Pergamon

Tetrahedron Letters, Vol. 37, No. 12, pp. 1975-1978, 1996

0040-4039/96 $15.00 + 0.00

S0040-4039(96)00261-4

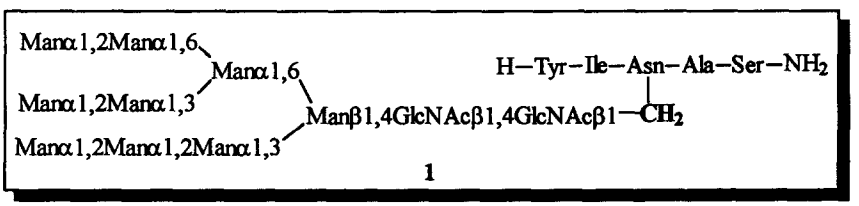
CHEMOENZYMATIC SYNTHESIS OF A HIGH-MANNOSE-TYPE *N*-GLYCOPEPTIDE ANALOG WITH *C*-GLYCOSIDIC LINKAGE

Lai-Xi Wang, Jian-Qiang Fan, Yuan C. Lee*

Department of Biology, The Johns Hopkins University, Baltimore, MD 21218, U. S. A.

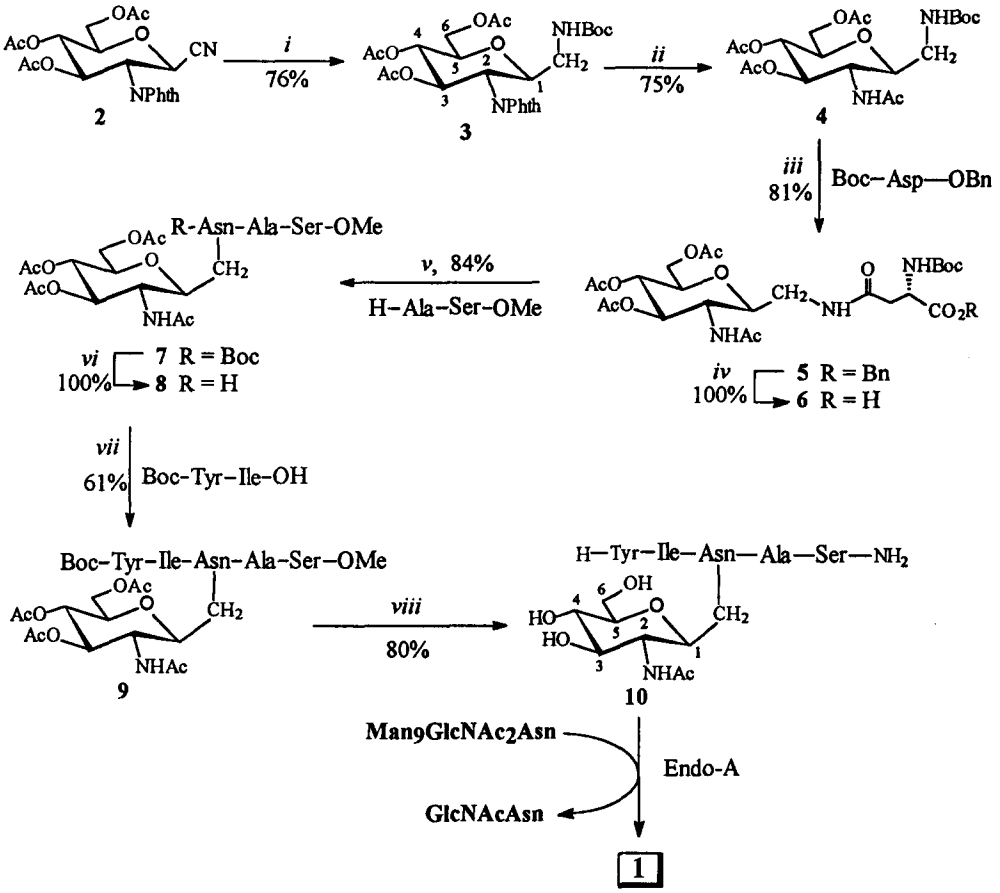
Abstract: *A synthesis of the title compound by chemical synthesis of a GlcNAc-CH_2-Asn containing peptide and enzymatic transfer of a $Man_9GlcNAc$ group to it was described.*

The diverse biological functions shown by glycoproteins have heightened interests in the synthesis of glycopeptides which represent partial structures of the glycoproteins.[1] The known strategies of glycopeptide synthesis include solid-phase synthesis using glyco-amino acid building blocks;[2] the convergent coupling of Asp-containing peptide with unprotected glycosylamine;[3] and the enzymatic synthesis making use of peptidases and glycosyltransferases.[4]



1

Our interests in the mechanism and specificity of peptide-N^4-(*N*-acetyl-β-D-glucosaminyl)-asparagine amidase, an important enzyme which is able to release the intact oligosaccharide from *N*-glycoproteins by cleaving the β-aspartyl-glucosylamine linkage,[5] led us to design and synthesize various substrate analogs. Among others, the insertion of a functional group between the crucial carbohydrate-peptide linkage of the natural glycopeptide substrates is of great interest. However, the above-mentioned approaches are not straightforward to these molecules. Here we describe a strategy of combining chemical and enzymatic methods for their synthesis. The essential enzyme was *Arthrobacter protophormiae* endo-β-*N*-acetyl-glucosaminidase (Endo-A), a hydrolyase with the activity of transfering a $Man_{5-9}GlcNAc$ residue to the 4-OH of a terminal GlcNAc residue.[6,7] A high-mannose-type *N*-glycopentapeptide analog (**1**) with the insertion of a methylene group at the crucial linkage region was selected as a model compound. Our synthetic strategy was



i: H_2, Pd/C, THF-EtOH (9:1), r.t., 16 h; (b) Boc_2O, Et_3N, THF, 35°C, 10 h.
ii: (a) hydrazine hydrate, EtOH, 80°C, 2 h; (b) Ac_2O, pyridine, r.t., 6 h.
iii: (a) **4**, 4M HCl in dioxane, CH_2Cl_2, r.t., 0.5 h;
(b) Boc-Asp-OBn, DCC, HOBt, Et_3N, CH_2Cl_2-THF (3:1), r.t., 10 h.
iv: H_2, Pd/C, THF-EtOH (1:2), r.t., 10 h. *v*: DCC, HOBt, Et_3N, DMF, r.t., 10 h.
vi: 4M HCl in dioxane, CH_2Cl_2, r.t., 1 h. *vii*: DCC, HOBt, Et_3N, DMF, r.t., 16 h.
viii: (a) methanolic ammonia, MeOH, r.t, 3 days; (b), 3M HCl, r.t., 1 h.

Ac: *Acetyl*; **Bn**: *Benzyl*; **Boc**: *tert-Butoxycarbonyl*; **DCC**: *Dicyclohexylcarbodiimide*; **HOBt**: *1-Hydroxybenzotriazole*; **Phth**: *Phthaloyl*.

based on the finding that the transglycosylation yield of Endo-A could be substantially enhanced by performing the enzymatic reaction in media containing organic solvents such as aqueous acetone.[7]

A chemical synthesis of the key intermediate, the *C*-glycopentapeptide (**10**), was summarized in the Scheme. The β-glycosyl cyanide **2**[8] was converted into **3**[9] by hydrogenation and subsequent *N*-protection of the resulting glycosylmethylamine. A large $J_{1,2}$-value (9.8 Hz) indicated that **3** was the desired β-*C*-glycoside. Treatment of **3** with hydrazine hydrate and subsequent acetylation gave **4**,[9] which was de-*N*-Boc-protected and condensed with Boc-Asp-OBn to provide the fully protected N^4-(2-acetamido-2-deoxy-β-D-glucopyranosylmethyl)-asparagine (**5**).[9] This compound can serve as a building block for synthesizing various glycopeptides containing this structural unit. To prepare **10**, a stepwise solution synthesis approach was used. Compound **5** was de-*O*-benzylated and coupled with dipeptide H-Ala-Ser-OMe to give the tripeptide **7**,[9] which was then elongated from the *N*-terminal to the pentapeptide derivative **9**[9] by coupling with the dipeptide Boc-Tyr-Ile-OH. Finally, deprotection of **9** via sequential treatment with methanolic ammonia and aqueous HCl successfully gave **10**.[9]

To transfer a high-mannose structure to the synthetic *C*-glycopentapeptide by Endo-A, we used $Man_9GlcNAc_2Asn$ prepared from soybean agglutinin by an established method[10] as the donor substrate. A mixture consisting of $Man_9GlcNAc_2Asn$ (2.4 μmol), **10** (12 μmol, 50 mM), and the enzyme (46.8 mU) in 25 mM NH_4OAc buffer (240 μL, pH 6.0) containing 35% acetone was incubated at 37°C for 20 min. The reaction was stopped by boiling in a 100°C water bath (3 min). The transglycosylation product was purified by HPLC on a Microsorb MW ODS column (4.6x250 mm) with 10% *aq*. MeCN containing 0.05% trifluoroacetic acid as eluant (flow rate: 1.0 mL/min.). The product eluted at 5.5 min was collected and lyophilized to provide **1** (1.56 mg, 0.634 μmol, 26%).[9] The excess **10** was eluted after 8 min under the condition and recovered. The ^{1}H-NMR spectrum of **1** showed eight α-Man H-1 signals at δ5.400--4.894 and one β-Man H-1 signal at δ4.783. A doublet at δ4.618 with a large *J* value (7.5 Hz) assignable to H-1 of the second GlcNAc indicated that the newly formed glycosidic linkage in **1** was in β-D-configuration. In addition, amino acid and sugar composition analysis of **1** gave satisfactory results. HR-FABMS of **1**: calculated for $C_{96}H_{157}N_9O_{63} + H^+$ ($M+H^+$) 2444.9436; Found 2444.9462.

In summary, a chemoenzymatic approach to the synthesis of a *C*-linked glycopentapeptide was described. This strategy should be also suitable for the preparation of other high-mannose-type glycopeptide analogs, and can be extended to the complex-type, when a suitable endo-enzyme becomes available.

Acknowledgement

We thank Dr. Yuanda Zhang of Caltech for measuring the mass spectra.

References and Notes

1. Meldal, M. in *Neoglycoconjugates: Preparation and Applications*, Lee, Y.C. and Lee, R.T., Eds., Academic Press, Orlando, **1994**, 145.
2. Meldal, M.; Bock, K. *Glycoconjugate J.*, **1994**, *11*, 59.
3. (a) Cohen-Anisfeld, S.T.; Lansbury, Jr., P.T. *J. Am. Chem. Soc.*, **1993**, *115*, 10531. (b) Wong, S.Y.C.; Guile, G.R.; Rademacher, T.W.; Dwek, R.A. *Glycoconjugate J.*, **1993**, *10*, 227.
4. (a) Wong, C.H.; Schuster, M.; Wang, P.; Sears, P. *J. Am. Chem. Soc.*, **1993**, *115*, 5893. (b) Schuster, M.; Wang, P.; Paulson, J.C.; Wong, C.H. *J. Am. Chem. Soc.*, **1994**, *116*, 1135 and references cited therein.
5. (a) Takahashi, N.; Nishibe, H. *Biochem. Biophys. Acta*, **1981**, *657*, 457. (b) Plummer, Jr., T.H.; Tarentino, A.L. *Glycobiology*, **1991**, *1*, 257. (c) Gosselin, S.; Payie, K.; Viswanatha, T. *Glycobiology*, **1993**, *3*, 419.
6. Takegawa, K.; Yamaguchi, S.; Kondo, A.; Iwamoto, H.; Nakoshi, M.; Kato, I.; Iwahara, S. *Biochem. Int.*, **1991**, *24*, 849.
7. Fan, J.Q.; Takegawa, K.; Iwahara, S.; Kondo, A.; Kato, I.; Abeygunawardana, C.; Lee, Y.C. *J. Biol. Chem.*, **1995**, *270*, 17723.
8. Myers, R.W.; Lee, Y.C. *Carbohydr. Res.*, **1986**, *154*, 145.
9. New compounds gave satisfactory microanalysis and/or HR-MS. selected ^{1}H-NMR data are listed below (*J* in Hz). **3**: δ_H ($CDCl_3$) 4.490 (dt, 1 H, *J*1.2, 10.3, H-1), 4.300 (t, 1 H, *J* 9.8, H-2), 3.256--3.288 (m, 2 H, CH_2N), **4**: δ_H ($CDCl_3$) 3.952 (q, 1 H, *J*9.4, H-2), 3.590 (m, 1 H, H-1), 3.462 and 3.051 (m, 2 H, CH_2N), 2.095, 2.038, 2.031, and 1.965 (each s, each 3 H, 4 Ac), 1.445 (s, 9 H, Boc); **5**: δ_H (DMSO-*d*6) 5.093 (s, 2 H, CH_2Ph), 4.351 (m, 1 H, α-C*H* Asp), 3.529 (m, 1 H, H-1), 3.352 and 2.948 (m, 2 H, CH_2N), 1.358 (s, 9 H, Boc); **6**: δ_H (DMSO-*d*6) 12.445 (s, 1 H, CO_2H), 1.362 (s, 9 H, Boc); 7: δ_H (DMSO-*d*6) 4.401--4.285 (m, 2 H, α-C*H* in Asp and Ser), 4.160 (m, 1 H, α-C*H* Ala), 3.618 (s, 3 H, OMe), 3.507 (m, 1 H, H-1), 1.366 (s, 9 H, Boc), 1.202 (d, 3 H, *J*7.0, β-CH_3 Ala); **9**: δ_H (DMSO-*d*6) 7.023--6.628 (m, 4 H, Tyr), 3.610 (s, 3 H, OMe), 3.510 (m, 1 H, H-1), 3,348 and 2.982 (m, 2 H, CH_2N), 1.297 (s, 9 H, Boc), 1.181 (d, 1 H, *J*7.0, β-CH_3 Ala), 0.803 (m, 6 H, 2 CH_3 in Ile); **10**: δ_H (D_2O) 7.046--6.807 (m, 4 H, Tyr), 4.590 (t, 1 H, *J*6.8, α-CH Tyr), 4.350 (t, 1 H, *J*4.9, α-CH Ser), 4.265 (q, 1 H, *J*7.3, α-CH Ala), 4.065 (m, α-CH Asn), 3.860 (m, α-CH Ile), 1.959 (s, 3 H, NAc), 1.358 (d, 1 H, *J*7.2, β-CH_3 Ala), 0.805--0.764 (m, 6 H, 2 CH_3 in Ile); **1**: δ_H (D_2O) 7.112 and 6.855 (each d, 4 H, *J*8.3, Tyr), 5.400, 5.344, 5.312, 5.143, 5.094, 5.083, 5.078, and 4.894 (each br. s, 8 H, H-1 of α-Man), 4.783 (s, 1 H, H-1 of β-Man), 4.618 (d, 1 H, *J*7.5, H-1 of GlcNAc-2), 3.174 (dd, 1 H, *J*7.5 and 14, 1/2 CH_2N), 2.086 and 2.044 (s, each 3 H, 2 NAc), 1.765 (m, 1 H, β-CH Ile), 1.426 (d, 3 H, *J*7.3, β-CH_3 Ala), 1.314 and 1.100 (m, 2 H, γ-CH_2 Ile), 0.868--0.838 (m, 6 H, 2 CH_3 Ile).
10. (a) Fan, J.Q.; Kondo, A.; Kato, I.; Lee, Y.C. *Anal. Biochem.*, **1994**, *219*, 224. (b) Dorland, L.; van Halbeek, H.; Vliegenthart, J.F.G.; Lis, H.; Sharon, N. *J. Biol. Chem.*, **1981**, *256*, 7708.

(*Received in USA* 13 *October* 1995; *revised* 17 *January* 1996; *accepted* 20 *January* 1996)