Carbohydrate Research, 116 (1983) C9–C12
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

Preliminary communication

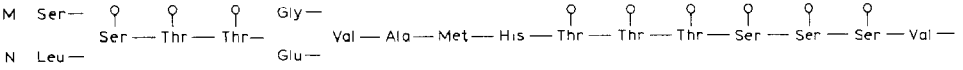
Synthesis of glycopeptides having clusters of *O*-glycosylic disaccharide chains [β-D-Gal-(1→3)-α-D-GalNAc] located at vicinal amino acid residues of the peptide chain*

VICENTE VEREZ BENCOMO and PIERRE SINAŸ**

Laboratoire de Biochimie Structurale, E.R.A. 739, U.E.R. de Sciences Fondamentales et Appliquées, 45046 Orléans Cédex (France)

(Received February 21st, 1983; accepted for publication, March 24th, 1983)

The Thomsen–Friedenreich (T) receptor is defined as the cryptic structure that is uncovered after neuraminidase treatment of vertebrate erythrocytes[1]. This enzyme removes sialic acid residues from the oligosaccharide chains which are mainly located at the amino-terminal region of glycophorin A (PAS-1), the major glycoprotein of the human red-cell membrane. This region, which is known to carry MN blood-group specificities, has the following primary structure[2].



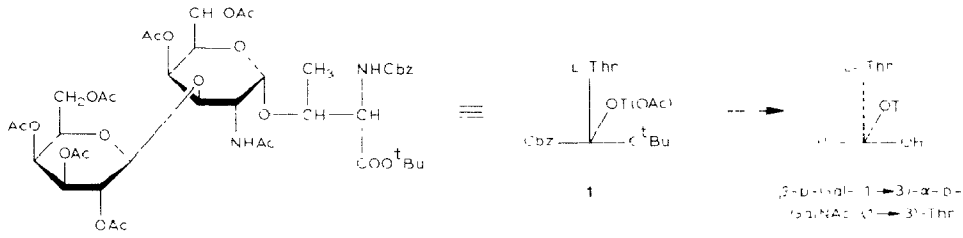
One characteristic feature of this structure is a genetically determined polymorphism at the first and fifth positions which, in part, may account for the observed MN-specificities[3]. On the other hand, the amino-terminal segment uniquely displays two typical clusters of *O*-glycosylic oligosaccharide chains (♀) located at vicinal amino acid residues (Ser or Thr) of the polypeptide chain. After removal of sialic acid residues, the oligosaccharide chain is the disaccharide β-D-Gal-(1→3)-α-D-GalNAc, which is considered to be the antigenic determinant of the Thomsen–Friedenreich (T) antigen[1].

The chemical synthesis of such heavily clustered glycopeptides is important, in order to define the influence of a high density of carbohydrate antigenic-determinants on the expression of the T-antigenicity. All glycopeptides active toward *Vicia graminea* lectin have such heavy clusters[4]. Furthermore, knowledge of the precise chemical structure of the Thomsen–Friedenreich receptor, a tumour-associated antigen that is not oncofetal in origin[5], could be of importance in cancer research. We now describe a synthesis of such heavy clusters.

We recently reported[6] the synthesis of *O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-(1→3)-*O*-(2-acetamido-4,6-di-*O*-acetyl-2-deoxy-α-D-galactopyranosyl)-(1→3)-*N*-(benzyloxycarbonyl)-L-threonine *tert*-butyl ester (**1**) and the L-serine analogue **2**.

*Dedicated to Professor Edgar Lederer on the occasion of his 75th birthday.
**To whom enquiries should be sent.

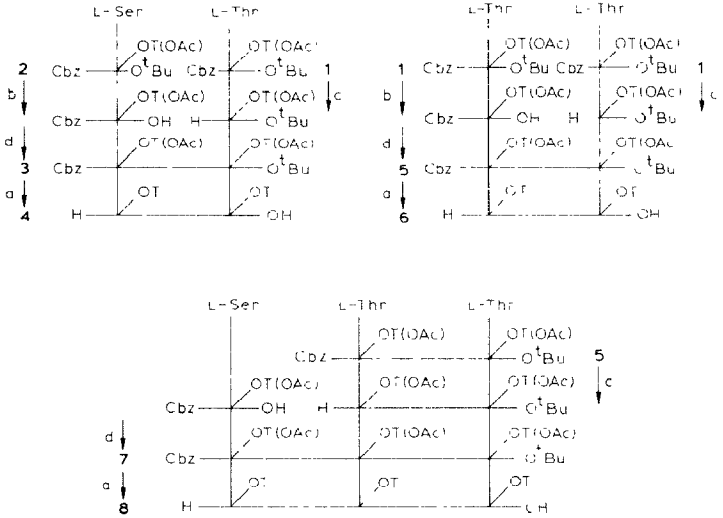
0008-6215/83/$03.00



Scheme 1 Reagents (1) MeOH–Et_3N, room temperature, 48 h, (2) CF_3COOH, room temperature, 1 h, (3) H_2, 10% Pd/C, MeOH–AcOH (10 : 1), room temperature, 24 h

Application to **1** of the sequence shown in Scheme 1 gave, after purification on Sephadex G-10, *O*-β-D-galactopyranosyl-(1→3)-*O*-(2-acetamido-2-deoxy-α-D-galactopyranosyl)-(1→3)-L-threonine [β-D-Gal-(1→3)-α-D-GalNAc-(1→3)-Thr] (80%), $[\alpha]_D$ +83° (*c* 0.5, water); lit.[7] $[\alpha]_D$ +90° (*c* 1.3, water). *O*-Deacetylation of **1** with Et_3N–MeOH did not promote β-elimination of the disaccharide[8], and the *O*-glycosylic linkage was resistant to brief treatment with trifluoroacetic acid[9]. The homogeneity and identity of this disaccharide–amino acid were confirmed by p.c. and ^{1}H-n.m.r. spectroscopy (400 MHz, D_2O).

The selective deblocking of the *N*- and *C*-terminals of **1** and **2**, and of the tetrasaccharide-dipeptide **5**, was performed as shown in Scheme 2. The peptide linkage was formed in the presence of 2-ethoxy-*N*-(ethoxycarbonyl)-1,2-dihydroquinoline[11] (EEDQ), to give the protected clusters **3** (63%), **5** (65%), and **7** (62%).



Scheme 2 Reagents (a) see Scheme 1, (b) CF_3COOH, room temperature, 1 h, (c) 10% Pd/C, cyclohexene–EtOH (1 : 2), reflux, 1 h[10], (d) EEDQ, dichloromethane, −10°→room temperature, 15 h

Compound **3** had $[\alpha]_D$ +73° (*c* 1.5, chloroform). ^{1}H-N.m.r. data (400 MHz, $CDCl_3$): δ 7.34 (s, 5 H, Ph), 6.53 (d, 1 H, *J* 9 Hz, NH), 6.21 (d, 1 H, *J* 8 Hz, NH), 5.06 (d, 1 H, $J_{1,2}$ 3.8 Hz, H-1), 4.63 (d, 1 H, $J_{1',2'}$ 8 Hz, H-1′), 4.58 (d, 1 H, $J_{1',2'}$ 8 Hz, H-1′), 2.23–1.93 (14 s, 42 H, 14 Ac), 1.52 (s, 9 H, ^{t}Bu), and 1.29 (d, 3 H, $J_{CH,Me}$ 7 Hz, Thr Me).

Anal. Calc. for $C_{71}H_{98}N_4O_{39}$: C, 52.27; H, 6.05; N, 3.43. Found: C, 52.06; H, 6.09; N, 3.32.

Compound **5** had $[\alpha]_D$ +75° (*c* 1, chloroform). ^{1}H-N.m.r. data (400 MHz, $CDCl_3$): δ 7.35 (s, 5 H, Ph), 6.55, 6.11, and 5.83 (3 d, 3 H, *J* 9 Hz, NH), 5.15 (d, 1 H, $J_{1,2}$ 3.4 Hz, H-1), 5.02 (d, 1 H, $J_{1,2}$ 3.5 Hz, H-1), 4.66 (d, 1 H, $J_{1',2'}$ 7.6 Hz, H-1'), 4.64 (d, 1 H, $J_{1',2'}$ 7.8 Hz, H-1'), 2.28–2.00 (14 s, 42 H, 14 Ac), 1.57 (s, 9 H, ^{t}Bu), 1.48 and 1.40 (2 d, 6 H, $J_{CH,Me}$ 6.4 Hz, Thr Me).

Anal. Calc. for $C_{72}H_{100}N_4O_{39}\cdot H_2O$: C, 51.98; H, 6.18; N, 3.36. Found: C, 51.94; H, 6.07; N, 3.44.

Compound **7** had $[\alpha]_D$ +78° (*c* 0.9, chloroform).

Anal. Calc. for $C_{101}H_{140}N_6O_{57}\cdot 3H_2O$: C, 50.45; H, 6.12; N, 3.49. Found: C, 50.41; H, 6.17; N, 3.43.

Deprotection of the clusters was routinely performed, using the sequence described in Scheme 1, and purification was achieved on Sephadex G-10 (clusters **4** and **6**) or Sephadex G-25 (cluster **8**). Elution with water gave a single peak at 206 nm. Amino acid analyses were in agreement with the structures.

Compound **4** had $[\alpha]_D$ +67° (*c* 0.7, water). ^{1}H-N.m.r. data (400 MHz, D_2O): δ 4.91 (d, 1 H, $J_{1,2}$ 3 Hz, H-1), 4.89 (d, 1 H, $J_{1,2}$ 3 Hz, H-1), 1.93 and 1.87 (2 s, 6 H, 2 Ac), and 1.14 (d, 3 H, $J_{CH,Me}$ 6.3 Hz, Thr Me).

Compound **6** had $[\alpha]_D$ +97° (*c* 1.5, water). ^{1}H-N.m.r. data (400 MHz, D_2O): δ 5.03 (d, 1 H, $J_{1,2}$ 3.4 Hz, H-1), 4.92 (d, 1 H, $J_{1,2}$ 3.2 Hz, H-1), 1.97 and 1.95 (2 s, 6 H, 2 Ac), 1.25 and 1.22 (2 d, 6 H, $J_{CH,Me}$ 6.5 Hz, Thr Me).

Compound **8** had $[\alpha]_D$ +109° (*c* 0.5, water). ^{1}H-N.m.r. data (400 MHz, D_2O): δ 2.04, 2.01, and 1.99 (3 s, 9 H, 3 Ac), 1.33 and 1.18 (2 d, 6 H, $J_{CH,Me}$ 6.5 Hz, Thr Me).

Using the same methodology, analogues were prepared with either L-serine or L-leucine as the *N*-terminal amino acid. These compounds are useful for studies of the molecular basis of the sialic acid-independent MN-antigenicity[12] and to define the receptor site of *Vicia graminea* lectin[4]. Such studies will be reported elsewhere.

ACKNOWLEDGMENTS

We thank the Institut National de la Santé et de la Recherche Médicale (C.R.L. No. 82 1027) for financial support, the Ministère des Relations Extérieures (France) for a fellowship (to V.V.B.), Drs. Lukacs and Lallemand (Gif-sur-Yvette, France) for performing high-field n.m.r. measurements, and Professor Monsigny for amino acid analyses.

REFERENCES

1 G. Uhlenbruck, *Immunol. Commun.*, 10 (1981) 251–264.
2 M. Tomita and V. T. Marchesi, *Proc. Natl. Acad. Sci. U.S.A.*, 72 (1975) 2964–2968.
3 E. Lisowska and K. Waśniowska, *Eur. J. Biochem.*, 88 (1978) 247–252; O. O. Blumenfeld and A. M. Adamany, *Proc. Natl. Acad. Sci. U.S.A.*, 75 (1978) 2727–2731.
4 M. Duk, E. Lisowska, M. Kordowicz, and K. Waśniowska, *Eur. J. Biochem.*, 123 (1982) 105–112.
5 G. F. Springer, P. R. Desai, M. S. Murthy, H. Tegtmeyer, and E. F. Scanlon, *Prog. Allergy*, 26 (1979) 42–96.

6 V. Verez-Bencomo, J.-C. Jacquinet, and P. Sinaÿ, *Carbohydr. Res.*, 110 (1982) C9–C11.
7 H. Paulsen and J.-P. Hölck, *Carbohydr. Res.*, 109 (1982) 89–107.
8 V. A. Derevitskaya, M. G. Vafina, and N. K. Kochetkov, *Carbohydr. Res.*, 3 (1967) 377–388.
9 H. J. Koeners, C. Schattenkerk, J. J. Verhoeven, and J. H. Van Boom, *Tetrahedron*, 37 (1981) 1763–1771.
10 A. E. Jackson and R. A. W. Johnstone, *Synthesis*, (1976) 685–687.
11 E. Belleau and G. Malek, *J. Am. Chem. Soc.*, 90 (1968) 1651–1652.
12 W. J. Judd, P. D. Issitt, B. G. Pavone, J. Anderson, and D. Aminoff, *Transfusion (Philadelphia)*, 19 (1979) 12–18.