Carbohydrate Research 419 (2016) 51–59

Contents lists available at ScienceDirect

Carbohydrate Research

journal homepage: www.elsevier.com/locate/carres

Novel structural features of the immunocompetent ceramide phospho-inositol glycan core from *Trichomonas vaginalis*

Christian Heiss [a,*], Zhirui Wang [a], Ian Black [a], Parastoo Azadi [a], Raina N. Fichorova [b], Bibhuti N. Singh [c,**]

[a] Complex Carbohydrate Research Center, The University of Georgia, 315 Riverbend Road, Athens, GA 30602, USA
[b] Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, 221 Longwood Avenue, Boston, MA 02115, USA
[c] Department of Biochemistry and Molecular Biology, SUNY Upstate Medical University, 750 E Adams St, Syracuse, NY 13210, USA

ARTICLE INFO

Article history:
Received 24 September 2015
Received in revised form 28 October 2015
Accepted 2 November 2015
Available online 10 November 2015

Keywords:
Trichomonas vaginalis
Glycoconjugates
Phospho-inositol glycan core
NMR

ABSTRACT

The ceramide phosphoinositol glycan core (CPI-GC) of the lipophosphoglycan of *Trichomonas vaginalis* is a major virulent factor of this common genitourinary parasite. While its carbohydrate composition has been reported before, its structure has remained largely unknown. We isolated the glycan portions of CPI-GC by nitrous acid deamination and hydrofluoric acid treatment and investigated their structures by methylation analysis and 1- and 2-D NMR. We found that the α-anomer of galactose is a major constituent of CPI-GC. The β-anomer was found exclusively at the non-reducing end of CPI-GC side chains. Furthermore the data showed that the rhamnan backbone is more complex than previously thought and that the inositol residue at the reducing end is linked to a 4-linked α-glucuronic acid (GlcA) residue. This appears to be the most striking and novel feature of this GPI-anchor type molecule.

1. Introduction

Human trichomoniasis is caused by an extracellular urogenital protozoan parasite, *Trichomonas vaginalis*. It is the most common non-viral sexually transmitted infection (~180 million people) worldwide. *T. vaginalis* adheres to and damages vaginal epithelial cells[1] and lives in the vagina and the male urethra.[2] In addition to being a serious cause of discomfort and the leading cause of vaginitis, the infection has been linked to a myriad of reproductive problems, e.g. preterm delivery, low birth weight, infertility, bacterial vaginosis, risk of HIV infection and transmission, HPV infection and cancer.[3–8] The infection is often recurrent with no lasting immunity, implying the importance of the innate immune defenses. Almost half of the women diagnosed with trichomoniasis are asymptomatic while the others develop a severe inflammatory reaction, which is an additional risk factor for HIV acquisition and other viral infections.[9,10]

Parasitic protozoa contain a variety of complex carbohydrates on their surfaces, e.g., glycolipids, glycoproteins, and glycosylated phosphatidylinositol glycolipids (GPI).[11–15] These surface glycoconjugates are known to play intrinsic roles in cell to cell interactions, host cell invasion and evasion of host responses.[11,12,15]

T. vaginalis contains a predominant GPI-anchored cell surface lipophosphoglycan (TV-LPG) that is involved in adherence of parasites to cervical epithelial cells,[13,16] and that LPG from vaginal isolates elicits a selective and species-specific production of cytokines by human female genital tract epithelial cells.[1,17] We have previously reported the biochemical nature of TV-LPG derived from a proinflammatory *T. vaginalis* isolate (UR1)[1,13,15] and have also demonstrated that the specific molecular domain of LPG, the ceramide phospho-inositol glycan core (CPI-GC), is required for triggering immunoinflammatory pathways in human vaginal and cervical epithelia[15] and amplifies inflammatory responses to vaginal bacteria.[18] A schematic diagram of *T. vaginalis* LPG and its substructure, CPI-GC, was previously published by Singh et al.[15] TV-LPG is analogous to the LPGs from other parasites such as *Leishmania* but is clearly distinct from them in its monosaccharide composition and other structural features. *Leishmania* LPG is a polymer of [6)-β-Gal-(1→4)-α-Man-1-*P*-(O→] repeating units attached to a glycan core that is embedded into the membrane via a 1-O-alkyl-2-lyso-phosphatidylinositol anchor.[19] TV-LPG contains no protein/peptide (unlike other GPI-anchored molecules) and no mannose, and is anchored via inositol-phosphate-ceramide.[14,15] The CPI-GC and TV-LPG contain terminal non-reducing β-galactosyl residues and lactosamine repeats, providing the molecular basis of CPI-GC interactions with the galectin family of host proteins, which have

* Corresponding author. Complex Carbohydrate Research Center, The University of Georgia, 315 Riverbend Road, Athens, GA 30602, USA. Tel.: +1 706 542 3008; fax: +1 706 542 4412.
E-mail address: cheiss@uga.edu (C. Heiss).
** Corresponding author. Department of Biochemistry and Molecular Biology, SUNY Upstate Medical University, 750 E Adams St, Syracuse, NY 13210, USA. Tel.: +1 315 464 5398; fax: +1 315 464 8750.
E-mail address: singhb@upstate.edu (B.N. Singh).

http://dx.doi.org/10.1016/j.carres.2015.11.001
0008-6215/

emerged as regulators of innate and adoptive immunity and host-pathogen interactions.[20–22]

TV-CPI-GC has been shown to be composed of the monosaccharide building blocks rhamnose (Rha), galactose (Gal), glucosamine (GlcN), and xylose (Xyl). Recently, Ryan et al.[23] published a thorough MS analysis of oligosaccharides released by mild acid hydrolysis of LPG (which they called lipoglycan) derived from a laboratory *T. vaginalis* strain (B7RC2). By combining these results with those of methylation analysis of partially depolymerized LPG, these authors were able to propose a new model of the LPG structure. Building upon this model, we have used extensive NMR analysis to refine and correct some aspects of the LPG structure. In particular, we propose the Gal in the core (CPI-GC) to have predominantly α-anomeric configuration, in contrast to the previously suggested β-configuration. In addition, we detected the presence of major amounts of phosphoethanolamine and phosphocholine in CPI-GC. Finally, we propose a novel reducing end motif, comprising a →4)-α-GlcA-(1→4)-*myo*-inositol (Ins) linkage.

Understanding the chemical nature of this major immunocompetent molecule will have important implications for the development of effective therapeutic drugs to treat human trichomoniasis.

2. Results

2.1. NMR analysis of nitrous acid-delipidated CPI-GC (CPI-GC-N)

For structural analysis of the glycan portion of CPI-GC, we removed its lipid portion by treatment with nitrous acid[13,15] to achieve sufficient solubility in aqueous solution. The released lipid was removed by Folch extraction, and the resulting glycan core, which was named "CPI-GC-N" (N stands for *N*itrous acid) was subjected to methylation analysis and 1- and 2-D NMR analysis.

The methylation analysis of the glycan core revealed a complex mixture of linkages (Table 1). The predominant linkages were terminal and 3-linked Gal and 3-linked and 4-linked GlcN (the linkage analysis does not distinguish between GlcNAc and GlcN). Rha was present in several different linkages, including terminal, 2-linked, 3-linked, 2,3-linked, and 2,4-linked. All monosaccharides had the pyranose ring form.

The 1-D proton spectrum (Fig. 1) showed two large peaks in the α-anomeric region, one major broad peak in the β-anomeric region, and several minor anomeric peaks in between. All of the major anomeric peaks had irregular line shapes, pointing to a structure with

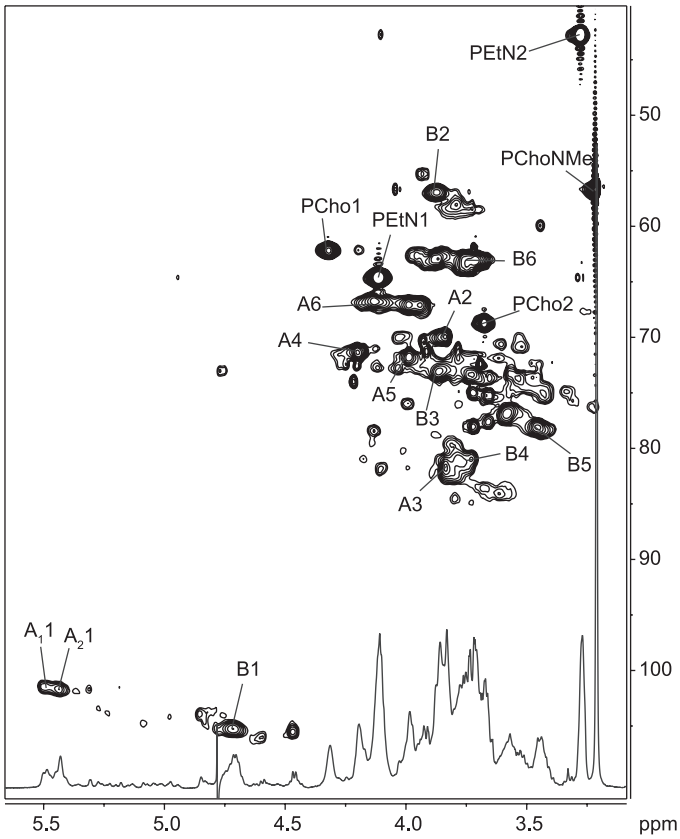

Fig. 1. Partial 1D proton and 2D $^{1}H–^{13}C$- HSQC spectrum of CPI-GC-N. Only the major residues belonging to the [3)-α-Galp-(1→4)-β-GlcNAc-(1→] repeats, as well as the PEtN and PCho substituents, are labeled. The numbers indicate the protons/carbons within a given residue, not the different residues.

Table 1
Results of methylation analysis of CPI-GC-N and CPI-GC-F in per cent GC–MS peak area

Residue	CPI-GC-N	CPI-GC-F
Terminal Rha*p*	4.9	2.0
Terminal Xyl*p*	5.7	1.4
2-linked Rha*p*	8.1	9.6
3-linked Rha*p*	5.9	7.1
Terminal Glc*p*	4.6	1.2
Terminal Gal*p*	16.3	2.3
3,4-linked Rha*p*	NF	0.8
2,3-linked Rha*p*	4.0	9.1
3-linked Glc*p*	2.3	7.1
2,4-linked Rha*p*	NF	6.7
2,3,4-linked Rha*p*	NF	2.5
3-linked Gal*p*	11.5	12.6
4-linked Glc*p*	3.8	0.2
3,4-linked Gal*p*	3.2	2.8
Terminal Glc*p*NAc	5.6	7.2
4-linked Glc*p*NAc	7.3	2.1
3-linked Glc*p*NAc	12.6	17.3
6-linked Glc*p*NAc	1.7	5.5
3,4-linked Glc*p*NAc	2.5	2.3

NF: not found.

a very complex repeating unit or no repeating unit at all. The ring region extended from 4.4 to 3.4 ppm. Just upfield from this region we observed two large singlets whose proton and carbon chemical shifts (from HSQC) were in agreement with phosphocholine (PCho)[24] and phospoethanolamine (PEtN)[25] diesters, as opposed to monoesters.[26,27] Around 2 ppm were found several overlapping N-acetyl peaks belonging to the GlcNAc residues and between 1.4 and 1.2 ppm a number of overlapping 6-deoxysugar methyl signals (not shown).

In the 2-D COSY spectrum of CPI-GC-N, at least 20 separate spin systems were discernible. Using TOCSY, HSQC (Fig. 1), and NOESY spectra, 14 of these could be assigned to specific monosaccharides. The partial assignment can be found in Table 2. The residues α-Gal and α-Rha were clearly discernible by their TOCSY patterns. Gal typically shows three cross peaks correlated with H-1 because of the large vicinal coupling constants of H-1 through H-4 and the small coupling constant between H-4 and H-5. Also typical for α-Gal is the rather low field chemical shift of H-4. Two spin systems, with H-1 at 5.51 and 5.45 ppm, showed an α-galacto pattern in the TOCSY spectrum.[28] The α-stereochemistry of these residues was indicated by the downfield chemical shift of their H-1 protons and by anomeric one-bond C—H coupling constants of 183 and 178 Hz, respectively.[29] The C-3 nuclei of both galactose residues resonated at 82.3 ppm indicating glycosylation in the 3-position. This was confirmed by an H-4 chemical shift of 4.20 ppm, consistent with 3-linked α-galactose. A GlcA residue was found with H-1 at 5.32 ppm and H-5 at 4.77 ppm. This residue was not found in the methylation analysis because detection of uronic acids by GC–MS requires prior carboxyl reduction, which we did not perform.

The major signal in the β-anomeric region was found, by 2-D COSY, to consist of at least two anomeric protons, both of them from