Lptn (Figure 2) reveals a cluster of serine and threonine residues located at the

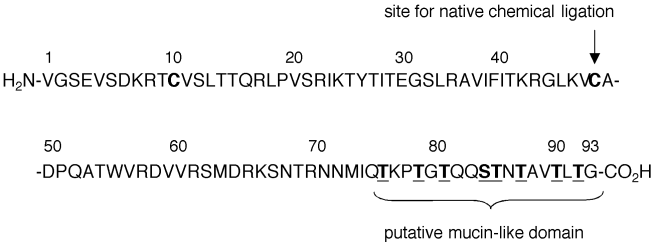


Figure 2. Amino acid sequence of human lymphotactin. Potential sites of *O*-linked glycosylation are underlined.

[a] Prof. C. R. Bertozzi, L. A. Marcaurelle
Departments of Chemistry and Molecular and Cellular Biology
Howard Hughes Medical Institute
University of California, Berkeley, CA 94720 (USA)
Fax: (+1) 510-643-2628
E-mail: bertozzi@cchem.berkeley.edu

[b] Dr. L. S. Mizoue, Prof. T. M. Handel
Department of Molecular and Cellular Biology
University of California, Berkeley, CA 94720 (USA)

[c] Dr. J. Wilken, Dr. S. B. H. Kent
Gryphon Sciences
South San Francisco, CA 94080 (USA)

[d] L. Oldham
DNAX Research Institute of Molecular and Cellular Biology
Palo Alto, CA 94304 (USA)

Supporting information for this article is available on the WWW under http://www.wiley-vch.de/home/chemistry/ or from the author. Supporting information contains one-dimensional ^{1}H NMR spectra and two-dimensional HSQC spectra of unglycosylated lymphotactin.

  0947-6539/01/0705-1129 $ 17.50+.50/0

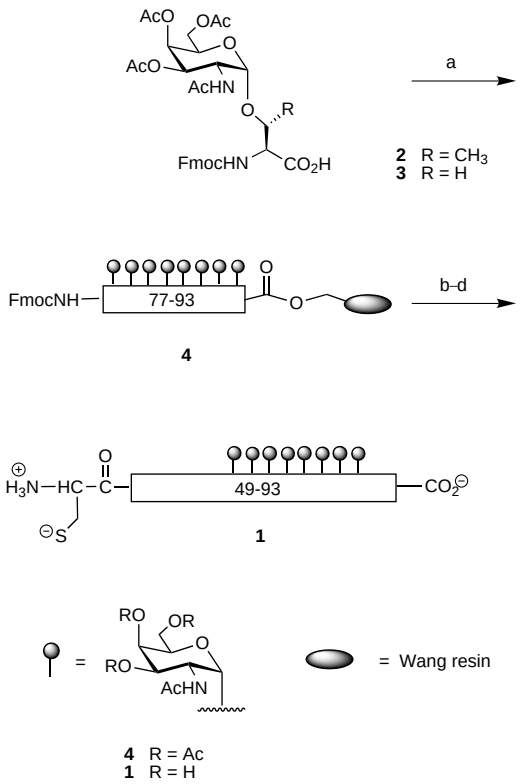
C-terminus, suggestive of the presence of a mucin-like domain. Indeed, Lptn has been shown to be *O*-glycosylated in vivo, although the structural and functional significance of this glycosylation is yet unknown.[5a]

While chemokines generally adopt highly ordered structures, NMR studies of unglycosylated Lptn suggest that the protein is only partially folded.[6] This is most likely due to the absence of one of the two intrachain disulfide bonds that are conserved among all other chemokines.[5] Since glycosylation can have a profound effect on the structure of a protein,[7] we reasoned that the addition of *O*-linked glycans to the putative mucin-like C-terminus of Lptn might lead to a more ordered structure. Other chemokines are also known to be influenced by glycosylation.[8, 9] In particular, fractalkine, a membrane-bound chemokine, contains a long mucin-like stalk which extends its receptor binding domain away from the cell surface.[8] The presence of a membrane-tethered mucin domain enables it to function not only as a chemoattractant, but also as a cell adhesion molecule.[10] As a first step in evaluating the structural and functional significance of the mucin-like domain of Lptn, we undertook the total chemical synthesis of the full-length glycoprotein, which we report herein.

For our initial studies we chose to incorporate only a single GalNAc residue at each of the eight potential glycosylation sites which are clustered near the C-terminus (Figure 2). It has been demonstrated in the context of small synthetic mucins that clusters of α-linked GalNAc can lead to the formation of highly stable structures similar to those induced by higher order glycans.[7b] In order to synthesize this Lptn glycoform we took advantage of the well-known technique of native chemical ligation.[11] This method, which is compatible with the presence of *O*-linked glycans,[12] required the synthesis of a 47-residue peptide α-thioester (1–47), and a 46-residue glycopeptide (48–93) with an N-terminal cysteine. Due to the acid lability of glycosidic linkages, *O*-linked glycopeptides in general must be synthesized using Fmoc-chemistry.[13] Thus, Fmoc-based SPPS was employed for the synthesis of the glycopeptide fragment. Since the amino terminus of Lptn is not predicted to be modified by glycosylation,[5a] we were able to employ established Boc-based methods to generate the peptide α-thioester.[14]

Results and Discussion

The synthesis of the glycopeptide fragment **1** on Wang resin is depicted in Scheme 1. The eight GalNAc residues were installed at Thr^{76}, Thr^{79}, Thr^{81}, Ser^{84}, Thr^{85}, Thr^{87}, Thr^{90}, and Thr^{92} using glycosyl amino acids **2** and **3**[15] as building blocks. Peptide couplings were performed manually with O-benzotriazol-1-yl-*N*,*N*′,*N*″,*N*‴-tetramethyluronium (HBTU)/1-hydroxybenzotriazole (HOBt) for the first 20 residues of the glycopeptide so that the reactions could be monitored by the Kaiser test.[16] Double couplings were performed as needed. Upon completion of glycopeptide **4** (residues 77–93) the remainder of the synthesis was carried out on an automated peptide synthesizer using dicyclocarbodiimide (DCC)/HOBt-mediated couplings. Cleavage of the glycopeptide from the resin with reagent K,[17] followed by treatment with 10% aqueous hydrazine hydrate in the presence of dithiothreitol (DTT),[18] afforded the deacetylated glycopeptide **1** (160 mg, 24% overall yield).



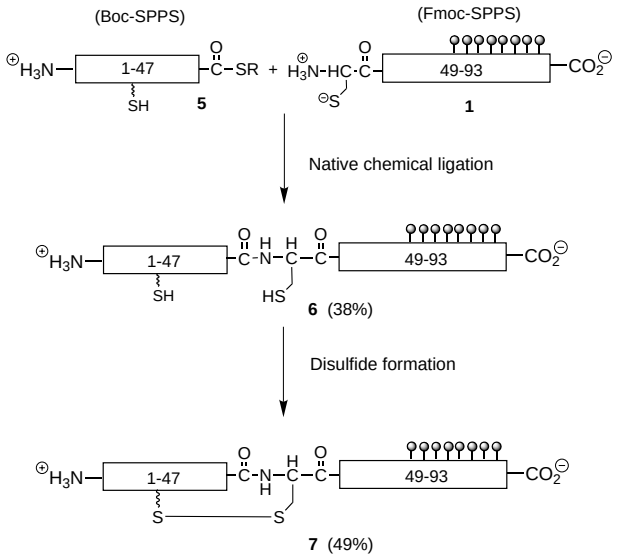
Scheme 1. Synthesis of glycopeptide **1**. a) Manual SPPS on Fmoc-Gly-Wang resin using N^{α}-Fmoc-amino acids (5 equiv), coupling with HBTU/HOBt, DIEA (5 equiv) in DMF, 1 h; b) automated SPPS using N^{α}-Fmoc-amino acids (10 equiv), coupling with DCC/HOBt (10 equiv) in NMP; c) reagent K (TFA/phenol/H_2O/thioanisole/ethanedithiol) 4 h; d) 10% N_2H_4 in H_2O, excess DTT, 30 min.

The peptide α-thioester **5** (Scheme 2) was generated by Boc-SPPS using established protocols. Of significance, attempts to synthesize the thioester fragment by Fmoc-SPPS[12] were problematic, perhaps due to the presence of the C-terminal valine residue. Treatment of the activated "safety-catch" resin with various thiol nucleophiles did not lead to the desired product. Likewise, attempts to generate the target thioester fragment using recombinant methods[19] were also unsuccessful. Thus, traditional Boc-based methods for generating peptide α-thioesters proved to be superior for the synthesis of **5**.

Ligation of glycopeptide **1** and peptide α-thioester **5** was carried out under standard conditions (Scheme 2).[20] Briefly, the peptide α-thioester **5** (42 mg, 7.8 μmol) and glycopeptide **1** (53 mg, 7.9 μmol) were reacted at pH 7.0 in aqueous buffer containing 6 M guanidinium hydrochloride (GuHCl) and 0.5% thiophenol. Due to the presence of a bulky valine residue at the ligation site the reaction proceeded slowly,[21] requiring 60 h to approach completion. Nevertheless, the ligated product **6** was isolated in 38% yield (35 mg, 2.9 μmol) after purification by RP-HPLC. The identity of the reduced glycoprotein **6** was confirmed by ESI-MS (calculated 11897.3, found 11898.7 ± 0.5).

Disulfide formation was achieved by stirring the purified glycoprotein **6** overnight at pH 8 in 2 M GuHCl in the presence

0947-6539/01/0705-1130 $ 17.50+.50/0 *Chem. Eur. J.* **2001**, *7*, No. 5



Scheme 2. Synthesis of glycosylated Lptn (**7**) by native chemical ligation.

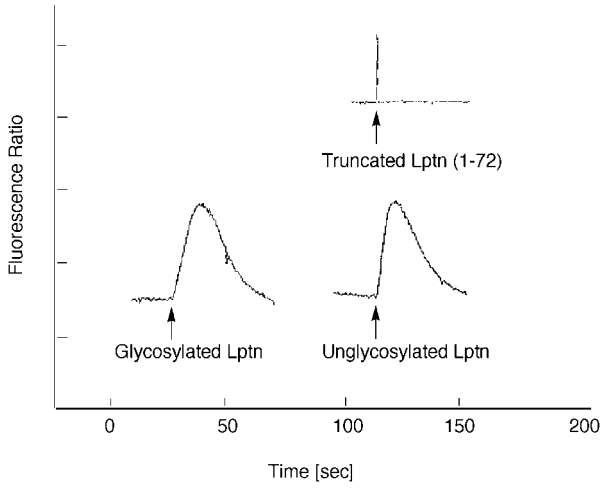


Figure 4. Calcium mobilization assay using human embryonic kidney cells (HEK 293) transfected with the human Lptn receptor XCR1. Arrows indicate points at which glycosylated, unglycosylated and truncated Lptn (1–72) were injected at a concentration of 1 mM. Both glycosylated and unglycosylated Lptn (1–93) induced a Ca^{2+} flux response while truncated Lptn was inactive.

of 8 mM cysteine and 1 mM cystine. Purification by semi-preparative RP-HPLC gave the oxidized glycoprotein **7** (17 mg, 1.5 μmol, 49% yield based on the reduced glycoprotein). The purity of the glycoprotein was confirmed by ESI-MS and analytical RP-HPLC (Figure 3).

In order to investigate the functional consequences of glycosylation of human Lptn, an unglycosylated variant was also synthesized.[22] Both variants were evaluated for their ability to bind the Lptn receptor XCR1 expressed on human embryonic kidney cells (HEK 293), and trigger a signal transduction cascade that produces an increase in intracellular calcium levels.[23] Glycosylated and unglycosylated Lptn were found to be equally active, inducing a Ca^{2+} flux response at a concentration of 1 μM in each case (Figure 4). By comparison, a truncated variant of Lptn (1–72)[5b] showed no activity at the same concentration. Since the C-terminus of Lptn is required for its activity, it is intriguing that the addition of eight *O*-linked glycans to this region of the protein has no adverse effects. By contrast, the activity of another chemokine, MCP-1, is significantly reduced when the protein is *O*-glycosylated.[9]

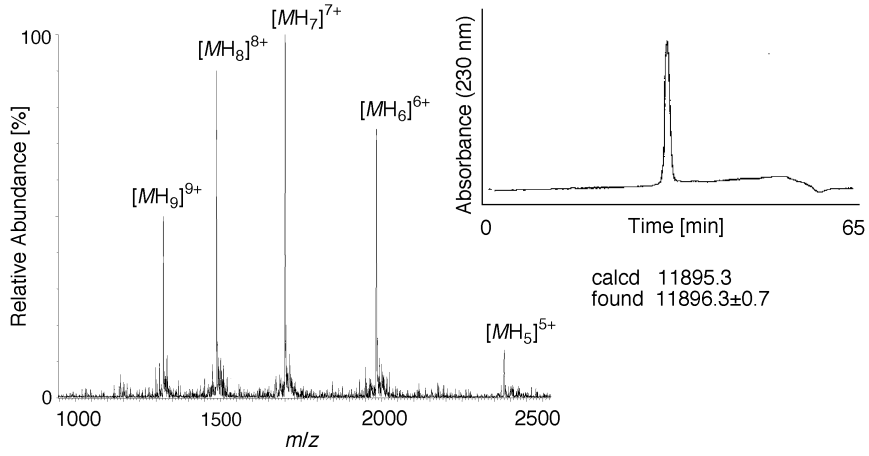


Figure 3. Electrospray mass spectrum and RP-HPLC trace of purified, oxidized glycosylated Lptn (**7**). RP-HPLC elution was performed using a gradient of 10–60% CH_3CN in H_2O (0.1% TFA) over 50 min.

Having confirmed the bioactivity of the synthetic glycosylated Lptn, we investigated its solution structure by NMR in order to assess the effects of the clustered *O*-linked glycans. One-dimensional proton spectra were acquired at various temperatures as shown in Figure 5. The most upfield amide and downfield methyl resonances (indicated with arrows) were exchange-broadened at higher temperature, consistent with a partially folded, dynamic state of the protein. As the temperature was lowered, these resonances sharpened considerably, indicating that the protein was becoming more ordered. However, similar results were obtained with the unglycosylated form of the protein;[6] this suggests that glycosylation does not significantly affect the structure.

Conclusion

A structural and functional role for glycosylation may not have been revealed in our studies for two reasons. First, the native *O*-linked glycans are undoubtedly more complex and may impart structural restraints on the protein that a single GalNAc residue cannot achieve. The Ca^{2+} flux assay probes one dimension of bioactivity, but other glycosylation-dependent functions may be possible. For example, similar to the adhesive properties of other mucin-like molecules,[1] the mucin domain of Lptn may promote its accumulation and

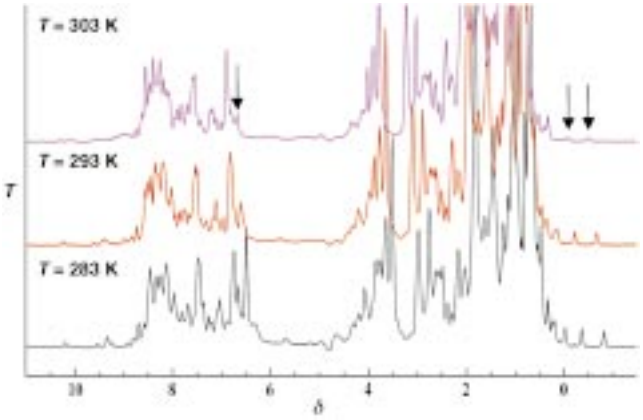


Figure 5. ^{1}H NMR spectra of glycosylated Lptn at various temperatures. Arrows indicate peaks that sharpen considerably as the temperature is lowered. Unglycosylated Lptn showed similar spectra at the indicated temperatures (Supporting Information).

retention on cell surfaces, a phenomenon thought to be important for the chemotactic function of many chemokines.[24] The synthetic glycosylated Lptn described here and more elaborated variants will provide chemically-defined reagents for exploring this hypothesis.

Experimental Section

Synthesis of glycopeptide 1: Glycopeptide **1** was synthesized on Fmoc-Gly-Wang resin (Novabiochem) on a 0.1 mmol scale. Automated peptide synthesis was carried out on an ABI 431A synthesizer. Peptide cleavage/deprotection was accomplished by treatment with reagent K at rt for 4 h. The crude glycopeptide was precipitated with *tert*-butyl methyl ether and purified by preparative RP-HPLC with a gradient of 10–60% CH_3CN in H_2O (0.1% TFA) over 50 min. After lyophilization the acetylated glycopeptide was treated with 10% aqueous N_2H_4 in the presence of DTT for 30 min at rt and directly purified by preparative RP-HPLC (same conditions as above) and lyophilized. The identity of glycopeptide **1** was confirmed by ESI-MS (calculated 6706.2, found 6706.5).

Calcium mobilization assay: The biological activities of the glycosylated, unglycosylated, and truncated forms of human Lptn were tested in calcium mobilization assays. Briefly, human embryonic kidney cells (HEK 293) transfected with human XCR1 were loaded for 60 min at 37 °C with 3 μM indol-1AM (Molecular Probes, Eugene, OR). Cells were washed and resuspended in Hank's balanced salt solution (HBSS) with 1% fetal calf serum (Gibco/BRL, Grand Island, NY) to a final concentration of 10^6 cells per mL. Calcium mobilization was measured using a Photon Technology International spectrophotometer with excitation at 350 nm and dual simultaneous recording of fluorescence emission at 400 nm and 490 nm. Relative intracellular calcium levels are expressed as the 400 nm/490 nm emission ratio. Experiments were performed at 37 °C with constant mixing in a cuvette containing 10^6 cells in 2 mL HBSS with 20 mM HEPES, pH 7.3, and 1.6 mM $CaCl_2$.

Structural analysis of synthetic Lptn: One-dimensional proton spectra were obtained by dissolving the lyophilized protein in 90% H_2O, 10% D_2O, 0.05% NaN_3 and adjusting the pH to 5.0. Spectra were recorded with presaturation on a 600 MHz Bruker DMX spectrometer at various temperatures.

Acknowledgement

This research was supported by a grant to T.M.H. from the National Institute of Health and by a grant to C.R.B. from the National Science Foundation (CHE-9734439) and the University of California BioSTAR Project (S97-73) in partnership with Gryphon Sciences. L.A.M. was supported by a graduate fellowship from Roche Bioscience.

[1] a) K. L. Carraway, S. R. Hull, *Glycobiology* **1991**, *1*, 131; b) G. J. Strous, J. Dekker, *J. Crit. Rev. Biochem. Mol. Biol.* **1992**, *27*, 57; c) J. Hilkens, M. J. L. Ligtenberg, H. L. Vos, S. V. Litvinov, *Trends Biochem. Sci*. **1992** *17*, 359; d) Y. Shimizu, S. Shaw, *Nature* **1993**, *366*, 630.
[2] M. L. Thomas, *Annu. Rev. Immunol*. **1989**, *7*, 339.
[3] J. L. Goldstein, M. S. Brown, R. G. W. Anderson, D. W. Russell, W. J. Schneider, *Annu. Rev. Cell Biol*. **1985**, *1*, 1.
[4] P. Reddy, I. Caras, M. Krieger, *J. Biol. Chem*. **1989**, *264*, 17329.
[5] a) B. Dorner, S. Müller, F. Entschladen, J. M. Schröder, P. Franke, R. Kraft, P. Friedl, I. Clarke-Lewis, R. A. Kroczek, *J. Biol. Chem*. **1997**, *272*, 8817; b) J. A. Hedrick, V. Saylor, D. Figueroa, L. Mizoue, Y. Xu, S. Menon, J. Abrams, T. Handel, A. Zlotnik, *J. Immunol*. **1997**, *158*, 1533.
[6] One-dimensional proton NMR and two-dimensional HSQC spectra of unglycosylated Lptn are provided as Supporting Information.
[7] a) S. E. O'Connor, B. Imperiali, *Chem. Biol.* **1998**, *5*, 427; b) D. H. Live, L. J. Williams, S. D. Kuduk, J. B. Schwarz, P. W. Glunz, X.-T. Chen, D. Sames, R. A. Kumar, S. J. Danishefsky, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 3489; c) R. Liang, A. H. Andreotti, D. Kahne, *J. Am. Chem. Soc.* **1995**, *117*, 10395; d) D. F. Wyss, J. S. Choi, J. Li, M. H. Knoppers, K. J. Willis, A. R. N. Arulanandam, A. Smolyar, E. L. Reinherz, G. Wagner, *Science* **1995**, *269*, 1273.
[8] a) J. F. Bazan, K. B. Bacon, G. Hardiman, W. Wang, K. Soo, D. Rossi, D. R. Greaves, A. Zlotnik, T. J. Schall, *Nature* **1997**, *385*, 640; b) A. M. Fong, H. P. Erickson, J. P. Zachariah, S. Poon, N. J. Schamberg, T. Imai, D. D. Patel, *J. Biol. Chem.* **2000**, *275*, 3781.
[9] P. Proost, S. Struyf, M. Couvreur, J.-P. Lenaerts, R. Conings, P. Menten, P. Verhaert, A. Wuyts, J. V. Damme, *J. Immunol*. **1998**, *160*, 4034.
[10] T. Imai, K. Hieshima, C. Haskell, M. Baba, M. Nagira, M. Nishimura, M. Kakizaki, S. Takagi, H. Nomiyama, T. J. Schall, O. Yoshie, *Cell* **1997**, *91*, 521.
[11] a) P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 766; b) T. W. Muir, P. E. Dawson, S. B. H. Kent, *Methods Enzymol*. **1997**, *289*, 266.
[12] Y. S. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman, C. R. Bertozzi, *J. Am. Chem. Soc*. **1999**, *121*, 11684.
[13] For the synthesis of *O*-linked glycopeptides, see: a) E. Meinjohanns, M. Meldal, H. Paulsen, A. Schleyer, K. Bock, *J. Chem. Soc. Perkin Trans. 1* **1996**, 985; b) N. Mathieux, H. Paulsen, M. Meldal, K. Bock, *J. Chem. Soc. Perkin Trans. 1* **1997**, 2359; c) Y. Nakahara, H. Iijima, T. Ogawa, *Tetrahedron Lett.* **1994**, *35*, 3321; d) B. Liebe, H. Kunz, *Tetrahedron Lett*. **1994**, *35*, 8777; e) *Glycopeptides and Related Compounds* (Eds.: D. G. Large, C. D. Warren), Marcel Dekker, New York, **1997**, and references therein.
[14] H. Hojo, S. Aimoto, *Bull. Chem. Soc. Jpn*. **1991**, *64*, 111.
[15] a) S. D. Kuduk, J. B. Schwarz, X.-T. Chen, P. Glunz, D. Sames, G. Ragupathi, P. Livingston, S. J. Danishefsky, *J. Am. Chem. Soc.* **1998**, *120*, 12474; b) K. A. Winans, D. S. King, V. R. Rao, C. R. Bertozzi, *Biochemistry* **1999**, *38*, 11701 and references therein.
[16] E. Kaiser, R. L. Colescott, C. D. Bossinger, P. I. Cook, *Anal. Biochem*. **1970**, *34*, 595.
[17] D. S. King, C. G. Fields, G. B. Fields, *Int. J. Pept. Protein Res.* **1990**, *36*, 255.
[18] In the absence of DTT the disulfide-bound dimer was obtained.
[19] T. W. Muir, D. Sondi, P. A. Cole, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6705.
[20] J. Wilken, D. Hoover, D. A. Thompson, P. N. Barlow, H. McSparron, L. Picard, A. Wlodawer, J. Lubkowski, S. B. H. Kent, *Chem. Biol*. **1999**, *6*, 43.
[21] T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10068.
[22] Unglycosylated Lptn was also prepared by native chemical ligation.
[23] T. Yoshida, T. Imai, M. Kakizaki, M. Nishimura, S. Takagi, O. Yoshie, *J. Biol. Chem*. **1998**, *273*, 16551.
[24] a) G. S. Kuschert, F. Coulin, C. A. Power, A. E. Proudfoot, R. E. Hubbard, A. J. Hoogewerf, T. N. Wells, *Biochemistry* **1999**, *38*, 12959; b) A. J. Hoogewerf, G. S. Kuschert, A. E. Proudfoot, F. Borlat, I. Clark-Lewis, C. A. Power, T. N. Wells, *Biochemistry* **1997**, *36*, 13570.

Received: November 6, 2000 [F2848]

