

# Molecular Structure and Rheological Properties of Short-Side-Chain Heavily Glycosylated Porcine Stomach Mucin

Gleb E. Yakubov,<sup>\*,†</sup> Aristeidis Papagiannopoulos,<sup>‡</sup> Elodie Rat,<sup>†</sup> Richard L. Easton,<sup>§</sup> and Thomas A. Waigh<sup>‡</sup>

Unilever Corporate Research, Colworth Park, Sharnbrook, Bedfordshire MK44 1LQ, United Kingdom, Biological Physics, Department of Physics and Astronomy, University of Manchester, Manchester M60 1QD, United Kingdom, and M-Scan Limited, Millars Business Centre, Wokingham, Berkshire RG41 2TZ, United Kingdom

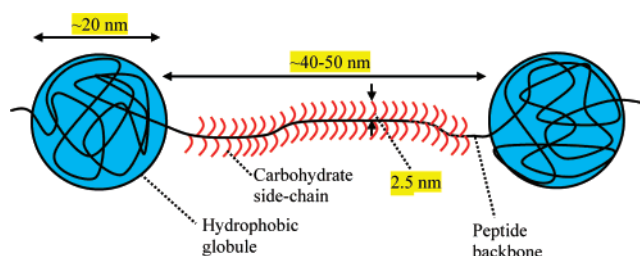
Received May 31, 2007; Revised Manuscript Received August 7, 2007

The current accepted model for high-molecular-weight gastric mucins of the MUC family is that they adopt a polydisperse coil conformation in bulk solutions. We develop this model using well-characterized highly purified porcine gastric mucin Orthana that is genetically close to the human MUC6 type. It has short side chains and low levels of sialic acid residues and includes minute amounts of cysteine residues that, if abundant, can be responsible for the self-polymerization of mucin. We have established that the mucin structure in bulk solutions corresponds to a daisy-chain random coil. Dynamic light scattering experiments probe the internal dynamics of globular subunits (individual daisies) at the  $\sim 9$  nm length scale, whereas viscosity and light scattering measurements indicate that the size of the whole mucin chains is much larger,  $\sim 50$  nm. The bulk viscosity ( $\eta$ ) scales with mucin concentration ( $c$ ) in a manner similar to that found for short-side-chain synthetic comb polyelectrolytes and is characterized by a transition between semidilute ( $\eta \sim c^{1/2}$ ) and entangled ( $\eta \sim c^{3/2}$ ) regimes.

## 1. Introduction

A number of comprehensive reports suggest that the high-molecular-weight gastric and salivary mucins of the MUC2, MUC5AC, MUC5B, and MUC6 types adopt a polydisperse coil conformation in bulk solutions.<sup>1,2</sup> Here we report the development of this established model using well-characterized highly purified short-side-chain porcine gastric mucin similar to the human MUC6 type. “Orthana” mucin is characterized by short side chains and low levels of sialic acid residues and includes minute amounts of the cysteine residues that are responsible for the self-polymerization of mucin. We have established that the structure in bulk solutions is found to correspond to a daisy-chain morphology, with the majority of the molecules adopting a dumbbell-like configuration (two globules per chain). Most probably this globular structure is determined by the interplay between electrostatic and hydrophobic interactions inside the nonglycosylated subunits of the mucin molecule and occurs when the energies of these interactions are comparable with the surface free energy of the globules, which is mainly due to the presence of extensively hydrated comb segments.<sup>3–5</sup> The hydrophobic/hydrophilic sequence of the protein chain is able to split the random coil into a set of smaller weakly charged globules (daisies) connected by the strings—a daisy-chain or necklace morphology (Figure 1). The bulk viscosity of the mucins scales with increasing concentration and is found to be characterized by a transition between semidilute and entangled regimes in a similar manner to previous experiments with short-side-chain synthetic comb polyelectrolytes.<sup>6</sup>

The role of mucins in biological systems is diverse. One of the important uses of mucins is to provide biolubrication in many



**Figure 1.** Schematic diagram of a dumbbell mucin molecule showing two globular structures per chain separated by a heavily glycosylated spacer. The model is also corroborated by atomic force microscopy (AFM) and transmission electron microscopy (TEM) experiments.<sup>5</sup> Approximate sizes are shown from a combination of AFM, TEM, DLS, and SLS results.

biological fluids including<sup>7–10</sup> saliva, tears, sweat, gastric secretions, etc. Mucins form gels at high concentrations<sup>11</sup> or low values of pH<sup>7,12–15</sup> due to the formation of disulfide bonds between the thiol groups of their cysteine residues. The detailed compositions and structures of mucins are highly variable and depend on their origin,<sup>16,17</sup> and hence various models of mucin conformation in solution have been reported in the literature. The observed differences in conformation could be due to the variability in the molecular weights chosen, the methods of preparation,<sup>18,19</sup> the isolation/fractionation procedures, or the process of sample handling. The possible presence of various contaminants (e.g., lipids<sup>20</sup>) could also be important.

There are a number of credible models for the structures of mucins that include: large aggregates of mucin monomer units that take the form of extended rods or threads<sup>7,21</sup> of roughly 5 nm in diameter and range in length from 100 to 5000 nm, cross-linked tertiary gel structures formed via disulfide bridges,<sup>22</sup> hydrophobic<sup>12</sup> or carbohydrate-to-carbohydrate bonds,<sup>23</sup> a random coil of linear flexible glycoprotein within a spheroidal solvent domain with no branching,<sup>18,19,24</sup> a semiflexible random

\* Author to whom correspondence should be addressed. E-mail: gleb.yakubov@unilever.com.

<sup>†</sup> Unilever Corporate Research.

<sup>‡</sup> University of Manchester.

<sup>§</sup> M-Scan Limited.

coil that is stiffened by glycosylated fragments where the structure of these oligosaccharide side chains has a significant impact on the final structure of the mucin in solution,<sup>25</sup> a “zipper-like” engagement of bottle-brush branched mucins,<sup>23,26</sup> a nematic liquid crystalline structure,<sup>18</sup> and an anisotropic extended conformation at pH < 4.<sup>17</sup>

However, the biochemistry of mucins displays some common features; mucin is composed of a linear peptide backbone and radially arranged (comblike) oligosaccharide chains (50–80% by weight).<sup>7,10,12,20</sup> The glycosylated regions are enriched with serine or threonine, which are linked to oligosaccharides (typically less than 20 sugar units) via bonding with GalNAc (O-glycosylation),<sup>7,17,20,27</sup> while the adjacent amino acids usually consist of small residues such as proline, glycine, and alanine.<sup>20</sup> The unglycosylated regions, which are normally found at both C- and N-termini, contain a large number of cysteines and charged amino acids.<sup>7,20,28</sup>

Another unique property of mucins, which is closely associated with their rheological and lubricating behavior, is their gel-forming ability.<sup>11</sup> Previous studies have shown that the aggregation of mucins depends upon several parameters, including type, purity, length of oligosaccharides side chains,<sup>29</sup> and concentration of mucin,<sup>13</sup> as well as physiological environments such as pH, ionic strength,<sup>7,12–15</sup> and the presence of Ca<sup>2+</sup> ions.<sup>30</sup> Intra- and/or intermolecular disulfide bonding have been suggested as the main driving force for such behavior, because the treatment of mucins with cysteine-specific enzymes or reducing agents such as dithiothreitol (DTT) prevents the aggregation of mucins.<sup>7,14</sup> At low concentrations, mucins generally do not exhibit such extensive aggregation, regardless of pH and ionic strength.<sup>8</sup> However, it was shown<sup>31</sup> that mucin exhibited effective boundary lubricating properties even at low concentrations, where gel formation is generally not possible.

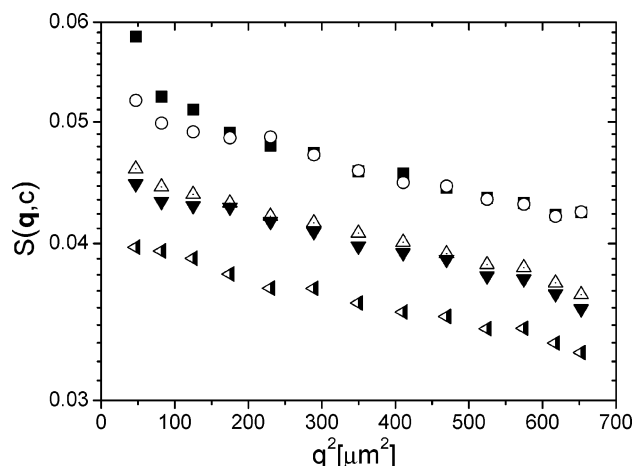
This is the first of two articles that considers the molecular structure and rheological properties of chemically well-defined stomach mucins. The second article examines the charge and interfacial adsorption of the molecules.<sup>5</sup>

## 2. Experimental Section

**2.1. Mucin Preparation.** Pharmaceutical grade porcine gastric “Orthana” mucin was purchased from A/S Orthana Kemisk Fabrik, Kastrup, Denmark. “Orthana” mucin is used in a saliva substitute formulation, “Saliva Orthana”, and originates from the linings of pork stomach. The commercial preparation was extensively dialyzed to remove all salts and other low-molecular-weight additives and finally lyophilized and stored for use as required. All solutions were made by dissolving weighed portions of the lyophilized material in ultrapure (resistance 18.2  $\Omega$  cm) water; the sample was shaken for 2 h and subsequently filtered through a Sartorius “Minisart” filter (200 nm pore size). The solutions were used immediately after preparation.

**2.2. Rotational Rheometry.** We used a Rheometrics ARES-LS rheometer with parallel plate geometry. Plates of 50 mm in diameter and a 50  $\mu$ m gap height were used. The gap was zeroed using the method previously developed.<sup>32</sup> Samples were loaded into the geometry according to the recommendations.<sup>32,33</sup> Shear experiments were typically performed to obtain at least 10 measurement points per decade in shear rate (or stress), and a measurement time ranging from 3 to 10 s was used. The mucin solutions of various concentrations were prepared in deionized water. Seven mucin concentrations were examined in the range of 0.03–30 mg/mL, and all experiments were performed at 25 °C between 10 and 1000 s<sup>−1</sup> (Supporting Information).

**2.3. Microrheological Measurements.** Polystyrene particles (0.5  $\mu$ m carboxylic acid) were added to the mucin solutions, and their Brownian movement was monitored via absorption contrast optical



**Figure 2.** Static form factor,  $S(q, c)$ , from light scattering experiments vs  $q^2$  for 30 mg/mL mucin solutions of different ionic strengths (Zimm plot): filled squares, deionized water; open circles, 1 mM NaCl; open triangles, 10 mM NaCl; filled triangles, 154 mM NaCl; half-filled triangles, 1 M NaCl.

microscopy.<sup>35</sup> The particle tracking software (IDL) enabled the probe particles to be tracked in the acquired images in two dimensions. These two-dimensional trajectories were then converted to mean-square displacements (MSDs) and plotted against the time step. The intercept with the ordinate axis of the MSD ( $\langle r^2(x, y) \rangle$ ) plotted against time ( $t$ ) provides the diffusion coefficient ( $D$ ) of the probe spheres in a purely viscous material

$$D = \lim_{t \rightarrow 0} \frac{\langle r^2(x, y) \rangle}{4t} \quad (1)$$

All of the mucin solutions examined were found to be purely viscous, and additional optical fiber picorheology data indicates that this linear relationship between the MSD and time ( $t$ ) holds down to 10<sup>−5</sup> s. The diffusion coefficient is further related to the viscosity ( $\eta$ ) of the solution via the Stokes–Einstein equation

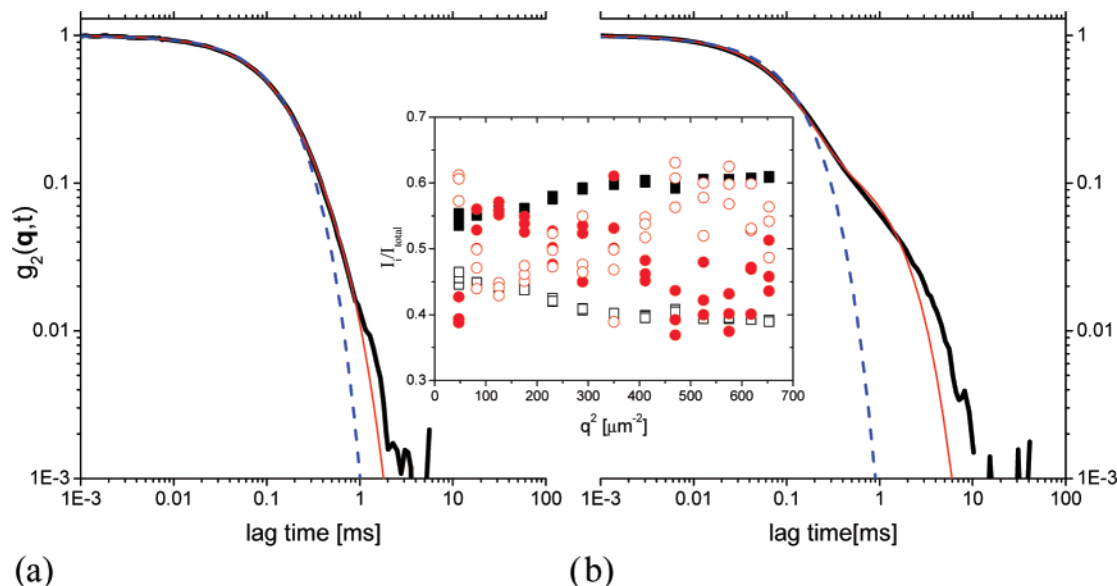
$$\eta = \frac{kT}{6\pi Db} \quad (2)$$

where  $b$  is the radius of the probe sphere and  $kT$  is the thermal energy.

Twelve solutions were prepared in phosphate-buffered saline (PBS) buffer (Sigma, U. K.) and deionized water following the protocol above (all experiments were performed at 25 °C) at a range of concentrations from 0.3 to 120 mg/mL. Sample preparation was completed by the addition of 10  $\mu$ L of a 0.3% suspension of polystyrene particles ( $b$  = 548 nm) to 100  $\mu$ L of each of the mucin solutions.

**2.4. Static Light Scattering.** Static light scattering (SLS) was performed using an ALV-5000 goniometer system (ALV GmbH, Germany). The scattered intensity was accumulated for 60 s for 11 scattering angles from 20° to 120° at a temperature of 27.8 °C. Two sets of measurements were performed to study the dependence of the mucin size on the concentration and the ionic strength. Four mucin concentrations were studied in the range of 0.3–30 mg/mL without any background electrolyte. This set of measurements was used to obtain a Zimm plot from which the radius of gyration ( $R_g$ ) of the mucin and the second virial coefficient ( $a_2$ ) that characterizes the interaction between mucin molecules in bulk solution were determined (Figure 2). The effect of the ionic strength was examined at a constant mucin concentration (30 mg/mL) and temperature in the range of 0.001–1 M NaCl in the presence of PBS buffer.

**2.5. Dynamic Light Scattering.** The measurements were performed using a PCS 4700 goniometer system (Malvern Instruments, U. K.) with dedicated software. The laser wavelength was 488 nm, and the power of the laser was adjusted to 20.4 mW to ensure consistent incident



**Figure 3.** Correlation functions from DLS experiments at an angle of 90° at (a) 1.0 and (b) 30 mg/mL and relative intensities of slow and fast modes from DLS experiments as a function of  $q^2$  from the fits to eq 3: single-exponential fit, dashed line; double-exponential fit, continuous line. Inset: Relative amplitudes ( $I/I_{\text{total}}$ ) of the two exponential fits to the correlation functions: filled circles, 1 mg/mL fast mode; open circles, 1 mg/mL slow mode; filled squares, 30 mg/mL fast mode; open squares, 30 mg/mL slow mode. All samples were held at 20 °C.

beam intensity without overheating the sample. Initial measurements were performed for 10 scattering angles between 30° and 120°. The correlation functions ( $g(t, q)$ ) were fit to the functional form

$$g(t, q) = A_{\text{fast}} e^{-t/\tau_{\text{fast}}} + A_{\text{slow}} e^{-t/\tau_{\text{slow}}} \approx A_{\text{slow}} e^{-t/\tau_{\text{slow}}} \left( 1 + \frac{A_{\text{fast}}}{A_{\text{slow}}} e^{-t/\tau_{\text{fast}}} \right) \quad (3)$$

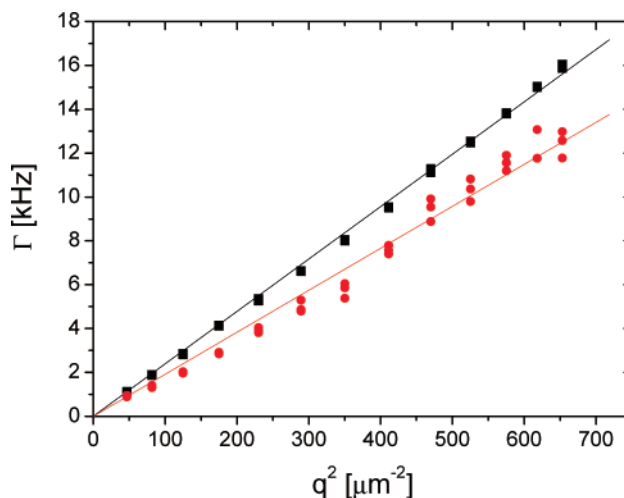
where  $A_{\text{fast}}$  and  $A_{\text{slow}}$  are the amplitudes of the fast and slow modes, respectively.  $\tau_{\text{fast}}$  and  $\tau_{\text{slow}}$  are the fast and slow relaxation times (Figure 3).  $q$  is the momentum transfer given by  $q = 4\pi/\lambda \sin \theta/2$ , where  $\lambda$  is the wavelength and  $\theta$  is the scattering angle. The relative amplitudes of the two modes ( $A_{\text{fast}}$  and  $A_{\text{slow}}$ ) are shown in the inset of Figure 3. The approximate equality in eq 3 is shown to highlight the applicability of the double-exponential approximation to a two-bead Zimm model where  $\tau_{\text{slow}} \gg \tau_{\text{fast}}$ .<sup>34</sup> Therefore,  $\tau_{\text{slow}}$  corresponds to the center of mass diffusion of a single mucin molecule, and  $\tau_{\text{fast}}$  is the first internal Zimm relaxation mode (also the rescaled monomer time in our two-bead model).

A clear linear dependence of the fast relaxation rate ( $\Gamma$ ) plotted against the square of the scattering vector ( $q^2$ ) was established, so four representative scattering angles were selected (30°, 60°, 90°, and 120°) to simplify the data analysis (Figure 4). The slow relaxation rate was predominately diffusive with a slightly higher than  $q^2$  dependence at low concentrations (Figure 5).

Mucin samples were injected into NMR tubes through a Minisart sterile filter with a pore size of 0.2  $\mu\text{m}$  to avoid dust contamination of the solution. The filter was chosen to reduce the binding of the mucin molecules to the filter material.

Different mucin solutions in deionized water (low-salt limit) were measured by dynamic light scattering (DLS). Mucin concentrations were examined in the range of 0.01–30 mg/mL and temperatures in the range of 12.2–60 °C (Figure 6 and 7). Mucin samples were also diluted in NaCl aqueous solutions, while the mucin concentration was kept constant (30 mg/mL) to ensure sufficient scattering intensity at similar concentration values to the rheological investigations. Salt concentrations in the range of 0.001–1 M NaCl and temperatures in the range of 12.2–60 °C were studied (Figure 8 and 9).

A broad range of pH's was examined using PBS solutions. The buffer ensured that the ionic composition did not change by a large amount

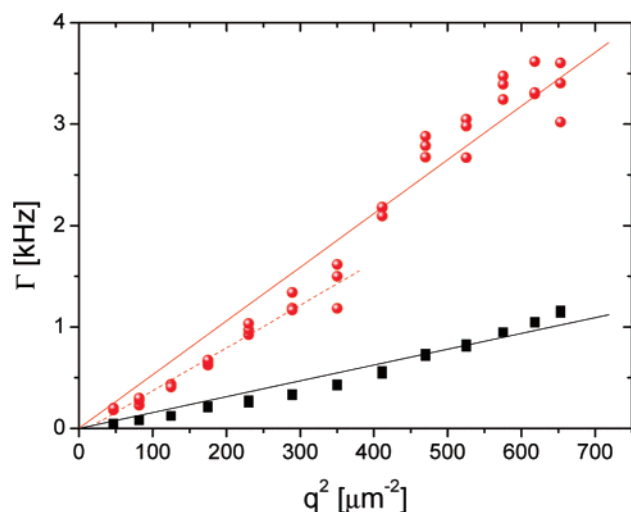


**Figure 4.** Linear dependence of the relaxation rate ( $\Gamma$ ) on the square of the scattering vector ( $q$ ) of mucin from DLS: filled squares, 30 mg/mL fast mode; filled circles, 1 mg/mL fast mode.  $R_h$  for 1 mg/mL was 10.3 nm, and for 30 mg/mL was 8.11 nm. All samples were held at 20 °C.

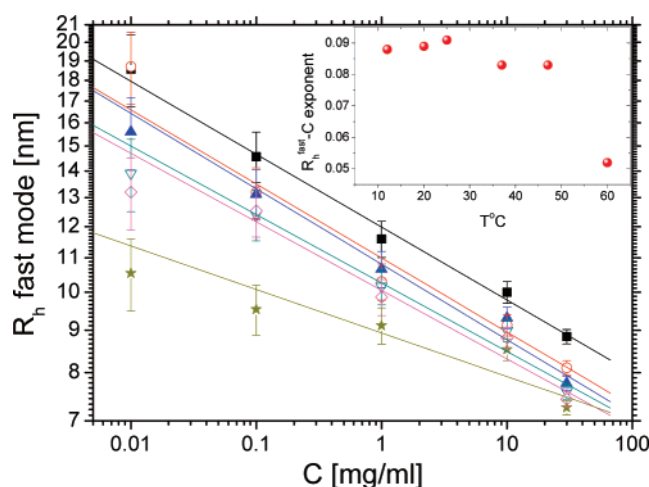
with the changing pH. The concentration of mucin was again 30 mg/mL. All experiments were performed at 25 °C. Seventeen pH's were examined in the pH range of 2.44–10.93.

The influence of  $\text{Ca}^{2+}$  ions was examined with 30 mg/mL mucin dissolved in 0.01 M  $\text{CaCl}_2$  aqueous solution.

**2.6. Monosaccharide Composition Analysis and Linkage Analysis of the O-Linked Carbohydrates.** We used both gas chromatography–mass spectrometry (GC-MS) and high-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) to examine the carbohydrate composition of the mucins. The latter method requires no subsequent derivatization once the monosaccharides are released. The glycosidic bond cleavage was done via methanolysis in conjunction with re-N-Acetylation and, in the case of the GC-MS experiments, derivatization by trimethylsilylation. An aliquot (1  $\mu\text{L}$ ) of each of the derivatized carbohydrate samples was dissolved in hexane (2 mL) and was analyzed by GC-MS using a Perkin-Elmer Turbomass Gold quadrupole mass spectrometer with an integrated gas chromatography–mass spectrometry (GC-MS).



**Figure 5.** Dependence of the slow mode on the square of the scattering vector ( $q$ ): filled circles, 1 mg/mL slow mode; filled squares, 30 mg/mL slow mode. Straight lines indicate approximately diffusive behavior.  $R_{\text{slow}}$  for 1 mg/mL was 37.2 nm (low  $q$  fit leads to 46.7 nm), and for 30 mg/mL was 123.1 nm. All samples were held at 20 °C.

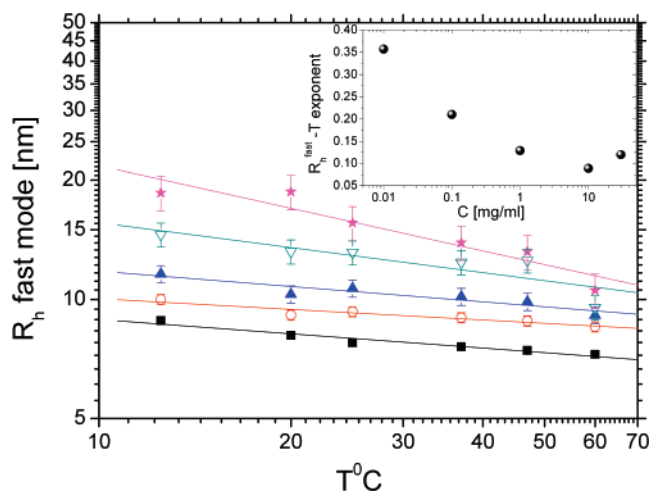


**Figure 6.** Dependence of the hydrodynamic radius from DLS experiments on the mucin concentration: filled squares, 12.2 °C; open circles, 20 °C; filled triangles, 25 °C; open triangles, 37 °C; open diamonds, 47 °C; filled stars, 60 °C. Inset shows the exponent of the power law fits.

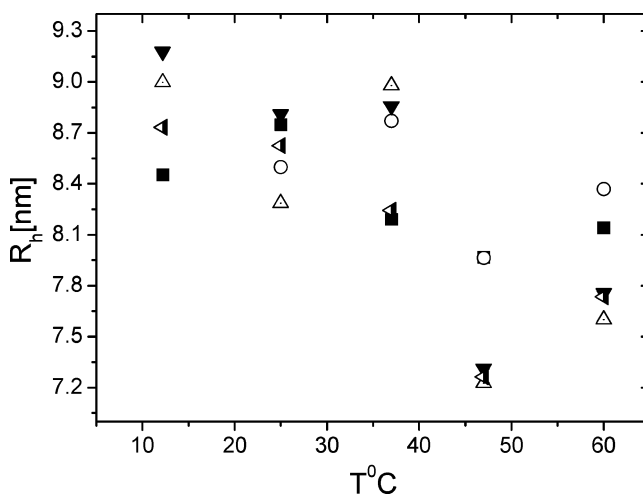
graph. The following conditions were used: column DB5; on-column injection; injector temperature, 95 °C; program, 1 min at 90 °C, then 25 °C/min to 140 °C, then 5 °C/min to 220 °C, then 10 °C/min to 300 °C, and finally 300 °C for 5 min; carrier gas, helium; mass spectrometry ionization voltage, 70 eV; ions monitored, 173 (for *N*-acetylhexosamines), 204 (for hexoses), 217 (for arabinol), and 298 (for *N*-sialic acid). On comparison of the data with that obtained from the standard mixtures containing known amounts of the expected monosaccharides, the sugars hydrolyzed from the samples were identified, and the quantity of each monosaccharide present was estimated.

Monosaccharide composition analysis by HPAEC-PAD was performed using two hydrolysis methods: One aliquot was hydrolyzed using 2 M trifluoroacetic acid (TFA) at 100 °C for 4 h, and the second aliquot was hydrolyzed using 6 M hydrochloric acid at 100 °C for 4 h. Monosaccharides released from the sample were analyzed alongside each standard mixture and a tube/reagent blank using a Dionex BioEC system. Chromeleon software, version 6.50, an AS50 autosampler with a thermal compartment, an ED50 electrochemical detector, and a GS50 gradient pump system were used.

The linkage analysis of the O-linked carbohydrates was performed with the permethylated glycan sample mixtures obtained from the



**Figure 7.** Dependence of the hydrodynamic radius from DLS measurements on the solution temperature: filled squares, 30 mg/mL; open circles, 10 mg/mL; filled triangles, 1.0 mg/mL; open triangles, 0.1 mg/mL; filled stars, 0.01 mg/mL. Inset shows the exponent of the power law fits.



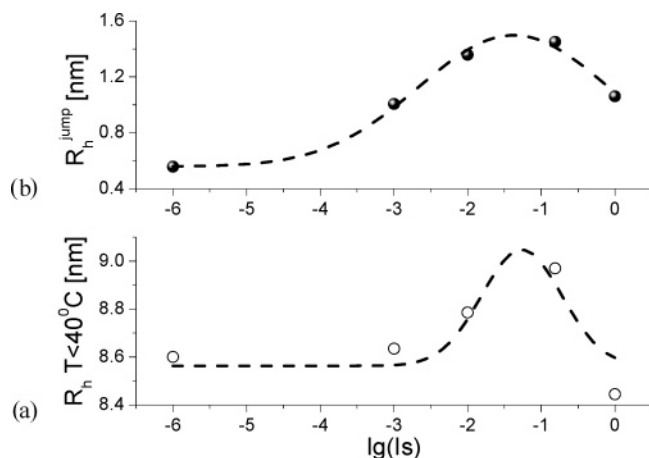
**Figure 8.** Dependence of the hydrodynamic radius on temperature for a series of sodium chloride concentrations from DLS measurements: filled squares, deionized water; open circles, 1 mM NaCl; open triangles, 10 mM NaCl; filled triangles, 154 mM NaCl; half-filled triangles, 1 M NaCl.

samples that had been hydrolyzed (2 M TFA, 2 h at 120 °C) and reduced (sodium borodeuteride ( $\text{NaBD}_4$ ) in 2 M  $\text{NH}_4\text{OH}$ , 2 h at room temperature). The sample was then acetylated using acetic anhydride (1 h at 100 °C). The acetylated sample was purified by extraction into chloroform. The partially methylated alditol acetates were then examined by GC-MS. A standard mixture of partially methylated alditol acetates was also run under the same conditions.

**2.7. Delayed Extraction Matrix-Assisted Laser Desorption Ionization Mass Spectrometry, Electrospray Mass Spectrometry, and Linkage Analysis of the O-Linked Carbohydrates.** Matrix-assisted laser desorption ionization (MALDI) time-of-flight (TOF) mass spectrometry was performed using a Voyager STR Biospectrometry Research Station Laser Desorption mass spectrometer coupled with delayed extraction technology. Dried permethylated glycans were redissolved in methanol/water (80:20) and analyzed using a matrix of 2,5-dihydroxybenzoic acid. Angiotensin and ACTH fragments in sinapinic acid were used as external calibrants.

Electrospray (ES) mass spectrometry was performed on a Micromass quadrupole TOF mass spectrometer. Dried permethylated glycans were redissolved in methanol/0.1% TFA (80:20 (v/v)) and loaded into a nanospray needle for analysis. Glu-fibrinopeptide was used to calibrate the instrument externally for both experiments (Supporting Information).





**Figure 9.** (a) Average hydrodynamic radius and (b) value of the temperature-induced jump in the hydrodynamic radius as a function of the ionic strength of the background electrolyte.

**2.8. Amino Acid Analysis.** We used an acid hydrolysis method with 6 N HCl followed by derivatization using phenylisothiocyanate and separation by reverse-phase high-performance liquid chromatography (RP-HPLC, Hewlett-Packard 1050 series), coupled with UV detection (254 nm), for amino acid analysis. The lyophilized aliquots of mucin sample were hydrolyzed directly using constant boiling HCl at either 110 °C for 24 h or 145 °C for 4 h. Quantification of the amino acid content was achieved by comparison of the data obtained with that from a standard mixture of amino acids.

### 3. Results and Discussions

**3.1. Analysis of Chemical Structure.** To establish the chemical structure of the porcine mucin, we performed monosaccharide analysis, amino acid analysis, and O-linked oligosaccharide screening. GC-MS and HPAEC-PAD for monosaccharide composition analysis of the “Orthana” mucin sample gave similar results. The data from the two techniques are summarized in Table 1. The percentage weight (wt %) of single monosaccharides and the number fraction (n %) of each of them has been calculated using the molecular weights of each of the monosaccharides, assuming the loss of water upon etherification/peptide bond formation. In each case, fucose, galactose, *N*-acetylglucosamine, and *N*-acetylgalactosamine were detected as the major monosaccharide species; they account for 97–100% of all of the monosaccharide residues (Table 1).

Trace levels of xylose and glucose are most likely derived from airborne dust contaminants associated with sample preparation, because they are not naturally occurring monosaccharides

in *O*-glycans. The presence of mannose is most likely due to a very low level of *N*-glycans present in the sample.

Sialic acid species were detected at trace levels by GC-MS but were not detectable by HPAEC-PAD. The overall weight fraction of sugars in the mucin sample is estimated as 71–76% (w/w). This value is in agreement with literature values for sugar/protein ratios for mammalian gastric/colonic mucins.

The amino acid analysis results for the “Orthana” mucin from the two hydrolysis conditions can be found in the Supporting Information. The percentage contribution to the mucin weight (wt %) of each amino acid and the number fraction (n %) were calculated assuming the loss of water upon peptide bond formation. The 110 °C hydrolysis conditions gave higher relative amounts for most of the amino acids, indicating that these conditions are optimal for the hydrolysis of the protein component of the mucin sample. The percentage correlation of variance (% CV) values for 145 °C were slightly worse than those for 110 °C, but the majority of the results was still well within the CV acceptance limit of 10%.

The overall protein content is found to be 20–21.5%. Assuming that the sugar fraction is 71–76%, one can assume that overall purity (percentage ratio of detectable sugars and amino acids relative to a dry aliquot of the material) of “Orthana” mucin is around 91–97%, depending on the analysis method.

The presence of histidine, arginine, and lysine groups that cumulatively account for 6–7% of all of the amino acid residues could be responsible for the existence of a substantial positively charged domain(s) within the mucin molecules. Together with delta negative sugars and polar amino acids, they contribute to the polyampholyte nature of “Orthana” mucin.

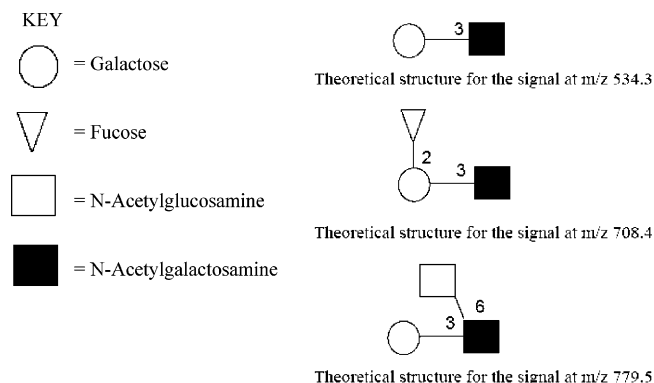
The relatively high fraction of hydrophobic amino acids—valine, isoleucine, leucine, and alanine (cumulatively 20–26%)—indicates the existence of hydrophobic domains. The relative distribution of alanine, proline, serine, threonine, valine, and histidine residues indicates that “Orthana” mucin is likely to belong to the MUC6 type of gastric mucins. The “absence” (below the detection limit) of asparagine residues indicates that *N*-glycosylation can only be found in trace amounts, in agreement with the low content of the mannose residues found in GC-MS experiments. The serine and threonine residues, the sites of *O*-glycosylation, are found to be abundant in quantities of 30–45%, suggesting that “Orthana” is heavily glycosylated.

MALDI-MS of the *O*-glycans released by reductive elimination showed the presence of three main signals that represent di- and trisaccharides (Supporting Information). Lesser amounts of tetra- to hexasaccharides were seen with only minor levels of hepta- to decasaccharides present. The postulated structures

**Table 1.** Comparison of the Amounts (nmol) of Each Monosaccharide Detected in “Orthana” Mucin by GC-MS and HPAEC-PAD Chromatography<sup>a</sup>

monosaccharide	$M_w$	$rM_w$	GC-MS			HPAEC-PAD		
			(nmol/60 $\mu$ g aliquot)	wt %	n %	(nmol/60 $\mu$ g aliquot)	wt %	n %
fucose	164.16	146.16	27	6.58	11.49	33	8.04	13.36
xylose	150.13	132.13	1.3	0.29	0.55	NA	NA	NA
mannose	180.16	162.16	1.8	0.49	0.77	ND	ND	ND
galactose	180.16	162.16	76	20.54	32.34	69	18.65	27.94
glucose	180.16	162.16	2.6	0.7	1.11	ND	ND	ND
GalNAc	222.21	204.21	57	19.4	24.26	60	20.42	24.29
GlcNAc	222.21	204.21	69	23.48	29.36	85	28.93	34.41
sialic acid	309.27	291.27	0.30	0.15	0.13	ND	ND	ND
total			235	71.63	100	247	76.04	100

<sup>a</sup> NA = not analyzed; ND = not detected.



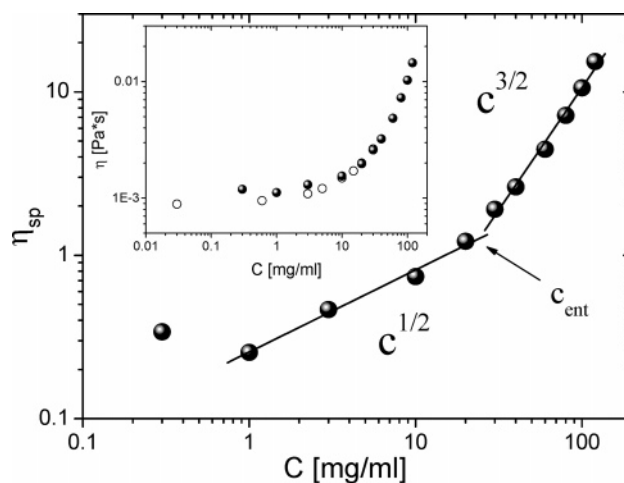
**Figure 10.** Postulated structures from analysis of the data obtained from MALDI-MS, ES-MS, and linkage analysis by GC-MS for the most abundant O-glycans. All monosaccharides are linked through carbon 1, and thus, for simplicity, this has not been labeled on the structures.

from the analysis of the data obtained from MALDI-MS, ES-MS, and linkage analysis by GC-MS for the most abundant O-glycans are presented in Figure 10. All monosaccharides are linked through carbon 1, and thus, for simplicity, this has not been labeled on the structures. The structures exhibit heterogeneity on the location and linkage of the fucose residue, forming either Lewis or Blood Group H type antigens.

The degree of glycosylation was estimated using the results from amino acid analysis and a molecular weight of 546 kDa ( $M_w$  from SLS). The mass fraction of the serine and threonine residues is  $\sim 45\%$  of the total amino acid composition of a pure mucin molecule, which implies that there are  $\sim 600$  glycosylation sites per mucin molecule. A similar estimation was done for oligosaccharide units using the molecular mass of the most abundant HexNAc–Hex–HexNAc trisaccharide,  $\sim 570$  Da. It was found that one mucin molecule contains approximately 700 oligosaccharide residues, which is higher than the number of O-glycosylation sites available. This discrepancy could be due to the presence of hexa-, hepta-, and decasaccharides in the carbohydrate population of the “Orthana” mucin or instrumental analysis errors. However, it can still be concluded that the absolute majority of the O-glycosylation sites of a protein backbone are glycosylated. This would characterize “Orthana” mucin as a highly glycosylated glycoprotein.

**3.2. Viscosity of Bulk Solutions of Hydrophobic Polyampholytes.** It was established from chemical analysis that “Orthana” mucin has predominantly neutral O-linked oligosaccharide side chains that provide essential hydrophilicity and allows mucin to be dissolved relatively easily in water at concentrations of up to 200 mg/mL with no evidence of phase separation. The presence of positively charged domains and hydrophobic amino acid residues suggests that mucin is a heterogeneous hydrophobic polyampholyte. Mucin possesses both positively and negatively charged domains, so they must exist in a state of balanced charge regulation that prevents the mucin molecule from inverting its charge; the exact mechanism of such a charge regulation is unclear, but it may be realized through multiple mesoscale dipole moments within the mucin domains similar to those suggested by Dobrynin et al. for the necklace structure of polyampholytes in bulk solutions.<sup>36</sup>

The rheology measurements reveal that the viscosities of “Orthana” mucin solutions are relatively low compared with other higher-molecular-weight mucins and/or polymers of a similar molecular mass, e.g., chitosan. At a mucin concentration of 30 mg/mL, the solution is only 3 times more viscous than



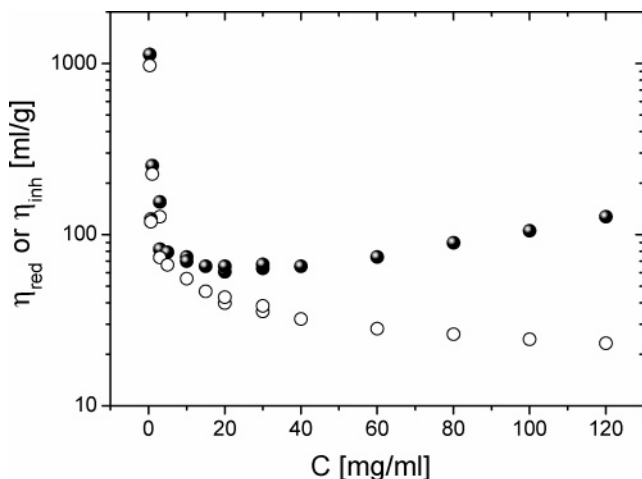
**Figure 11.** Specific viscosity of mucin solutions. Inset: The dependences of the viscosity on mucin concentration obtained by rotational rheometry (open circles) and microrheology (filled circles) techniques. The behaviors for semidilute unentangled ( $\eta \sim c^{1/2}$ ) and semidilute entangled ( $\eta \sim c^{3/2}$ ) polyelectrolyte solutions are shown. The entanglement concentration is labeled ( $C_{ent}$ ).

water. For shear rates of  $10\text{--}1000\text{ s}^{-1}$ , “Orthana” mucin solutions remain Newtonian across a broad range of concentrations (Supporting Information). We have also performed interfacial rheology and microgap high-shear rheological measurements that have shown that “Orthana” mucin does not build up any substantial interfacial elasticity. The mucins remain Newtonian up to shear rates of  $10^5\text{ s}^{-1}$  for gaps as low as 10  $\mu\text{m}$ .

The dependency of the viscosity on the mucin concentration over a wide range of mucin concentrations was determined via microrheological measurements<sup>35</sup> and is presented in the inset of Figure 11. The corresponding data for steady shear viscosity measured using the ARES rheometer are superimposed on the plot (red spheres). It is clear that the microrheological measurements are in reasonable agreement with rotational rheometry. To explore the molecular origin of the viscosity data, we plotted the specific viscosity ( $\eta_{sp} = \eta - \eta_{\text{solvent}}/\eta_{\text{solvent}}$ ) versus concentration, as given in Figure 11. One can clearly observe two well-defined power law scaling regimes. At high mucin concentrations, the specific viscosity scales as  $\eta_{sp} \sim c^{3/2}$ , and below  $c = 24\text{ mg/mL}$  as  $\eta_{sp} \sim c^{1/2}$ . This behavior agrees with theoretical predictions<sup>37,38</sup> made for both unentangled ( $\nu = 1/2$ , i.e., Fuoss law) and entangled ( $\nu = 3/2$ ) flexible linear polyelectrolyte solutions in low-salt conditions.

In the low mucin concentration (Fuoss) regime, the reduced viscosity ( $\eta_{red} = \eta_{sp}/c$ ) scales as  $\eta_{red} \sim c^{-1/2}$ , and therefore the intrinsic viscosity  $[\eta]$  cannot be extracted. This unusual behavior of polyelectrolyte solutions can be rationalized at low concentrations by using the technique of isoionic dilutions.<sup>39,40</sup> In this method, the viscosity is studied as a function of polymer concentration for a given value of the ionic strength (i.e., as the solution is diluted, salt is added to keep the ionic strength constant). The increase in the reduced viscosity at low  $c$  is then linear, as for neutral polymers. In neutral polymer solutions, this linear behavior is usually represented in the form  $\eta_{red} = [\eta] + k_H[\eta]^2c$ , where  $k_H$  is the Huggins constant. Generally, in polyelectrolyte solutions,  $[\eta]$  is large and varies strongly with the ionic strength ( $I_s$ ).

The present measurements were done in both water and PBS (0.154 M salt concentration). The PBS buffer is characterized by extremely low values of the Debye screening length.



**Figure 12.** Reduced (filled circles) and inherent (open circles) viscosities as a function of mucin concentration.

However, the Fuoss regime still occurred following the addition of salt in the PBS buffer with no significant change in the viscosity observed. Such behavior is probably associated with the weak overall charge of the mucin molecules, and thus the enthalpic component of the charge–charge interactions appears to be less important than entropic contributions related to the chain flexibility.

To examine the low concentration mucin data in more detail, the reduced and inherent ( $\eta_{\text{inh}} = 1/c \ln(\eta_{\text{sp}} + 1)$ ) viscosities versus mucin concentration were measured using both microrheology and bulk rheology methods as shown in Figure 12. The increase of the viscosity at low concentrations is the Fuoss behavior. To calculate the overlap concentration ( $c^*$ ) that is inversely related to the intrinsic viscosity,  $c^* = 1/[\eta]$ , we have used a power law expansion for viscosity according to the Dobrynin–Colby–Rubenstein model<sup>36,37</sup>

$$\eta = \eta_{\text{solvent}} F(c/c^*)$$

$$F(c/c^*) \approx (c/c^*)^{1/2} \begin{cases} (c/c^*)^{-1/10}, & c < c^* \\ 1, & c^* < c < c_{\text{ent}} \\ \tilde{a}(c/c^*), & c > c_{\text{ent}} \end{cases} \quad (4)$$

In the expression above, the parameter  $1/\tilde{a} \approx n^4$  describes an entanglement onset, where each chain has to overlap with  $n$  others. The reduced viscosity is therefore given by the definition

$$\eta_{\text{red}} = (\eta - \eta_{\text{solvent}})/c\eta_{\text{solvent}} = \frac{1}{c}(F(c/c^*) - 1) \quad (5)$$

The fitting procedure enabled the effective overlap concentration to be determined as  $c^* = 1/[\eta] = 19 \text{ mg/mL}$ , and thus the intrinsic viscosity is found to be  $[\eta] = (c^*)^{-1} = 53.6 \text{ mL/g}$ .

The Dobrynin–Colby–Rubenstein model<sup>37</sup> gives an empirical justification for the observation that semidilute polyelectrolyte solutions obey Zimm–Rouse dynamics over a broad range of concentrations above the overlap concentration. In neutral polymers, it is observed that entanglements become important when each chain interacts with about 10 other chains.<sup>41</sup> However, for comb polymers much fewer overlapping chains are typically required for entangled rheology.<sup>42</sup> With such short side chains (number of side monomers  $N_s$  in a side chain is 2–4), the viscosity of the entangled mucin solution follows a  $\eta \sim c^{3/2}$  scaling in contrast to the  $\eta \sim c^2$  scaling predicted and observed with long flexible side-chain combs.<sup>38</sup>

**Table 2.** Overlap Concentrations That Are Relevant for the Structure and Dynamics of Mucin Molecules<sup>a</sup>

(mg/mL)		
$c^*$		1.54
$c_{\text{brush}}^{\text{I}}$	$R_{\text{comb}} = 2 \text{ nm}$	45.1
	$R_{\text{comb}} = 2.5 \text{ nm}$	23.1
	$R_{\text{comb}} = 2 \text{ nm}$	30.1
	$L_{\text{comb}} = 4 \text{ nm}$	
$c_{\text{brush}}^{\text{II}}$	$R_{\text{comb}} = 2.5 \text{ nm}$	15.4
	$L_{\text{comb}} = 5 \text{ nm}$	
$c_{\text{daisy}}^{\text{I}}$	$N_{\text{daisy}} = 2$	274.5
	$N_{\text{daisy}} = 3$	183.0

<sup>a</sup>  $c^*$  is the semidilute overlap concentration,  $c_{\text{brush}}$  is the brush overlap concentration, and  $c_{\text{daisy}}$  is the overlap concentration of the globular structures.

The concentration for the overlap of the whole mucin chains ( $c^*$ ) can be calculated in the standard manner

$$c^* = \frac{M_{\text{mucin}}}{4/3\pi R_g^3} = \frac{546\,000(\text{g/mol}) \times 10^{-3}(\text{m}^3/\text{L})}{4/3\pi \times (52 \times 10^{-9})^3(\text{m}^3) \times 6.02 \times 10^{23}(\text{mol}^{-1})} = 1.54 \text{ mg/mL} \quad (6)$$

At higher concentrations, a number of other overlap concentrations are expected to be important. The overlap concentrations of the daisy beads ( $c_{\text{daisy}}$ ) and the side chains of the comb subunits ( $c_{\text{brush}}$ ) depend on the mucin structure. For  $c_{\text{brush}}$ , one can suggest spheroidal ( $c_{\text{brush}}^{\text{I}}$ ) and cylindrical ( $c_{\text{brush}}^{\text{II}}$ ) approximations for the shape of the oligosaccharide side chain, giving slightly different expressions for the brush overlap concentration (their applicability depends on the degree of backbone flexibility)

$$c_{\text{brush}}^{\text{I}} = \frac{M_w(\text{whole mucin})}{N_{\text{combs}} \times 4/3\pi R_{\text{comb}}^3} \quad (7)$$

$$c_{\text{brush}}^{\text{II}} = \frac{M_w(\text{whole mucin})}{N_{\text{combs}} L_{\text{comb}} \pi R_{\text{comb}}^3} \quad (8)$$

where  $N_{\text{comb}}$  is the number of comb subunits on a mucin chain,  $R_{\text{comb}}$  is the comb radius (size of the side chains), and  $L_{\text{comb}}$  is the comb length.

The daisy-chain overlap concentration can be defined as

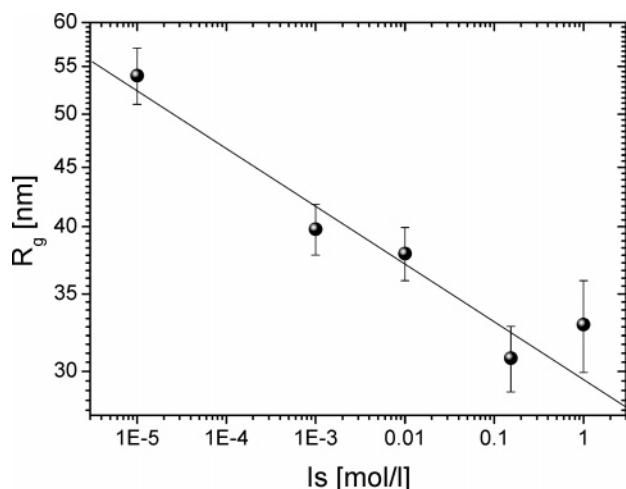
$$c_{\text{daisy}} = \frac{M_w(\text{whole mucin})}{N_{\text{daisy}} V_{\text{daisy}}} \quad (9)$$

where  $N_{\text{daisy}}$  is the number of hydrophobic globules on a chain and  $V_{\text{daisy}}$  is the volume of a globule. The calculated values are summarized in Table 2.

$c^*$  corresponds to the transition between the dilute and the semidilute rheology in Figure 12, and  $c_{\text{brush}}$  coincides to the entanglement concentration ( $c_{\text{ent}}$ ).<sup>13</sup> The globular daisy structures would be expected to overlap at much higher concentrations ( $c_{\text{daisy}}$ ) than those probed in the rheology experiments ( $\sim 200 \text{ mg/mL}$ ).

As expected for nonassociating mucin solutions,<sup>8,43</sup> no evidence was found of nonlinear viscoelastic behavior.<sup>44,45</sup> Also the gelation of the mucin solutions at higher concentrations or low pH values was not observed<sup>12</sup> due to the virtual absence of cysteine residues.





**Figure 13.** Radius of gyration of the mucin molecules as a function of ionic strength from SLS experiments. The straight line is a power law fit,  $R_g \sim I_s^{0.05}$ .

**3.3. Structural Analysis of Bulk Solutions.** Static light scattering data demonstrated little dependence of the form factor ( $S(q, c)$ ) on the bulk mucin concentration. From the analysis of Zimm plots, the radii of gyration are 51.8, 57.9, 56.0, and 52.0 nm at concentrations of 0.3, 1, 3, and 30 mg/mL, respectively. The calculated molecular weight ( $M_w$ ) is 546 400 g/mol and the second virial coefficient ( $a_2$ ) is  $2.34 \times 10^{-4} \text{ cm}^3/\text{mol}$ .

The molecular weight ( $M_w$ ) and radius of gyration ( $R_g$ ) data for the “Orthana” mucin are in good agreement with previous data.<sup>13</sup> On the basis of the  $M_w$  and  $R_g$  data, it is possible to calculate the intrinsic viscosity for the mucin solution, which is defined as  $[\eta] = \phi R_g^3/M_w = 53.5 \text{ mL/g}$ , where  $\phi$  is a universal constant equal to  $2.1 \times 10^{23} \text{ mol}^{-1}$ . This value of the intrinsic viscosity is very close to the result from bulk rheology measurements ( $[\eta]_{\text{rheol}} = 53.6 \text{ mL/g}$ ).

The Rubinstein–Dobrynin (DR) scaling model of polyelectrolyte solutions has been successful in interpreting the dependence of the diffusion coefficient, chain size, relaxation time, and viscosity as a function of the polyelectrolyte concentration in dilute, semidilute, and entangled solutions. For linear flexible dilute polyelectrolytes Rubinstein and Dobrynin predict that the radius of gyration ( $R$ ) is given by

$$R \approx bN^{3/5}(cb^3)^{-1/5}(1 + 2Ac_s/c)^{-1/5} \quad (10)$$

where  $c$  is the monomer concentration,  $b$  is the monomer size,  $N$  is the number of monomers in a chain,  $A$  is the charge fraction, and  $c_s$  is the salt concentration. The result was tested in a recent article on linear flexible polyelectrolytes (maleic anhydride copolymers<sup>46</sup>), and the salt dependence was in reasonable agreement with experiment.

A change in conformation due to the presence of background electrolyte was previously reported<sup>7</sup> for the long-chain gastric mucin of the MUC5B family ( $M_w > 2 \text{ MDa}$ ), but no exact scaling parameter was determined. In Figure 3, the angular dependencies of static form factors,  $S(q, c)$ , for “Orthana” mucin solutions with different salt levels are shown. The evaluated values for  $R_g$  versus ionic strength are presented in Figure 13. The  $R_g$  value decreases with an increase in salt concentration, as is expected from the theoretical model (eq 10) that predicts the contraction of polymer chains with a decrease in the electrostatic screening length. The experimental value of the scaling factor is found to be  $1/20$ . The value of this exponent is nearly an order of magnitude less than that predicted by the

DR polyelectrolyte theory, and it could be explained by the low density of charges on the mucin chains (small number of sialic acid groups) as well as a substantial fraction of hydrophobic domains. A low value of the scaling parameter ( $\nu = 1/10 - 0$ ) has been found for polyampholytes,<sup>47</sup> hydrophobic polyelectrolytes,<sup>48</sup> and stiff polymers.<sup>49</sup> The relative increase in  $R_g$  for 1 M NaCl solutions can be attributed to the reswelling of the mucin coil due to the entropically driven penetration of hydrated counterions into the coil volume. This effect is characteristic of polyampholytes;<sup>36</sup> the value of critical electrolyte concentration ( $c_s^*$ ) observed by McCormick and Salazar<sup>50</sup> is 0.1 M NaCl for mucin and is in agreement with our experimental results (Figure 13).

The DLS measurements of mucin solutions of different concentrations were performed at six different temperatures: 12.2, 20, 25, 37, 47, and 60 °C. In accordance with the rheological data, we have examined the transitions between dilute, semidilute, and entangled regimes of mucin solutions with DLS. Experiments were completed with four different mucin concentrations: 0.1, 1.0, 10.0, and 30.0 mg/mL. For all concentrations, we observed a linear  $q^2$  dependence for the relaxation rate ( $\Gamma$ ) of the fast mode (Figure 5). The slow mode has a predominantly  $q^2$  dependence, indicating a predominantly diffusive motion. The slightly higher than  $q^2$  dependence of the slow mode with the lowest concentration mucin could be attributed to slight aggregation of the mucin molecules.<sup>46</sup> The relative amplitudes of the two modes are relatively independent of  $q$  (Figure 3) probably due to the close proximity of the sizes of the two modes ( $\sim 10$  and  $\sim 50$  nm).

The dependencies of the hydrodynamic radius ( $R_h$ ) on the temperature and the mucin concentration are summarized in Figures 7 and 8. We conclude that the observed hydrodynamic radius of  $\sim 8$ – $16$  nm of the fast mode does not correspond to the radius of gyration ( $R_g$ ) determined from the SLS and intrinsic viscosity data ( $\sim 55$  nm). Indeed, the slower of the two DLS modes is in much better agreement with the values of the radius of gyration at low mucin concentrations. For the lowest dilute mucin concentration (0.1 mg/mL), we observe the highest value of the hydrodynamic radius for the fast mode:  $R_h \approx 15$  nm. The classical Zimm model for flexible linear chains predicts that  $R_g = 1.5R_h$ ,<sup>51</sup> and for polyelectrolytes we expect  $R_g$  to be about  $1.3R_h$ .<sup>13</sup> Our data do not agree with these predictions,<sup>51</sup> and this cannot be due to the highly anisotropic shape of the mucin molecule as previously reported for ovine submucillary mucins,<sup>25</sup> because it would not account for the 500% discrepancy between SLS and DLS measurements. Instead, we consider that the “fast” DLS mode corresponds to internal dynamics of the mucin molecules, corresponding to subunits of size  $R_h$ , an internal daisy-chain conformation that is governed by both the charge and the hydrophobicity of the mucin molecules. The possibility that two different molecular weight mucin species (namely, big and small) are present in the solution or that a high polydispersity exists in the mucin sample can be excluded, because bigger species give a significant contribution to the correlation function and would be discriminated by the CONTIN analysis (Figure 3).

Figure 6 demonstrates that increasing the mucin concentration by an order of magnitude from 0.1 to 1.0 mg/mL does not substantially increase the values of the fast  $R_h$  observed. For the semidilute regime at 1.0 mg/mL, we might expect to attribute the fast  $R_h$  to the cooperative diffusion mesh size ( $\xi_h$ ). For a linear polyelectrolyte chain, the mesh size should scale with polymer concentration as  $\xi_h = c^{-1/2}$ . From Figure 6, we obtain an experimental value of the scaling parameter for “Orthana”



muin of  $\xi_h = c^{-0.085}$ . This low value of the scaling parameter for the mucin solutions suggests that there is no dependency of the mesh size on the mucin concentration and again indicates the fast mode is an internal mode for subunits inside the mucin molecules.

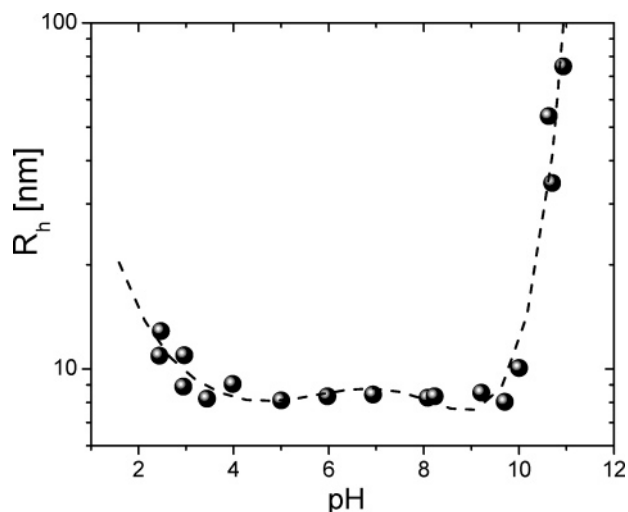
The measured fast  $R_h$  decreases with temperature (Figure 7), and the values of  $dR_h/dT$  (Figure 7, inset) are found to decrease with increasing mucin concentration. If we again attempt to associate the fast  $R_h$  with  $\xi_h$ , then the observed behavior indicates that (i) water is a marginal solvent for mucin,<sup>52,53</sup> because in the good solvent  $\xi_h$  should not change with temperature,<sup>36,52</sup> and (ii) the quality of the solvent increases (e.g.,  $R_h$  decreases) with increasing mucin concentration. Clearly, these conclusions contradict each other.

From the above observations, we conclude that the observed fast radius of gyration ( $R_h$ ) does not correspond to the cooperative diffusion of the semidilute mesh of the mucin molecules. Analysis of the literature regarding DLS studies of mucins suggests that a number of previous reports<sup>12,13,19,24</sup> have misinterpreted the dynamic mode in favor of either correlator after pulse error,<sup>12</sup> cooperative diffusion of the mesh ( $\xi_h$ ),<sup>19,24</sup> or dynamics of a sticky entangled mesh ( $\xi_h$  for the sticky molecules).<sup>13</sup> However, the entangled dynamics in sticky mucin samples is characterized by a nonlinear  $q^2$  dependence of  $\Gamma(q)$  that was not observed in our experiments for the fast mode. The present study is unique as we have explored the dynamics of mucin solutions across a broad range of temperatures and concentrations, thereby showing that the fast mode cannot be interpreted as cooperative whole molecule diffusion, which would be extremely sensitive to changes in polymer concentration.

To examine the hypothesis associated with the anisotropy of the mucin molecule (i.e., rod shape/stretched coils), we have performed depolarized DLS measurements. These confirmed that the "Orthana" mucin molecule has an isotropic shape with an aspect ratio ( $\rho$ ) of 0.75–1.5.

The scattering results are consistent with a daisy bead model, introduced in the works of Rubinstein and co-workers, that describes the configuration of polyampholytes in solutions of different solvent quality<sup>36</sup> (Figure 1). Densely packed domains (beads) are connected via fast moving flexible chain regions that allow the beads to move independently within the time scales of their diffusive dynamics. The internal friction coefficients are thus dominated by the globular daisy subunits, and the Stokes–Einstein equations for their Zimm dynamics can be applied to a first approximation, providing a characteristic hydrodynamic size of the globules (the measured value of the fast  $R_h$ ). We thus have two giant beads connected by a flexible spacer that constitute a rescaled two-bead Zimm model. The relaxation time of a single bead corresponds to the rescaled monomer relaxation time ( $\tau_0 \approx \eta_s b^3/kT = \tau_{\text{fast}}$ , where  $b$  is the radius of gyration of a globule<sup>34</sup>), and the center of mass diffusion ( $\tau_{\text{slow}}$ ) is given by the Zimm time for the two-bead chain ( $\tau_Z \approx \tau_0 N^{3\nu}$ ). We assume that the statistics act as for a polymer in a good solvent ( $\nu = 0.588$ ), and we obtain a reasonable estimate of  $\tau_{\text{slow}}/\tau_{\text{fast}} = 3.4$ . More accurate calculations are possible by explicitly including the dynamics of the flexible spacer, a triblock copolymer type model for the Zimm modes.<sup>54</sup>

The presence of two types of domains (daisy beads with connecting chains) within a single mucin molecule requires further investigation to characterize whether these domains consist of glycosylated blocks or the collapsed naked protein backbone fragments. The effect of the screening length on the



**Figure 14.** Dependence of the hydrodynamic radius of the mucin molecules from the fast mode on the pH of the solutions from DLS experiments.

dynamics of daisy beads provides more information on the chemical origin of the daisy bead, the glycosylated unit versus the folded protein globule. The possibility of a tightly packaged disulfide bridged protein domain in the globules can be excluded due to the absence of cysteine residues in the molecule. In Figures 8 and 9, the experimental data for the dependence of the fast  $R_h$  on the salt concentration are presented. It can be clearly seen that the added salt has a negligible effect on the conformation of the daisy beads and their diffusive mobility. Such a small effect excludes the possibility of an electrostatically governed folded globular structure for the daisy bead but rather suggests that daisy beads are formed around hydrophobic domains separated by heavily glycosylated fragments.

**3.4. Effects of pH and  $\text{Ca}^{2+}$ .** Previous studies have shown that pH has a strong effect on mucin configurations.<sup>7</sup> Although nonsticky "Orthana" mucin does not exhibit a macrogel phase, in general the aggregation of mucins is expected at both low and high values of pH.<sup>12</sup> This is due to cross-linking via hydrophobic domains at low charge fractions. All measurements were performed with a single-ion buffer, and therefore all ion-specific effects can be excluded.

The pH dependence of fast  $R_h$  is presented in Figure 14. For pH's in the range of  $3 < \text{pH} < 10$ , no significant changes in the fast  $R_h$  were observed. At  $\text{pH} > 10$ , one can observe an abrupt increase in the hydrodynamic radius, most probably due to the partial hydrolysis of the mucin. Because the observed hydrodynamic radius exceeds the  $R_g$  value of a single mucin molecule, the agglomeration or aggregation of 2–3 mucin molecules must take place. At the low values of pH ( $< 3$ ), we observe a slight increase in the fast  $R_h$  of mucin that can be attributed to the swelling of daisy beads neutralized in the acidic environment.

The formation of a gel at low values of pH is usually attributed to the interaction between the hydrophobic domains of mucin molecules<sup>55</sup> when charges are suppressed in an acidic environment. However, "Orthana" mucin does not form a macrogel at low values of pH, most probably due to the daisy-chain structure that protects the hydrophobic domains.

The addition of  $\text{Ca}^{2+}$  ions did not induce any bulk-phase separation in "Orthana" mucin solutions, and no significant change in  $R_h$  ( $\delta R_h < 1$  nm) was observed. This is due to the absence of silicic acid groups that can form complexes with multivalent ions.<sup>56</sup>

#### 4. Conclusions

The analysis of the molecular structure of purified (>91%) porcine gastric "Orthana" mucin established that the overall weight fraction of sugars was in the range of 71–76%. The monosaccharide composition is predominantly fucose, galactose, *N*-acetylglucosamine, and *N*-acetylgalactosamine residues; they account for 97–100% of all monosaccharide residues. Sialic acid species were detected at trace levels.

MALDI-MS analysis of the *O*-glycans released by reductive elimination showed the presence of three main signals representing both di- and trisaccharides. Lesser amounts of tetra- to hexasaccharides were seen with only minor levels of hepta- to decasaccharides present. The basic amino acids—histidine, arginine, and lysine—cumulatively account for 6–7% of all of the amino acid residues. The hydrophobic amino acids—valine, isoleucine, leucine, and alanine—account for 20–26% of the residues. No evidence of cysteine residues was found. The serine and threonine residues, the sites of *O*-glycosylation, are found to be in abundant quantities (30–45%). Together with the oligosaccharide analysis, it is suggested that the majority of the *O*-glycosylation sites of a protein backbone are glycosylated. "Orthana" mucin is a highly glycosylated glycoprotein.

The experimental investigation of "Orthana" mucin molecular conformation in aqueous solutions reveals a daisy-chain configuration of the comb subunits. The daisy beads are roughly spherical and are separated by flexible chains, presumably consisting of glycosylated fragments. The daisy beads are most likely due to intramolecular assemblies of numerous hydrophobic domains that are surrounded by hydrophilic neighboring parts of the mucin molecule. The daisy-chain conformation stems from the polyampholyte nature of the mucin, which has a relatively high proportion of basic amino acids and negatively charged sugars, as well as the presence of hydrophobic domains.

The viscosity of aqueous solutions of the mucin samples shows Newtonian behavior for shear rates up to  $10^5 \text{ s}^{-1}$ . The concentration dependence of the specific viscosity ( $\eta$ ) of the mucin solutions indicates the existence of semidilute ( $\eta \sim c^{1/2}$ ) and entangled ( $\eta \sim c^{3/2}$ ) regimes in agreement with previous results for short-side-chain synthetic comb polyelectrolytes.

**Acknowledgment.** Ann-Marie Williamson and Tom McLeish are kindly acknowledged for helpful discussions and support, John Windus is acknowledged for bioinformatics support in genetic allocation of "Orthana" mucin, Georgina Davies and Jason Stokes are kindly acknowledged for an immense help with setting up the bulk rheological measurements and further discussions, Bill Frith and Jeroen Bongaerts are kindly acknowledged for helpful discussion of DLS results, Damiano Rossetti is kindly acknowledged for interfacial characterization of mucin, and Asish Nandi is kindly acknowledged for the help with sample preparation.

**Supporting Information Available.** Comparison of the hydrolysis conditions for the amino acid analysis, dependence of the viscosity on the shear rate, and MALDI mass spectrum of the permethylated *O*-glycans released from "Orthana" mucin. This material is available free of charge via the Internet at <http://pubs.acs.org>.

#### References and Notes

- Harding, S. E. *Adv. Carbohydr. Chem. Biochem.* **1989**, *47*, 345–381.
- Fogg, F. J. J.; Hutton, D. A.; Jumel, K.; Pearson, J. P.; Harding, S. E.; Allen, A. *Biochem. J.* **1996**, *316*, 937–942.
- van Wiechen, P. H.; Booi, H. C. *J. Eng. Math.* **1971**, *5*, 89–98.
- Kantor, Y.; Kardar, M.; Ertas, D. *Physica A* **1998**, *249*, 301–306.
- Yakubov, G. E.; Papagiannopoulos, A.; Rat, E.; Waigh, T. A. *Biomacromolecules*, submitted for publication.
- Papagiannopoulos, A.; Fernyhough, A. M.; Waigh, T. A. *J. Chem. Phys.* **2005**, *123*, 214904.
- Bansil, R.; Stanley, E.; LaMont, J. T. *Annu. Rev. Physiol.* **1995**, *57*, 635–657.
- Lee, S.; Mueller, M.; Rezwan, K.; Spencer, N. D. *Langmuir* **2005**, *21*, 8344–8353.
- Berg, C. H.; Lindh, L.; Arnebrant, T. *Biofouling* **2004**, *20*, 65–70.
- Tabak, L. A. *Annu. Rev. Physiol.* **1995**, *57*, 547–564.
- Carlstedt, I.; Sheehan, J. K.; Corfield, A. P.; Gallagher, J. T. *Essays Biochem.* **1985**, *20*, 40–76.
- Cao, X.; Bansil, R.; Bhaskar, K. R.; Turner, B. S.; LaMont, T.; Niu, N.; Afdhal, N. H. *Biophys. J.* **1999**, *76*, 1250–1258.
- Waigh, T. A.; Papagiannopoulos, A.; Voice, A.; Bansil, R.; Unwin, A. P.; Dewhurst, C. D.; Turner, B.; Afdhal, N. *Langmuir* **2002**, *18*, 7188–7195.
- Bhaskar, K. R.; Gong, D.; Bansil, R.; Pajevic, S.; Hamilton, J. A.; Turner, B. S.; LaMont, T. *Am. J. Physiol.* **1991**, *261*, G827–G832.
- Durrer, C.; Irache, J. M.; Duchene, D.; Ponchel, G. *J. Colloid Interface Sci.* **1995**, *170*, 555–561.
- Veerman, E. C. I.; van den Keijbus, P. A. M.; Nazmi, K.; Vos, W.; van der Wal, J. E.; Bloemena, E.; Bolscher, J. G. M.; Amerongen, A. V. N. *Glycobiology* **2003**, *13*, 363–366.
- Wu, A. M.; Csako, G.; Herp, A. *Mol. Cell. Biochem.* **1994**, *137*, 39–55.
- Shogren, R. L.; Jamieson, A. M.; Blackwell, J.; Jentoft, N. *Biopolymers* **1986**, *25*, 1505–1517.
- Gupta, R.; Jentoft, N.; Jamieson, A. M.; Blackwell, J. *Biopolymers* **1990**, *29*, 347–355.
- Strous, G. J.; Dekker, J. *Crit. Rev. Biochem. Mol. Biol.* **1992**, *27*, 57–92.
- Sheehan, J. K.; Oates, K.; Carlstedt, I. *Biochem. J.* **1986**, *239*, 147–153.
- Bhushana-Rao, K. S. P.; Masson, P. L. In *Mucus in Health and Disease I*; Elstein, M.; Parke, D. V., Eds.; Advances in Experimental Medicine and Biology 89; Plenum Press: New York, 1977; pp 275–282.
- Meyer, D. Y.; Silberberg, A. In *Respiratory Tract Mucus*; Ciba Foundation Symposium 54; Elsevier: Amsterdam, 1978; pp 203–218.
- Jumel, K.; Fogg, F. J. J.; Hutton, D. A.; Pearson, J. P.; Allen, A.; Harding, S. E. *Eur. Biophys. J.* **1997**, *25*, 477–480.
- Shogren, R.; Gerken, T. A.; Jentoft, N. *Biochemistry* **1989**, *28*, 5525–5536.
- Allen, A. In *Physiology of the Gastrointestinal Tract*; Johnson, L. R., Ed.; Raven Press: New York, 1981; Vol. 1, pp 617–639.
- Herp, A.; Wu, A. M.; Moschera, J. *Mol. Cell. Biochem.* **1979**, *23*, 27–44.
- Silberberg, A. *Biorheology* **1987**, *24*, 605–614.
- Sellers, L. A.; Allen, A.; Morris, E. R.; Ross-Murphy, S. B. *Biochem. J.* **1988**, *256*, 599–607.
- Nylander, T.; Arnebrant, T.; Baier, R. E.; Glantz, P. O. *Prog. Colloid Polym. Sci.* **1998**, *108*, 34–39.
- Gassin, G.; Heinrich, E.; Spikes, H. A. *Tribol. Lett.* **2001**, *11*, 95–102.
- Davies, G. A.; Stokes, J. R. *J. Rheol.* **2005**, *49*, 919–922.
- Stokes, J. R.; Graham, L. J. W.; Lawson, N. J.; Boger, D. V. *J. Fluid Mech.* **2001**, *429*, 117–153.
- Rubinstein, M.; Colby, R. H. *Polymer Physics*; Oxford University Press: New York, 2003.
- Waigh, T. A. *Rep. Prog. Phys.* **2005**, *68*, 685–742.
- Dobrynin, A. V.; Colby, R. H.; Rubinstein, M. *J. Polym. Sci., Part B: Polym. Phys.* **2004**, *42*, 3513–3538.
- Dobrynin, A. V.; Rubinstein, M. *J. Polym. Sci.* **2005**, *30*, 1049–1118.
- Doi, M.; Edwards, S. F. *The Theory of Polymer Dynamics*; Calendon Press: Oxford, U. K., 1986.
- Pals, D. T. F.; Hermans, J. J. *J. Polym. Sci.* **1948**, *3*, 897.
- Terayama, H.; Wall, F. T. *J. Polym. Sci.* **1955**, *16*, 357.
- Muthukumar, M. *J. Chem. Phys.* **1997**, *107*, 2619–2635.
- Papagiannopoulos, A.; Waigh, T. A.; Hardingham, T.; Heinrich, M. *Biomacromolecules* **2006**, *7*, 2162–2172.
- Raynal, B. D. E. *Biochem. J.* **2002**, *362*, 289–296.
- Davis, S. S. *Rheol. Acta* **1971**, *10*, 28–35.

- (45) van der Reijden, W. A.; Veerman, E. C. I.; Amerongen, A. V. N. *Biorheology* **1993**, *30*, 141–152.
- (46) Di Cola, E.; Waigh, T. A.; Colby, R. H. *J. Polym. Sci., Part B: Polym. Phys.* **2007**, *45*, 774–785.
- (47) Vaynberg, K. A.; Wagner, N. J.; Sharma, R. *Biomacromolecules* **2000**, *1*, 466–472.
- (48) Colby, R. H.; Boris, D. C.; Krause, W. E.; Tan, J. S. *J. Polym. Sci., Part B: Polym. Phys.* **1997**, *35*, 2951–2960.
- (49) Podgornik, R.; Parsegian, V. A. *Biophys. J.* **1998**, *74*, A177–A177.
- (50) McCormick, C. L.; Salazar, L. C. *Polymer* **1992**, *33*, 4384–4387.
- (51) Netz, R. R.; Joanny, J. F. *Macromolecules* **1998**, *31*, 5123–5141.
- (52) Brochard, F.; de Gennes, P. G. *Macromolecules* **1977**, *10*, 1157–1161.
- (53) Doi, M.; Onuki, A. *J. Phys. II* **1992**, *2*, 1631–1656.
- (54) Wang, R. N. *Macromolecules* **1975**, *8*, 364–371.
- (55) Shi, L.; Caldwell, K. D. *J. Colloid Interface Sci.* **2000**, *224*, 372–381.
- (56) Raynal, B. D. E.; Hardingham, T. E.; Sheehan, J. K.; Thornton, D. J. *J. Biol. Chem.* **2003**, *278*, 28703–28710.

BM700607W