The *Giardia intestinalis* filamentous cyst wall contains a novel β(1-3)-N-acetyl-d-galactosamine polymer: a structural and conformational study

Gerrit J. Gerwig¹, J. Albert van Kuik², Bas R. Leellang², Johannis P. Kamerling², Johannes F.G. Vliegenthart¹², Craig D. Karr³, and Edward L. Jarroll³

¹Bijvoet Center, Department of Bio-Organic Chemistry, Section of Glycoscience and Biocatalysis, Utrecht University, Padualaan 8, NL-3584 CH Utrecht, The Netherlands; and ²Department of Biology, Northeastern University, Boston, MA 02115, USA

Received on March 5, 2002; revised on April 2, 2002; accepted on April 3, 2002

Assembly of a protective cyst wall by *Giardia* is essential for the survival of the parasite outside the host intestine and for transmission among susceptible hosts. The structure of the *Giardia intestinalis* filamentous cyst wall was studied by chemical methods, mass spectrometry, and ¹H nuclear magnetic resonance spectroscopy. Isolated cyst wall material contains carbohydrate and protein in a ratio of 3:2 (w/w), and the carbohydrate moiety is composed of a β(1-3)-N-acetyl-d-galactopyranosamine homopolymer. Conformational analysis by molecular dynamics and persistence length calculations of GalNAc oligomers in solution demonstrated a flexible structure consisting of left- and right-handed helical elements. It is most likely that in the solid state, the polysaccharide forms ordered helices or possibly multiple helical structures having strong interchain interactions. The highly insoluble nature of the *Giardia* cyst wall must be due to these strong interchain interactions and, probably, a strong association between the carbohydrate and the protein moiety.

**Key words:** cyst wall/*Giardia* ¹H NMR spectroscopy/mass spectrometry/polysaccharide

**Introduction**

The parasitic protozoan *Giardia intestinalis* (syn. *G. duodenalis*, *G. lamblia*) is a widespread human pathogen, being a major cause of enteric disease in the world (Craun, 1990). Human giardiasis ranges from an asymptomatic, self-limiting disease to a chronic infection (Wolfe, 1990). Signs and symptoms include flatulence, diarrhea (often with steatorrhea), epigastric pain and distress, nausea, and weight loss (Ferguson *et al.*, 1990). Transmission of giardiasis often occurs by fecal–oral routes via infected water. The life cycle of *Giardia* consists of two morphologically distinct forms: (1) the trophozoite that is responsible for pathogenesis within the gastrointestinal tract, and (2) a dormant cyst that serves in the transmission of the parasite to new hosts (Adam, 1991). *Giardia* cysts, produced in response to bile, measure approximately 8 × 12 μm and contain a pair of trophozoites. The cyst wall is composed of an inner (double) membranous and an outer (0.3 μm thick) filamentous portion (Erlandsen *et al.*, 1996). The filaments measure 7–20 nm in diameter and are arranged in a tightly packed meshwork (Erlandsen *et al.*, 1989). The highly insoluble cyst wall is essential for survival of the parasite outside the host. Cysts display greater resistance than do trophozoites to chemical agents used for water disinfection and to chemotherapeutic agents routinely used in the treatment of giardiasis (Paget *et al.*, 1998). Consequently, encystment represents a potentially important site of chemotherapeutic attack.

For the study of the cytodifferentiation from trophozoite to cyst as well as for our study of the biogenesis of the *Giardia* cyst wall, detailed information of the molecular structure of the cyst wall material is a prerequisite. It is widely accepted that many protozoan cyst walls contain chitin, but there is little sound biochemical evidence supporting this assumption. Indeed, a series of histochemical and biochemical studies suggested that the cyst wall of *Giardia* was composed of chitin (reviewed in Jarroll and Paget, 1995). However, in a series of studies (Jarroll *et al.*, 1989; Manning *et al.*, 1992; Macechko *et al.*, 1992) focused on *G. intestinalis* and *G. muris*, it was demonstrated that the filamentous cyst walls contain carbohydrate and protein and that the major monosaccharide constituent is N-acetyl-galactosamine (GalNAc). Subsequently, these investigators have shown that GalNAc is synthesized de novo from endogenous glucose (Macechko *et al.*, 1992) via a pathway of inducible enzymes, which are transcriptionally as well as allosterically regulated (Van Keulen *et al.*, 1998; Bulik *et al.*, 2000), and that GalNAc is fixed into an insoluble polysaccharide by the action of “cyst wall synthase” (Jarroll and Paget, 1995). Considering the insolubility of the cyst wall and the abundance of GalNAc, it was suggested that the filaments are composed of a carbohydrate(polysaccharide)–protein complex. With respect to the carbohydrate component, this report presents results demonstrating conclusively that the *Giardia* cyst wall filaments are not composed of chitin but of a previously undescribed (β1-3)-linked GalNAc homopolymer. Furthermore, insight is provided into conformational aspects of this polysaccharide.

---

¹To whom correspondence should be addressed; E-mail: j.f.g.vliegenthart@chem.uu.nl

© 2002 Oxford University Press
Results

Analysis of filamentous cyst wall material

*Giardia intestinalis* cysts were homogenized, and cyst pellets were collected by centrifugation, then treated with sodium dodecyl sulfate (SDS), amyloglucosidase, papain, DNase, RNase, and proteinase K resulting in an unsoluble fine-powdered white material, denoted cyst wall material CWM.

Gas-liquid chromatography (GLC) analysis after methanolyis of CWM and subsequent derivatization (re-N-acetylation, then trimethylsilylation) revealed D-GalNAc to be the only monosaccharide constituent. The total carbohydrate content amounted 63% (w/w). GLC analysis after acid hydrolysis of CWM and subsequent derivatization (trimethylsilylation) indicated the presence of galactosamine (GalN). Protein determination of CWM revealed a protein content of 36% (w/w). Taking into account the carbohydrate and protein content of CWM and the fact that CWM is hydrolyzable under the methanolysis/hydrolysis conditions used (Kamerling and Vliegenthart, 1989), a high percentage of GalN as native monosaccharide constituent in CWM can be ruled out. The identity of the N-acetyl group as N-acetyl was demonstrated by 1H nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) (*vide infra*). Therefore it can be assumed that the carbohydrate part of CWM is a polysaccharide built up (nearly completely) from GalNAc residues.

To generate soluble material, the highly insoluble CWM was partially hydrolyzed (PH) by repeated short treatments with 0.5 M trifluoroacetic acid (TFA). The mixture obtained (CWM PH) was fractionated by gel filtration chromatography (Figure 1) yielding five fractions, denoted CWM PH 1–5, which were rechromatographed to obtain better defined fractions. Fraction CWM PH 5 contained free GalNAc according to monosaccharide and 1H NMR analysis. MALDI-TOF-MS analysis of fraction CWM PH 2 (Figure 2) only showed the presence of the pseudomolecular ion [M+Na]+ of HexNAc2 at m/z 446.9. After reduction (CWM PH 4R), monosaccharide analysis revealed the occurrence of GalNAc and GalNAc-ol in the molar ratio of 0.9:1.0, whereas methylation analysis demonstrated the presence of terminal GalNAc and 3-substituted GalNAc-ol.

The 1H-NMR spectrum of CWM PH 4R, depicted in Figure 3A, was completely assigned using 2D total correlation spectroscopy (TOCSY) and rotating frame nuclear Overhauser effect spectroscopy (ROESY) measurements [GalNAc(β1-3)GalNAc(β1-3)GalNAc-ol]. Based on these results, fraction CWM PH 3R contains mainly GalNAc(β1-3)GalNAc(β1-3)GalNAc-ol. MALDI-TOF-MS analysis of fraction CWM PH 2 (Figure 2) showed the pseudomolecular ions [M+Na]+ of oligosaccharides ranging from HexNAc2 (m/z 852.8) up to HexNAc9 (m/z 1868.0) (Δ m/z = 203). Methylation analysis of fraction CWM PH 2 showed besides terminal GalNAc, only 3-substituted GalNAc.

Monosaccharide analysis of fraction CWM PH 1 demonstrated a carbohydrate (GalNAc) content of 40% (w/w), whereas methylation analysis again showed only terminal GalNAc and 3-substituted GalNAc. The protein determination of CWM PH 1 showed a protein content of 60% (w/w). In the MALDI-TOF-MS spectrum of CWM PH 1 (Figure 2) a series of pseudomolecular ions [M+Na]+ of oligosaccharides ranging from HexNAc2 up to HexNAc9 (m/z 4710.2) was observed. Because chain cleavage occurred during partial hydrolysis, the degree of polymerization in the native cyst wall material may be much higher.

The 1D 1H-NMR spectrum of CWM PH 1 (carbohydrate region) is shown in Figure 3B, and the included proton assignments were obtained through analysis of 2D TOCSY and nuclear Overhauser effect spectroscopy (NOESY) NMR spectra. According to the various data, it can be concluded that the carbohydrate...
moiety of CWM is composed of a (β1-3)-linked GalNAc homopolymer. The structural analysis has been performed on GalNAc oligosaccharides released by partial acid hydrolysis, but in view of the high protein content in CWM PH 1, it is likely that a covalent linkage between protein and GalNAc oligomers occurs in CWM.

Conformational analysis of the carbohydrate moiety

To gain insight into the conformational aspects of the GalNAc polysaccharide, especially in relation to its insolubility, molecular dynamics (MD) calculations with the constituting disaccharide GalNAc(β1-3)GalNAc were performed. The results were used to evaluate possible conformations of polysaccharide structures.

To explore the preferred conformations of the glycosidic linkages $\phi$ (O5′-C1′-O3-C3) and $\psi$ (C1′-O3-C3-C2) (IUPAC-IUB Joint Commission on Biochemical Nomenclature, 1983), a relaxed energy map of the methyl glycoside of the disaccharide in vacuum was constructed with the use of the CHEAT force field (Grootenhuis and Haasnoot, 1993; Kouwijzer and Grootenhuis, 1995). This map (Figure 4A) showed two conformational regions with a low energy, one ($\phi, \psi = -60, -120$) is...
denoted region I, and the other (ψ, ϕ = -60, 60) is denoted region II. Region I represents the global lowest energy. Next, minimal energy conformations selected from regions I and II were minimized again, this time with fully relaxed interglycosidic dihedral angles.

The conformational aspects of the disaccharide in solution were studied by two MD calculations in water, performed with the GROMOS force field (Spieser et al., 1999). One run was started with the minimum energy conformation obtained from region I and was continued for 5 ns. The other simulation started with the lowest energy conformation from region II and had a duration of 3.5 ns. During this last simulation, the ψ dihedral angle changed after 500 ps simulation time from 60° (region II) to the global minimum at -120° (region I). The GalNAc(β1-3)GalNAc interglycosidic dihedral angles ϕ and ψ for this run are shown in Figure 4A, superimposed over the energy contour map. From this map it appears that the global energy minimum consists of two subminima, denoted region Ia and region Ib. The MD simulation that was started from region I did not visit region II. This run also showed the two regions Ia and Ib.

To assess the value of the computer simulation results, the 1H NMR data of fraction CWM PH I were used. The NOESY spectra contained several medium to strong cross-peaks that in many cases were intraresidual. The H1'-H4 interaction would be detectable for the regions Ib (strong) and II (weak). Its absence however, demonstrated the dominant role of the Ia conformation. Signal overlap prevented a straightforward quantitative analysis of the NOEs.

To determine the population distribution of conformations more precisely, the method of adaptive umbrella sampling of the potential of mean force (Hooft et al., 1992) was used. With this method, free energy differences are calculated between conformations in relation to the interglycosidic dihedral angle ψ. From this information, the population distribution of the conformations can be determined and this result is shown in Figure 4B. Inspection of this distribution profile shows that the GalNAc disaccharide only occurs in the global lowest energy conformation (region I). This energy minimum can be subdivided into two subminima, which are populated in a ratio of 4:1. Although the two conformations do not differ a great deal with respect to the interglycosidic dihedral angles, this small difference can have a significant effect on the total conformation of the GalNAc polysaccharide.

To get an impression of the conformation of this polysaccharide in solution, some polysaccharide models with a size of 30 monosaccharide units were constructed from disaccharides and linkage information obtained from the MD runs. Polysaccharide models constructed from disaccharide residues collected from region Ia form a right-handed helical structure. The average size of a helical turn is approximately seven residues. In contrast, polymer models constructed from disaccharide residues and linkages collected from region Ib gave a left-handed helical structure with an average helical-turn size of about 8.4 residues. A polysaccharide model chain, constructed from disaccharide structures and linkages selected from both regions in a ratio of Ia:Ib = 4:1, showed a random coil with a much more folded (less rodlike) conformation than observed for the two previous polysaccharide models (Figure 5).

To estimate the flexibility and to compare the stiffness of the three polysaccharide model structures mentioned, the persistence length (Kroon-Batenburg et al., 1997) for all three polysaccharide model chains was calculated. The persistence length is the mean length of projection of a polymer chain along the direction of the first linkage and is a measure of the stiffness of the polymer. It is larger for stiffer chains. Polymers built exclusively from regions Ia or Ib have a persistence length of 52 and 80 Å, respectively. Calculations on a mixture of both regions in a ratio of Ia:Ib = 4:1, yielded a persistence length of 35 Å. Thus the combination of two different conformations results in a much more flexible chain. This flexibility is in agreement with the finding that polysaccharide fragments with a size of 23 monosaccharide units can still be dissolved in water. Compare, for example, cellulose, which has a persistence length in water in the order of 90–125 Å (Kroon-Batenburg et al., 1997), and a largest dissolving fragment of nine monosaccharides. From these findings, it is evident that the highly insoluble character of the cyst wall cannot be accounted for by the conformational aspects of a single GalNAc polysaccharide chain but must be the result of strong interchain interactions. It is likely that the polysaccharide forms ordered helices in the solid state, possibly forming multiple helical structures, as is observed for similar polysaccharides, for example, (1-3)-linked β-D-GlcNAc (Kouwijzer and Perez, 1998; Rao et al., 1998).

**Discussion**

The precursor for the GalNAc polymer is UDP-GalNAc, which is synthesized from endogenous glucose by inducible enzymes during encystment (Macechko et al., 1992; Van Keulen et al., 1998; Bulik et al., 2000). Evidently, GalNAc is polymerized into a β(1-3)-D-GalpNAc homopolymer by an
enzyme activity tentatively referred to as “cyst wall synthase” (Jarroll and Paget, 1995) and incorporated into an insoluble material resembling *Giardia*’s cyst wall filaments.

In our conformational studies, two preferential structures could be identified for GalNAc(β1-3)GalNAc. A 4:1 combination of these two conformations resulted in a flexible, random-coil model for the GalNAc polymer. From the conformational studies, it is clear that the highly insoluble nature of the cyst wall does not stem from the conformational properties of a single GalNAc polysaccharide chain but must be due to strong interchain interactions. The potential covalent linkage between the GalNAc polymers and the detected protein could also play a role. This is currently the subject of further study.

Of biological interest here is the fact that the cyst wall filaments of *Giardia* are not composed of chitin but rather of a novel polysaccharide for which no known degradative enzyme has been found as yet. Because many of the microorganisms that inhabit the intestines where these cysts are formed can produce chitinase, there could be a clear evolutionary advantage to *Giardia* not having its protective cyst wall composed of chitin. It is also interesting to speculate that since the cyst wall of *Giardia* is made of this novel GalNAc homopolymer, then there must be some hydrolase(s) capable of degrading it. Whether this hydrolase exists and is involved in the cytodifferentiation called excystation (trophozoites emerging from the cyst) is unknown. It is possible that excystation might occur simply by proteolysis of the protein present in the cyst wall as suggested by Ward et al. (1997). However, it is also likely that there exists a hydrolase capable of degrading this polymer, otherwise the amount of this material in the ecosystem would tend to accumulate irreversibly.

Materials and methods

*Isolation and purification of Giardia cyst wall material*

Approximately 40 L of *G. intestinalis* (strain MR4) trophozoites were grown in axenic culture and encysted for 24–48 h as described elsewhere (Macechko et al., 1992). Cysts were collected by centrifugation of encystment medium at 750 × g for 10 min at 25°C and subjected to homogenization to disrupt trophozoites and partially encysted cells. Cyst pellets were washed five times in 10 ml distilled water and collected at 500 × g. Cyst wall filaments preparations were made by a modification of the procedure of Manning et al. (1992). The pooled sample was divided into two equal portions and processed separately. Each cyst pellet was heated to 100°C in an equal volume of 10% SDS for 5 min.

The cysts were concentrated by centrifugation at 12,000 × g for 5 min (as in subsequent steps), and the supernatant was discarded. The cyst pellet was washed in 10 ml distilled water five times until there were no visible signs of the detergent. The resulting cyst pellet was resuspended in 3 ml of amylglucosidase buffer (20 mM acetate buffer, pH 4.5), transferred to a 20-ml glass screw-capped scintillation vial with a flat bottom and a stir bar. One hundred units of amylglucosidase (EC 3.2.1.3, Sigma A-3514, St. Louis, MO) was added to the vial that was placed on a stir plate in a 55°C warm air incubator for 60 min. This procedure was repeated once more, and each time the supernatant was discarded.

After the two amylglucosidase treatments, the SDS treatment was performed once again. The resulting cyst pellet was washed in 10 ml of freshly made papain buffer (100 mM phosphate buffered saline [PBS], pH 7.2, with 5 mM cysteine and 5 mM ethylenediamine tetra-acetic acid [EDTA]). The pellet was resuspended in 3 ml papain buffer and again placed in a glass scintillation vial with a stir bar. The sample was stirred with 50 U of papain (EC 3.4.22.2, Sigma P-4762) for 3 h at 60°C. This procedure was repeated once more, and each time the supernatant was discarded. Again the resulting cyst pellet was collected by centrifugation and washed, as in step one, with 10% SDS and water.

The pellet resulting from the papain treatment was subjected to DNase (EC 3.1.21.1, Sigma DN-25) and RNase (EC 3.1.27.5, Sigma R-4875) treatments. The pellet was resuspended in 3 ml of XNase buffer (0.1 M PBS, pH 7.2, with 5 mM MgCl₂) in a glass scintillation vial with a stir bar. The sample was incubated with 500 Kunitz units of DNase and 5 Kunitz units of RNase stirring for 2 h at 37°C. This treatment was repeated once more and the pellet washed with amylglucosidase buffer. At this point, all subsequent centrifugation steps were carried out at 44,000 × g for 10 min at 10°C. Following the DNase/RNase treatment, the pellet was subjected to another amylglucosidase treatment as described and then to a proteinase K (EC 3.4.21.14, Sigma P-0390) treatment. The pellet was washed once in proteinase K buffer (50 mM Tris, pH 8.0, 0.2% Triton X-100, 1 mM CaCl₂). The pellet was resuspended in the proteinase K buffer, placed in a glass scintillation vial with 50 U proteinase K, and stirred for 2 h at 60°C. This procedure was repeated once more, and the pellet was washed with an equal volume of 10% SDS containing 100 mM dithiothreitol (DTT) for 5 min at 100°C. The pellet was then washed five times in distilled water to remove the detergent and DTT. The resulting CWM was lyophilized for 48 h, yielding approximately 150 mg dry fine-powdered white material.

*Monosaccharide analysis and protein determination*

Monosaccharide analysis after methanolysis, re-N-acetylation, and trimethylsilylation, including absolute configuration determination using (-)-2-butyl glycosides, was performed by GLC-MS as described (Gerwig et al., 1979; Kamerling and Vliegenthart, 1989). Protein determinations were carried out using a Micro BCA™ Protein Assay Reagent Kit (Pierce).

*Acid hydrolysis*

Dried CWM (1 mg) and GalNAc (1 mg) were each treated with 1 ml of 2 M HCl for 3 h at 100°C and with 1 ml of 4 M TFA for 5 h at 100°C. After lyophilization, the residues were dissolved in 1 ml H₂O. Aliquots of 25 µl of the four samples were taken, lyophilized, and trimethylsilylated with 25 µl pyridine/hexamethyldisilazane/chlorotrimethylsilane 5:1:1 (v/v) for 30 min at room temperature. Of each sample, 1 µl was analyzed by GLC (Chrompack CP 9002 gas chromatograph with CP Sil 5CB column, 25 m × 0.32 mm); temperature program: 140–240°C at 4°C/min).

*Partial acid hydrolysis*

Dried CWM (8 mg) was treated with 0.5 ml of 0.5 M TFA for 15 min at 100°C. After centrifugation, the supernatant was collected and the sediment was treated again with 0.5 ml of 0.5 M TFA for 15 min at 100°C. This procedure was repeated...
2D ROESY experiments were carried out at 280 instrument. The spectra were processed using locally developed and 125 ms were recorded at 500 MHz on a Bruker DRX-500 (Vliegenthart et al., 1989). 2D TOCSY spectra with mixing times of 12, 25, 41, 50, and 125 ms were recorded at 50 MHz on a Bruker DRX-500 instrument. The spectra were processed using locally developed NMR software (Van Kuik, Bijvoet Center, Utrecht University). 2D ROESY experiments were carried out at 280 K with mixing times of 100, 200, 300, and 500 ms. The HOD signal was presaturated for 1 s during the relaxation delay.

**Minimum energy calculations**

Minimum energy calculations in vacuo were performed with CHEAT (Grootenhuis and Haasnoot, 1993; Kouwijzer and Grootenhuis, 1995), a CHARMM-based force field for carbohydrates in which hydroxyl groups are represented by extended atoms to prevent the formation of intramolecular hydrogen bonds and whereby energies are independent of the hydroxyl group orientation. A relaxed energy map was generated for the GalNAc(β1-3)GalNAc disaccharide by changing the inter-residual torsion angles in steps of 10° resulting in a 36 × 36 grid. For each grid point calculation, three intraglycosidic dihedral angles per monosaccharide ring and both interglycosidic angles are kept fixed by setting constraints during a first minimization. Next, all constraints in the ring were removed and the molecule was minimized a second time. Contour levels were plotted in steps of 1 kcal/mol from the global minimum.

**Molecular dynamics calculations**

Molecular dynamics simulations were carried out using the GROMOS program package (Van Gunsteren, 1987) and the updated carbohydrate force field for GROMOS (Spiesser et al., 1999) on local PentiumII class computers running Linux. Each molecule was surrounded by SPC/E (Berendsen et al., 1984) water molecules and placed in a truncated octahedral periodic box. All bond lengths were kept fixed using the SHAKE procedure (Ryckaert et al., 1977). A cut-off radius of 0.8 nm and a time step of 2 fs were used. Simulations were performed with loose coupling to a pressure bath at 1 atm and a temperature bath at 300 K (Berendsen et al., 1984) with time constants of 0.5 and 0.1 ps, respectively. To confirm calculated data, (β1-3)-linked GalNAc oligosaccharide fragments with DP >5 and DP >20 were analyzed by 1H NMR spectroscopy.

**Potential of mean force calculations on the GalNAc-GalNAc ψ glycosidic linkage**

To calculate the free energy differences of the conformations around the GalNAc(β1-3)GalNAc glycosidic ψ angle, potential-of-mean-force calculations (Hooft et al., 1992) were run for the methyl β-glycoside of the disaccharide with the GROMOS force field. All simulations were divided into jobs of 10 ps. The ψ values were collected into 72 classes, each with a width of Δψ = 5°. The derivatives were fitted to a 12-term Fourier series. The first 0.2 ps of each job were discarded.

**Persistence length calculations**

The persistence length calculations of the GalNAc polymer were performed with data selected from appropriated parts of the MD simulations of the GalNAc(β1-3)GalNAc disaccharide (Kroonen-Batenburg et al., 1997). For each calculation, 1000 frames were selected and used to create chains of 300 residues.

**Acknowledgments**

We acknowledge partial support for this project from the NIH AI 41230 and the Netherlands Organisation for Scientific Research (NWO).
Abbreviations

CWM, cyst wall material; DTT, dithiothreitol; EDTA, ethylenediamine tetra-acetic acid; GalNAc, N-acetyl-D-galactosamine; GalN, D-galactosamine; GLC, gas-liquid chromatography; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; MD, molecular dynamics; MS, mass spectrometry; NMR, nuclear magnetic resonance; NOE/SY, nuclear Overhauser effect spectroscopy; PBS, phosphate buffered saline; PH, partially hydrolyzed; ROESY, rotating frame nuclear Overhauser effect spectroscopy; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid; TOCSY, total correlation spectroscopy.

References


G. intestinalis cyst wall contains a β(1-3)-GalNAc polymer