Noam Greenspoon*,[†] and Ellen Wachtel[‡]

Contribution from the Departments of Organic Chemistry and of Polymer Research, The Weizmann Institute of Science, Rehovot, Israel 76100. Received September 4, 1990. Revised Manuscript Received May 20, 1991

Abstract: The surfactant sodium 3-[[1,2-bis[[(2-ethylhexyl)oxy]carbonyl]ethyl]thio]succinate (1) aggregates in chloroform/cyclohexane (1:1, v/v) mixture to form inverse micelles. The small-angle X-ray scattering from these solutions indicates that the aggregates are approximately spherical with a radius of gyration that varies from 6.3 to 18.5 Å, depending upon the amount of added water. These inverse micelles can bind and solubilize monosaccharides in nonpolar organic solvents. ¹H and ¹³C NMR, as well as ESR experiments on these solutions, indicate that monosaccharide derivatives are bound to the surfactant head group at low water content and are solubilized in the water pool at high water content.

Introduction

A large number of cellular recognition events involve specific interactions between particular classes of the carbohydrate structures on one cell surface and the receptor proteins on another cell surface.¹ The importance of carbohydrate recognition is evident from the fact that ca. 70% of antibodies obtained upon immunization with whole cells are directed against oligosaccharide structures.² Complex carbohydrate recognition by protein receptors plays an essential role in immunology, cell adhesion, and fertilization.³ Additionally, the transport of carbohydrates through lipophilic membranes into the cell, which in several cases is ATP-independent and selective for only one type of carbohydrate,⁴ involves a carbohydrate transporter protein, which is embedded in the cell membrane. This protein binds the saccharide from the blood and transports it into the cytoplasm.

Furthermore, specific interactions of some carbohydrates with phospholipids⁵ stabilize the permeability barrier of membranes and play an important role in the ability of certain organisms to survive states of complete dehydration.

Despite their proven participation in all these processes, the exact role of the oligosaccharide is still poorly understood. Similarly, not much is known about the basis for specificity and affinity in the binding. An understanding of these interactions at the molecular level would be very desirable, as would the design and synthesis of a molecule that will bind and solubilize sugars in organic solvents as a simple model for their binding and interactions in biological systems.⁶ An X-ray study of the arabinose-binding protein suggests that an important driving force for the complexation of the carbohydrate is the formation of a hydrogen bond network involving the arabinose, its first hydration shell, and the polar groups of the amino acid residues Asp, Asn, Arg, Glu, and Gln in the binding site.⁷ We have found that the novel synthetic surfactant⁸ sodium 3-[[1,2-bis[[(2-ethylhexyl)oxy]carbonyl]ethyl]thio]succinate (1)⁸ can solubilize monosaccharide derivatives in chloroform or chloroform/cyclohexane (1:1, v/v) mixture even at a very low water content.

Results and Discussion

The environment within the water pool formed in inverse micelles serves to mimic the polar pockets in enzymes.⁹ We are interested in studying this kind of environment as a simple model for the binding site of carbohydrate-binding proteins. Bis(2ethylhexyl) sulfosuccinate (Aerosol-OT; AOT) is a widely used surfactant for the study of inverse micelles. The environment in the aqueous part of this kind of aggregate has been used to solubilize inorganic or organic solutes. Solubilization of macromolecules by inverse micelles was also studied and used as a means for partial immobilization of enzymes.¹⁰ In all these studies relatively large amounts of water were present in the water pool beside the solute. We were interested to see if we could solubilize sugar in the water pool at a very low water content and detect

Chart I



Table I. Triaxial Radius of Gyration R_G and the Largest Vector That Can Be Contained within the Micelles, D_{max} , for a 50 mM Solution of 1 at Various Water and Glucose Contents

[water]/[1]	[glucose]/[1]	R _G , Å	D _{max} , Å	
trace ^a	0	6	18	_
2 ^b	0	8	20	
56	0	9	25	
15 ^b	0	18	49	
tracea	0.1	9	23	
2 ^b	0.1	11	29	

^a0.3, as determined by a Karl-Fischer titration. ^bBased on amounts actually added to the solution.

the interaction of the carbohydrate with the polar head group as a simple model for the binding of sugars to carbohydrate-binding proteins and for the interaction of phospholipids with sugars. We tried to form AOT inverse micelles by adding solid glucose monohydrate or 2 μ L of a 2 M aqueous D-glucose solution to a 50 mM solution of AOT in dry chloroform or cyclohexane. We were unable to detect any glucose in these solutions by NMR. Only when the [added water]/[glucose] ratio was >10 could we detect glucose in these solutions. In this case, the glucose was solubilized in the water pool and had no detectable interaction with the polar head groups. We posited that only an inverse micelle-forming surfactant that can participate in a hydrogen-bonding network can bind and solubilize a sugar in nonpolar solvents.¹¹ Indeed we found that 1, which is prepared by a Michael addition of

- (1) Bock, G.; Harnett, S. Carbohydrate Recognition in Cellular Function.
- Bock, G., Hallich, S. Carson, and S. Ciba Found. Symp. 1989, 145.
 Lemieux, R. U. Chem. Soc. Rev. 1989, 18, 347.
 Wassarman, P. M. Carbohydrate Recognition in Cellular Function.
- (3) wassarman, P. M. Carbonyarate Recognition in Cellular Function. *Ciba Found. Symp.* 1989, 145, 135.
 (4) (a) Widdas, W. F. *Biochim. Biophys. Acta* 1988, 947, 385. (b) Bali,
 D. L.; Horuk, R. *Biochim. Biophys. Acta* 1988, 947, 571.
 (5) (a) Farbie, C. H. J. P.; De Kruijff, B.; De Gier, J. *Biochim. Biophys. Acta* 1990, 1024, 380. (b) Carpenter, J. F.; Crowe, J. H. *Biochemistry* 1989, 28, 3916. (c) Hincha, D. K. *Biochim. Biophys. Acta* 1989, 987, 231.
 (6) Aoyama, Y.; Tanaka, Y.; Sugahara, S. J. Am. Chem. Soc. 1989, 111, 5397
- 5397
- (7) (a) Quiocho, F. A. Pure Appl. Chem. 1989, 61, 1293. (b) Quiocho,
 F. A. Annu. Rev. Biochem. 1986, 55, 287.
 (8) Fuhrhop, J.-H.; David, H.-H.; Mathieu, J.; Liman, U.; Winter, H.-J.;
- Boekema, E. J. Am. Chem. Soc. 1986, 108, 1785.
- (9) Fendler, J. Membrane Mimetic Chemistry; John Wiley and Sons: New York, 1982; p 48. (10) Luisi, P. L.; Steinman-Hofman, B. Methods Enzymol. 1987, 136, 188.
- (11) Aoyama, Y.; Tanaka, Y.; Sugahara, S. J. Am. Chem. Soc. 1989, 111, 5397.

7233

Department of Organic Chemistry.

¹Department of Polymer Research.



Figure 1. Distance distribution function P(r) for 50 mM solutions of 1 calculated from the X-ray data as described in the Experimental Section with (a) trace water, (b) [added water]/[surfactant] = 2, (c) [added water]/[surfactant] = 0.1. The maximum value of r for which P(r) is nonzero is D_{max} , the largest vector that can be contained within the micelle. The extent to which P(0) is not identically zero indicates the error in subtracting the solvent scattering.



Figure 2. Distance distribution function P(r) for a 50 mM solution of micelles of 1 with [added water]/[surfactant] = 15. The central dip may be indicative of the presence of an internal water pool.

thiomalic acid to bis(2-ethylhexyl)maleate and purified by crystallization from acetone, forms inverse micelles in chloroform or a 1:1 cyclohexane/chloroform mixture. We also found that 1 can solubilize glucose in these solvent systems at low water content. The results of small-angle X-ray scattering experiments on a 50 mM solution of 1 are consistent with the existence of inverse micelles. A semilogarithmic plot of the scattered intensity as a function of the square of the amplitude of the scattering vector exhibits a linear region. Linear regression analysis in this region and the Guinier approximation¹² give the z-average radius of gyration, R_{G} . Reducing the concentration by a factor of 2 and by a factor of 4 showed no obvious concentration effects on $R_{\rm G}$. The distance distribution function P(r), which defines the frequency of occurrence of vectors of a given length weighted by the electron density at either end of the vector within the aggregate, is derived by using the indirect transformation procedure,¹³ assuming a homogeneous size distribution. The maximum micelle dimension, D_{max} , and the radius of gyration are given in Table I, and selected P(r) curves are shown in Figures 1 and 2. Since the primary contribution to the contrast comes from the polar head groups, D_{max} is essentially determined by the head-group position within the micellar structure. The symmetry of these curves indicates that the structures are reasonably isometric and supports the assumption of a homogeneous size distribution.



Figure 3. 270-MHz ¹H NMR spectra of *p*-nitrophenyl β -D-galactopyranoside 2 in inverse micelles: (a) *p*-nitrophenyl β -D-galactopyranoside in a solution of 1 (50 mM) in CDCl₃/cyclohexane- d_{12} (1:1, v/v); (b) after the addition of water ([added water]/[surfactant] = 7).

Addition of solid D-glucose monohydrate to a micellar solution of 1, stirring for 5 h, and removal of suspended particles by centrifugation gave a clear solution. ¹H NMR studies clearly indicate that glucose is incorporated within the inverse micelles. The anomeric protons appear as broad signals at 5.3188 (α) and 4.6810 ppm (β). The relative integration ratio suggests that nine surfactant molecules are required to solubilize one glucose molecule. Upon addition of D₂O ([added water]/[surfactant] ratio is 2), the broad sugar signals sharpen, indicating that the glucose in the interior of the reverse micelle is being hydrated and thus gains some motional freedom. The HOD peak shifts downfield, indicating the presence of both free and bound water.¹⁴

Figure 3 shows the ¹H NMR spectra of *p*-nitrophenyl β -Dgalactopyranoside solubilized in a 50 mM solution of 1 in the same solvent system. The two sets of aromatic protons, as well as the anomeric one, can be observed by 'H NMR. In the micelles, which were formed without the addition of water, $H_{1'}$, $H_{2'}$, and H₁ appear as broad signals (spectrum a), which sharpen and are shifted downfield upon addition of water (spectrum b, [added water]/[surfactant] = 7). The signal for $H_{2'}$ shifts by 0.03 ppm while that for $H_{i'}$ shifts by only 0.01 ppm. The anomeric proton shifts by 0.07 ppm, from 4.93 to 5.00 ppm (its chemical shift in D_2O is 5.03 ppm). This probably indicates that the carbohydrate part of the molecule is being dragged into the aqueous environment as the inverse micelle swells with water while the hydrophobic aromatic ring is still in the less polar shell of the micelle. The chemical shift for the aromatic protons in p-nitrophenol are sensitive to their environment as was observed in the binding of p-nitrophenolates to cyclodextrin. Upon binding, the protons ortho to the nitro group were shifted by 35 Hz whereas the protons meta to the nitro are shifted by only 14 Hz, indicating that the molecule enters into the cyclodextrin cavity with its "nitro end" first.¹⁵ The ¹³C NMR spectra of D-glucose-1-¹³C (90% enriched) in

The ¹³C NMR spectra of D-glucose-l-¹³C (90% enriched) in the inverse micelle are substantially different from the corresponding resonances of glucose in water both with respect to the chemical shift and the line width. As can be seen in Figure 4 both signals, corresponding to the α and β anomers, are shifted downfield and are broadened to a large extent. This shift and line broadening are very similar to the ones observed for carbohydrates bound to lectins,¹⁶ indicating perhaps a similar chemical

⁽¹²⁾ Guinier, A.; Fournet, G. Small Angle Scattering of X-Rays; Wiley: New York, 1955.

⁽¹³⁾ Glatter, O. J. Appl. Crystallogr. 1977, 10, 415.

⁽¹⁴⁾ De Marco, A.; Menegatti, E.; Luisi, P. L. J. Biochem. Biophys. Methods 1986, 12, 325.

⁽¹⁵⁾ Bergeron, R. J.; Channing, M. A.; Gibeily, G. J.; Pillor, D. M. J. Am. Chem. Soc. 1977, 99, 5146.



Figure 4. 125.75-MHz ¹³C NMR spectra of D-glucose-1-13C (90% enriched): (a) in D_2O ; (b) in 50 mM solution of dry 1 in $CDCl_3/cyclo-hexane$ (1:1, v/v). Sample preparation and spectral parameters are described in the Experimental Section.

Table II. Amplitude Ratio and Hyperfine Coupling Constant for a Water-Soluble Spin Probe in Water and in 50 mM Solutions of 1 with Various Water Contents

[water]/[1] ^a	h(0)/h(+)	h(-)/h(+)	hyperfine coupling, G	
H ₂ O	1.04	0.92	16.87	
0	0.97	0.52	15.75	
2	0.90	0.34	15.75	
4	0.80	0.25	15.88	
6	0.76	0.24	15.88	
8	0.82	0.29	15.88	
10	0.77	0.29	16.25	

"Based on amounts actually added to the solution. [1] = 50 mM for all except the water only experiment.

environment. Upon addition of water the peaks shift upfield and become sharper.

Additional indication that the carbohydrates are bound in the inverse micelle was provided by an ESR study of the water-soluble 6-deoxy(6-TEMPO)-4'-(aminocarbonyl)-D-glucose17 solubilized in a 50 mM solution of 1.

The line shape, line width, and hyperfine splitting of the spectra (Figure 5) suggest a partition of the spin probe between environments of different polarity.¹⁸ The line shapes can be quantified by the amplitude ratio, h(0)/h(-1) and h(-1)/h(+1), where h(-1), h(0), and h(+1) are the heights of the low-, center-, and high-field signals, respectively. These as well as the hyperfine coupling for the probe in a 50 mM solution of 1 are given in Table II. As the water content of the sample increases, the observed hyperfine splitting increases, indicating an increasing population in the more polar environment, in agreement with our observations in the NMR study. These values are indicative of a weakly immobilized spin probe and are very similar to values observed for nitroxides in a polymer network.¹⁹ In contrast, the spectra of the waterinsoluble probe N-acetyl-4-amino-TEMPO run under the same conditions (50 mM of surfactant in 1:1 chloroform/cvclohexane solution) show no dependence on the water content of the inverse micelles and, in fact, seem very similar to the spectra in the solvent itself.20-22



Figure 5. ESR spectra of 6-deoxy(6-TEMPO)-4'-(aminocarbonyl)-Dglucose (3). The solution was prepared as described in Figure 1. Key: (a) in water; (b) in inverse micelles of 1, prepared as described in the Experimental Section; (c) As for (b) after addition of water ([added water]/[surfactant] = 9).

Conclusion

The results obtained indicate that surfactant molecule 1 in a 1:1 chloroform/cyclohexane mixture at a concentration of 50 mM organizes into reasonably isometric structures with a size that depends upon the amount of water and/or carbohydrate present. Addition of carbohydrates causes the size of the micelle to increase, for a given amount of water. It is likely that the chemical environment in the interior of these aggregates, which is composed of carboxylic acid and carboxylates, resembles the environment in the binding site of carbohydrate-binding proteins. Our NMR and ESR data indicate that monosaccharide derivatives are being extracted into this environment and are bound, at low water content, to the polar head groups. In contrast bis(2-ethylhexyl) sulfosuccinate,²³ which can solubilize large quantities of water in isooctane, can solubilize glucose in chloroform only at high water content ([added water]/[sugar] > 10). This fact further supports our conclusion regarding the importance of hydrogen bond formation in solubilization and binding of monosaccharides.

Experimental Section

¹H NMR spectra were measured on a Bruker WH 270 spectrometer. Chemical shifts were measured relative to TMS. ¹³C NMR spectra were recorded on a Bruker AM500 spectrometer, operating at 125.75 MHz. Chemical shifts were measured relative to dioxane as an external standard (66.5 ppm). IR spectra were measured on a Nicolet 510 FT-IR spectrometer. ESR spectrum were measured on a Varian E3 instrument. Elemental analysis were performed in the microanalysis laboratory of the Hebrew University, Jerusalem. FAB-MS was recorded on a TSQ-70B, Finnigan spectrometer.

D-Glucose and p-nitrophenyl β -D-galactopyranoside were purchased from Aldrich and were used without further purification. 6-Deoxy(6-TEMPO)-4'-(aminocarbonyl)-D-glucose was prepared according to a reported procedure.17 Sodium 3-[[1,2-bis[[(2-ethylhexyl)oxy]carbonyl]ethyl]thio]succinate was prepared by a modification of a literature procedure;8 thiomalic acid (0.8 g, 5.3 mmol) was added to a solution containing bis(2-ethylhexyl) maleate (1.8 g, 5.29 mmol), sodium

⁽¹⁶⁾ Berman, E.; Brown, J.; Lis, H.; Sharon, N. Eur. J. Biochem. 1985, 152, 447

⁽¹⁷⁾ Sosnovsky, G.; Rao, N. U. M.; Lukszo, J.; Brasch, R. C. Z. Naturforsch. 1986, 41B, 1293. (18) Jost, P. C.; Griffith, O. H. Methods Enzymol. 1978, 49, 369.

⁽¹⁹⁾ Schlick, S.; Harvey, R. D.; Alonso-Amigo, M. G.; Klempner, D. Macromolecules 1989, 22, 822.

⁽²⁰⁾ Yoshioka, H. J. Colloid Interface Sci. 1983, 95, 82.

⁽²¹⁾ Menger, F. M.; Suito, G.; Sanzero, G. V.; Dodd, J. R. J. Am. Chem. Soc. 1975, 97, 909.
(22) Lim, Y. Y.; Fendler, J. H. J. Am. Chem. Soc. 1978, 100, 7490.
(23) Luisi, P. L.; Straub, B. E. Reverse Micelles; Plenum Press: New Neth Oct. York, 1984.

hydroxide (2.12 g, 5.29 mmol), and water (5 mL) in dimethylformamide (30 mL), and the reaction mixture was stirred overnight. The solvent was then removed under reduced pressure, the crude product was dissolved in water, and the pH was adjusted to 7.0 with 1 M HCl. The water was then removed under reduced pressure and the crude product was recrystallized twice from acetone to give 1.35 g of a white powder (50% yield). ¹H NMR (CDCl₃): 3.97 (br m, 5 H, OCH₂ and CHS), 3.68 (br m, 1 H, CHS); 2.91, 2.79, (2 H, CH₂CO, AB type); 2.55, 2.38 (2 H, CH₂CO, AB type); 1.54 (br m, 2 H, CH); 1.26 (br m, 16 H, CH₂); 0.85 (br m, 12 H, CH₃) ppm. Anal. Calcd for C₂₄H₄₁NaO₈S: C, 56.3; H, 8.3. Found: C, 56.4; H, 8.0. FAB-MS: m/z 513.2 (M⁺ + H).

NMR Experiments. D-Glucose (3 mg) was added to a solution of 1 (50 mL) in a 1:1 (v/v) mixture of cyclohexane- d_{12} and CDCl₃,²⁴ and the mixture was stirred for 2 h. It was then filtered, centrifuged at 8000g for 3 min, and put into a dry NMR tube. The H₁ α and H₁ β resonances at 5.3188 and 4.6810 ppm and with line widths of 31 and 27 Hz, respectively, could be observed.

Following this, $2 \ \mu L$ of D_2O was added (the total volume was 0.5 mL). The HOD peak shifted so that only $H_1\alpha$ could be detected at 5.2920 ppm with a line width of 10.15 Hz (in addition, the ring protons and H_6 appear as unresolved multiplets at 3-4.5 ppm). After further addition of 6 μL D_2O , when the micelles were close to collapse, the chemical shift was 5.2455 ppm with no change in the line width.

p-Nitrophenyl β -D-galactopyranoside (2 mg) was added to a solution of 1 (50 mM) in CDCl₃ as described above for glucose, and the ¹H NMR spectrum was recorded. The signal corresponding to H₁, was slightly broadened, whereas H₂, was very broad. Upon addition of 2 μ L of D₂O, the whole spectrum sharpened and there was a change in the chemical shifts of H₂ and the anomeric proton but not of H₁. Chemical shifts (in ppm) for p-nitrophenyl β -D-galactopyranoside in various environments are as follows:

proton	D ₂ O	1 in CDCl ₃	1 in CDCl ₃ + $2 \mu L$ of D ₂ O
H ₁	5.03	4.93	5.00
H,	7.06	7.13	7.16
H _{2′}	8.08	8.12	8.13

¹³C NMR Experiments. D-Glucose- $l^{-13}C$ (90% enriched; 2 mg) was added to 0.5-mL solution of 0.125 M 1 in CDCl₃; the solution was then transferred into a 5-mm NMR tube and a 125.75-MHz spectrum was measured. A total of 26 000 scans were collected with a spectral window of 29412 Hz and 16 384 data points. The FID was then Fourier transformed with zero filling to 32 768 data points and line broadening of 20 Hz in order to enhance the signal to noise ratio. The chemical shift for the C₁ β is 96.6848 ppm, and that for C₁ α is 92.4083 ppm. Line widths are 95 and 50 Hz, respectively.

In a similar experiment, a solution of D-glucose- $l^{-13}C$ (90% enriched) 2 M in D₂O was added (1 μ L) to a 0.7-mL solution of 1 (0.05 M) in CDCl₃. The chemical shift for the C₁ β is 96.1973 ppm and the line width is 29 Hz; C₁ α appears at 92.1344 ppm with a line width of 14 Hz. For β -D-glucose in D₂O, the chemical shift was found to be 95.8168 ppm with a line width of 4.5 Hz, and for α -D-glucose, the resonance of anomeric carbon is at 91.9978 ppm with a line width of 4.5 Hz.

ESR Experiments. Solutions were prepared by an addition of the solid spin probe to the micellar solution as described for glucose. ESR measurements were performed on a Varian E3 instrument, operating at 9.5 GHz with 100-KHz modulation. Instrument parameters were 10-mW microwave power, 0.8-G modulation amplitude, 3386-G applied field, 100-G scan range, 8-min scan time, and 10-s time contact.

Small-Angle \bar{X} -ray Scattering. Solutions were prepared as described for the NMR experiment and were flame-sealed in a 1.5-mm-diameter quartz (or Li glass) X-ray capillary. A Philips sealed-tube X-ray generator operating at approximately 45 kV and 55 mA produced Mo X radiation, which was monochromated by a Zr filter and a single Franks mirror and collimated by a series of slits and height limiters. The scattering profile was recorded by a linear position sensitive detector of the delay line type²⁵ and stored in a Z-80 based microprocessor unit as a 256-channel histogram. On completion of each experiment, the data were sent to an IBM 3090 computer for further analysis. The duration of each experiment was 8 h.

For each specimen, two parallel experiments were run—the micellar solution and the solvent alone, all other parameters remaining unchanged. The solvent scattering curve was subtracted from that of the micellar solution after normalization with respect to the high-angle background intensity. Data were desmeared by use of the Lake procedure²⁶ and Guiner analysis¹² was performed to obtain the triaxial radius of gyration R_G . Analysis of the background-subtracted, smeared profile was also performed using the indirect transformation procedure (ITP) of Glatter.¹³ The unsmoothed data were deconvoluted from the beam shape, and the noise was treated as statistical in origin. The spline fit curve I(q) was then used to calculate the distance distribution function

$$P(r) = (\frac{1}{2}\pi) \int_0^\infty I(q) q^2 (\sin qr/qr) \, \mathrm{d}q$$
 (1)

where $q = 4\pi/\lambda \sin \theta$, 2θ is the scattering angle, and $\lambda = 0.71$ Å for Mo radiation. This expression gives the frequency of occurrence of vectors of length r, weighted by the electron density at either end of the vector. The radius of gyration may be obtained from

$$R_{\rm G}^2 = \int_0^\infty P(r)r^2 \, {\rm d}r/2 \, \int_0^\infty P(r) \, {\rm d}r \qquad (2)$$

The infinite integrals were approximated by finite sums, restrictions on sample intervals and range $(q = 0.056-0.4427 \text{ Å}^{-1})$ being in accordance with information theory. Only small differences were observed between R_G determined by the Guinier method and that calculated by using eq 2. The particle shape is reflected in the shape of P(r) and the presence of interparticle interference would be discernible.

Acknowledgment. We thank R. de Roos for expert technical assistance; Dr. O. Glatter for kindly making the ITP program available to us; Prof. S. Schlick for useful discussions, and the Fund of Basic Research, administered by the Israel Academy of Sciences and Humanities, for financial support.

⁽²⁴⁾ This solvent system was chosen because it is difficult to perform SAXS experiments in chloroform due to its high electron density, and it is not desirable to use cyclohexane since 1 gives an obviously nonhomogeneous size distribution of micelles in this medium.

⁽²⁵⁾ Reich, M. H.; Kam, Z.; Eisenberg, H. Biochemistry 1982, 21, 5189. (26) Lake, J. A. Acta Crystallogr. 1967, 23, 191.