



**HAL**  
open science

## Asymmetrical, Symmetrical, Divalent, and Y-Shaped (Bola)amphiphiles: The Relationship between the Molecular Structure and Self-Assembly in Amino Derivatives of Sophorolipid Biosurfactants

Niki Baccile, Elisabeth I. P. Delbeke, Martha Brennich, Chloé Seyrig, Jonas Everaert, Sophie L. K. W. Roelants, Wim Soetaert, Inge N. A. Van Bogaert, Kevin M. Van Geem, Christian V. Stevens

### ► To cite this version:

Niki Baccile, Elisabeth I. P. Delbeke, Martha Brennich, Chloé Seyrig, Jonas Everaert, et al.. Asymmetrical, Symmetrical, Divalent, and Y-Shaped (Bola)amphiphiles: The Relationship between the Molecular Structure and Self-Assembly in Amino Derivatives of Sophorolipid Biosurfactants. *Journal of Physical Chemistry B*, American Chemical Society, 2019, 123 (17), pp.3841-3858. 10.1021/acs.jpcc.9b01013 . hal-02128241

**HAL Id: hal-02128241**

**<https://hal.sorbonne-universite.fr/hal-02128241>**

Submitted on 14 May 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Asymmetrical, symmetrical, divalent and Y-shaped  
(bola)amphiphiles: the relationship between molecular structure  
and self-assembly in amino derivatives of sophorolipid  
biosurfactants

Niki Baccile,<sup>a,\*†</sup> Elisabeth I. P. Delbeke,<sup>b,c,†</sup> Martha Brennich,<sup>d</sup> Chloé Seyrig,<sup>a</sup> Jonas Everaert<sup>b,e</sup> Sophie L. K. W. Roelants<sup>e,f</sup> Wim Soetaert<sup>e,f</sup> Inge N. A. Van Bogaert,<sup>g</sup> Kevin M. Van Geem,<sup>c</sup> Christian V. Stevens<sup>b,\*</sup>

<sup>a</sup> Sorbonne Université, Centre National de la Recherche Scientifique, Laboratoire de Chimie de la Matière Condensée de Paris, LCMCP, F-75005 Paris, France.

<sup>b</sup> SynBioC, Department of Green Chemistry and Technology, Ghent University, Coupure Links 653, 9000 Ghent, Belgium.

<sup>c</sup> Laboratory for Chemical Technology (LCT), Department of Materials, Textiles and Chemical Engineering, Ghent University, Technologiepark 125, 9052 Ghent, Belgium.

<sup>d</sup> European Molecular Biology Laboratory, Synchrotron Crystallography Group, 71 Avenue des Martyrs, 38042 Grenoble, France

<sup>e</sup> InBio, Department of Biotechnology, Ghent University, Coupure Links 653, 9000 Ghent, Belgium.

<sup>f</sup> Bio Base Europe Pilot Plant (BBEU), Rodenhuisenkaai 1, 9042 Ghent (Desteldonk), Belgium.

<sup>g</sup> Centre for Synthetic Biology, Department of Biotechnology, Ghent University, Coupure Links 653, 9000 Ghent, Belgium.

\* Corresponding authors: [niki.baccile@sorbonne-universite.fr](mailto:niki.baccile@sorbonne-universite.fr); [chris.stevens@UGent.be](mailto:chris.stevens@UGent.be)

† These authors equally contributed to this work

**Abstract:**

Conventional *head-chain* but also more exotic divalent, gemini or bolaform amphiphiles have in common well-defined hydrophilic and hydrophobic blocks with often a predictable self-assembly behaviour. However, new categories of amphiphiles, like microbial biosurfactants, challenge such

conventional understanding, because of the poorly-defined boundaries between the hydrophilic and hydrophobic portions. Microbial glycolipids, like sophorolipids, rhamnolipids or cellobioselipids, interesting biodegradable, nontoxic, alternatives to synthetic surfactants, all represent interesting examples of atypical amphiphiles with partially predictable self-assembly properties. However, their limited molecular diversity strongly limits their application potential. For this reason, we used them as ready-made platform to prepare a whole class of new derivatives. In particular, a broad range of amino derivatives of sophorolipid biosurfactant was recently prepared with the goal of producing biobased antimicrobial and transfection agents, of which the efficiency strongly depends on their molecular structure and unpredictable self-assembly behavior. The new compounds contain a set of asymmetrical and symmetrical bolaamphiphiles, the latter with three or four hydrophilic centers, divalent amphiphiles with asymmetric polar headgroups and even a Y-shaped amphiphiles, bearing two sophorose groups connected to one nitrogen atom. In this contribution we employ small angle x-ray scattering (SAXS) to establish a relationship between their peculiar molecular structures and the self-assembly properties in water. We find that all divalent and Y-shaped compounds form micelles, of which the hydrophilic shell is composed of a bulky sophorose- $C_x(x= 8,11)$ -amine moiety, with aggregation numbers between 30 and 100. On the contrary, most symmetrical and asymmetrical bolaamphiphiles display poor self-assembly properties, generally showing aggregation numbers below 20, especially in the presence of either short spacers or large spacers containing hydrophilic centers.

## **Introduction:**

In view of the transition towards a bio-based economy which has been initiated in the last two decades, bio-derived amphiphiles are regarded as interesting alternatives for classical, fossil-based surfactants.<sup>1-4</sup> In this respect, sophorolipids gained a lot of interest. They are glycolipid biosurfactants, consisting of sophorose (glucose  $\beta(1,2)$ ) as the hydrophilic headgroup and a fatty acid (mostly oleic acid) as the lipid tail.<sup>5,6</sup> Sophorolipids are produced by different yeast species via fermentation, with *Starmerella bombicola* being the standard production organism. Microbial production results in the production of different sophorolipid derivatives, which all differ in the number of acetyl groups on the sugar headgroup, the number of unsaturations in the lipid tail, the position of the glycosidic bond on the lipid tail ( $\omega$  or  $\omega-1$ ) and whether or not lactonization occurred towards respectively the closed lactone or open acid form. The major fermentation products are

diacetylated C18:1 sophorolipid lactone **A** and C18:1 sophorolipid acid **B** (Figure 1). Despite this apparent broad variety of compounds, the term sophorolipid generally refer to the deacetylated C18:1 open acidic and/or partially acetylated close lactone form, often coexisting together as well as with 10 % to 15 % of other congeners.<sup>7</sup> By far, acidic C18:1 sophorolipids are the most studied compound for its ease to recover after simple basic hydrolysis.

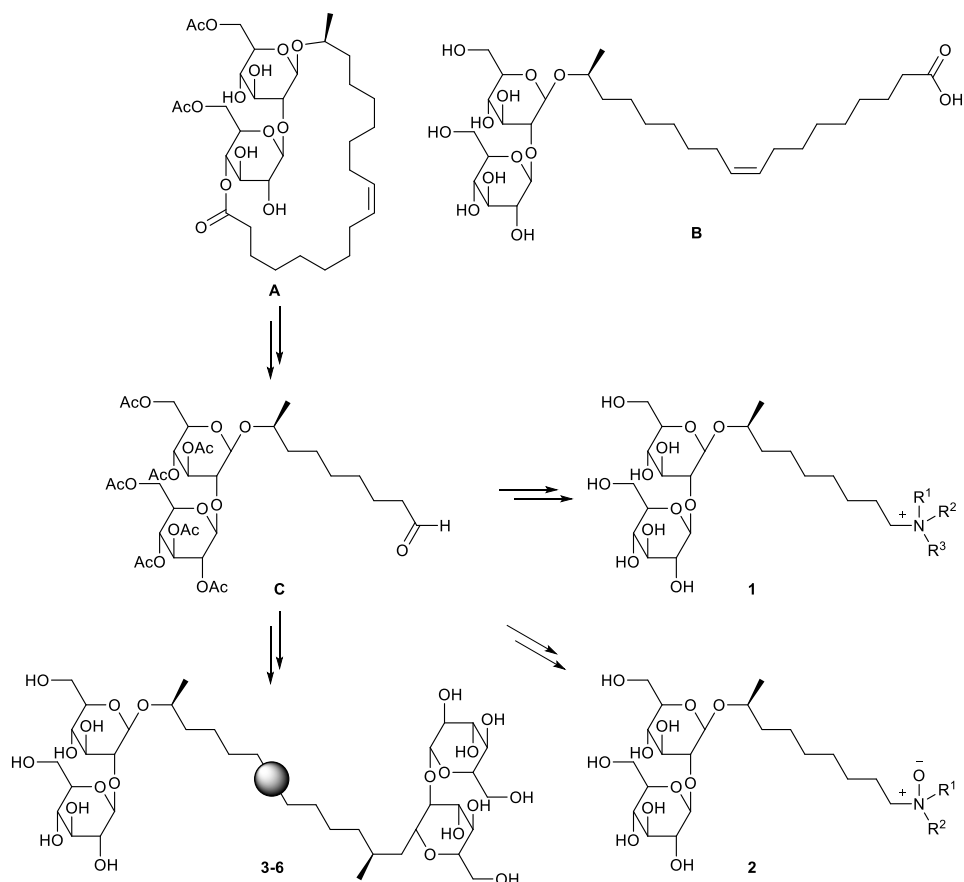
Acidic C18:1 sophorolipids can lower the surface tension of water from 72.8 to 30-40 mN/m and have a critical micelle concentration of 40-100 mg/L.<sup>5,8,9</sup> If the first reports on the self-assembly of sophorolipids have shown the possibility to form micelles as well as helical structures,<sup>10,11</sup> more recent work has pointed out the importance of the batch uniformity on the phase behaviour.<sup>7</sup> In fact, several recent works have shown the crucial impact of the lipid structure (number of glucose moieties, double, mono and no unsaturation, isomerization) on the phase behavior and in particular on the formation of chiral ribbons, bilayers, vesicles,<sup>12-17</sup> with a crucial importance of pH.<sup>14,18,19</sup> This short survey shows that, despite the rich phase behavior of this class of compounds, the diversity in terms of molecular structure remains very limited and generally confined to the general formula sugar-*spacer*-COOH, where the spacer is commonly a C18:1 or C18:0 hydrocarbon, although more complexity can be found in rhamnolipids and cellobioselipids.<sup>20,21</sup>

To explore new properties and broaden the application potential of the natural sophorolipids, several chemical derivatives including alkyl esters (better emulsifiers)<sup>22-25</sup> and polymers<sup>26</sup> have been recently developed, just to cite some.<sup>27</sup> One possible approach consists in modifying the double bond in the C18:1 backbone.<sup>28</sup> Several chemical modification paths were developed by us starting from the diacetylated C18:1 sophorolipid lactone **A** (Figure 1). In a first step, the synthesis of a functionalized sophorolipid aldehyde building block **C** was accomplished via an ozonolysis reaction to transform the double bond in a reactive site.<sup>29</sup> This modification reduces the chain length of the sophorolipid derivatives, hereby increasing their hydrophilic character. In a second step, the sophorolipid aldehyde building block **C** was transformed into a broad library of innovative sophorolipid derivatives. On the one hand, a whole range of nitrogen containing sophorolipid derivatives was synthesized, including sophorolipid quaternary ammonium salts **1** and sophorolipid amine oxides **2**.<sup>29,30</sup> The sophorolipid aldehyde building block **C** was used for the synthesis of a new class of nitrogen containing sophorolipid bolaamphiphiles **3-6**.<sup>31</sup> The modifications were also extended to petroselinic acid based sophorolipids to increase



variation in the lipid tail.<sup>32</sup> Most of these new compounds have only been studied in terms of their antimicrobial and plasmid transfection properties, but a recent work from our group<sup>31</sup> has shown that these properties are observed only for those sophorolipid derivatives which are able to form micellar aggregates. These results demonstrate the importance of acquiring a better knowledge over these new compounds (**1** to **6**, Figure 1), in light of their non-conventional molecular structures (symmetrical and asymmetrical bolaamphiphiles, divalent, Y-shaped), which challenge the conventional idea of a classical *head-tail*, but also divalent, gemini and bolaform surfactants, all characterized by well-defined (hydro)philic/phobic portions.<sup>33-41</sup> On the contrary, the hydrophilic/hydrophobic portions of the atypical amphiphiles studied here are much less defined, making unpredictable their self-assembly behavior, otherwise crucial to evaluate their application potential.

Compounds shown in Figure 2 contain one or two sophorose groups and one or two amine (or ammonium and amine oxide) centers, bearing several side groups. We identify the molecular geometries as follows: asymmetrical (**1a-1l** and **2a-2g**) and symmetrical (**3a-3d**, **4a-4d**, **5a-5d**, **6a-6c**) compounds; among the asymmetrical compounds, we identify the quaternary ammonium (**1a-1l**) and amine oxide derivatives (**2a-2g**), each group being characterized by a bolaform (**1a-1e**, **2a-2f**) and divalent (**1f-1l**, **2g**) molecular structure. Within the symmetrical compounds, we identify the ammonium/amine bolaform derivatives bearing three (**3a-3d**, **4a-4d**) and four (**5a-5d**, **6a-6c**) hydrophilic centers and two Y-shaped (**3d**, **4d**) amphiphiles. These compounds represent a set of non-conventional amphiphiles, if compared to a standard (bola)amphiphile, generally characterized by well-defined hydrophilic-hydrophobic(-hydrophilic) portions:<sup>41-43</sup> bolaamphiphiles **3-6** have several hydrophilic centers in their spacer while divalent **1f-1l**, **2g** and Y-shaped **3d**, **4d** amphiphiles bear an atypical sophorose-C<sub>x</sub>(x= 8,11)-amine headgroup. In all cases, the hydrophilic/hydrophobic character of each portion of the molecule is unpredictable, and does not allow to draw immediate conclusions on the global self-assembly properties.<sup>33,43,44</sup> In addition, the atypical bulkiness of the sophorose-C<sub>x</sub>(x= 8,11)-amine headgroup may screen the positive charged localized on the ammonium group, having unexpected effects on the micelles-micelles interactions.<sup>45</sup> In this work, we employ Small Angle X-ray Scattering to understand the impact of structure and chemical composition of these nonconventional amphiphiles on their aqueous self-assembly properties.



**Figure 1 - Diacetylated C18:1 sophorolipid lactone (A) and C18:1 sophorolipid acid (B). Chemical modification pathway towards sophorolipid quaternary ammonium salts (1), sophorolipid amine oxides (2) and sophorolipid bolaamphiphiles (3-6)**

### Materials and Methods:

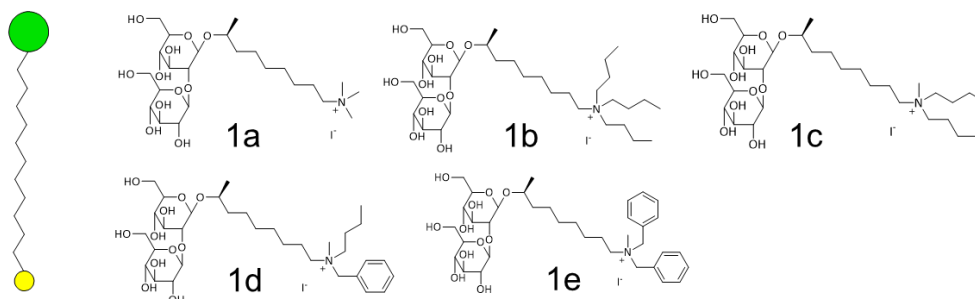
*Products:* all sophorolipid derivatives have been used as such, and their synthesis and characterization has been reported in previous studies. Quaternary ammonium sophorolipids **1a-1l** have been described in ref. 29,<sup>46</sup>. One should note that samples **1a-1g**<sup>29</sup> were deprotected with NaOH while samples **1h-1l**<sup>46</sup> were deprotected with Et<sub>3</sub>N. For comparison, samples **1f** and **1g** were also prepared in double using Et<sub>3</sub>N deprotection. Amine oxide sophorolipids **2a-2g** are described in ref. 30. Symmetric bolaform sophorolipids with four hydrophilic centers **5a-5d** and **6a-6c** have been reported in ref. 31. Symmetric bolaform sophorolipids with three hydrophilic centers **3a-3c** and **4a-4c** and Y-shaped **3d**, **4d** sophorolipid are reported in ref. 31. Amine oxide samples **2**<sup>30</sup> and the entire set of samples **3-6**<sup>31</sup> were all deprotected with Et<sub>3</sub>N. Water is of milliQ grade and HCl and NaOH solutions to modify the pH are prepared in water.

*General procedure for NaOH deprotection:* in a flame-dried round-bottomed flask, the peracetylated sophorolipid quaternary ammonium salt was dissolved in dry methanol and sodium methoxide (0.15 eq) was added. The reaction mixture was stirred overnight at room temperature, concentrated under reduced pressure and recrystallized from acetone to yield the pure deprotected sophorolipid quaternary ammonium salt **1**.

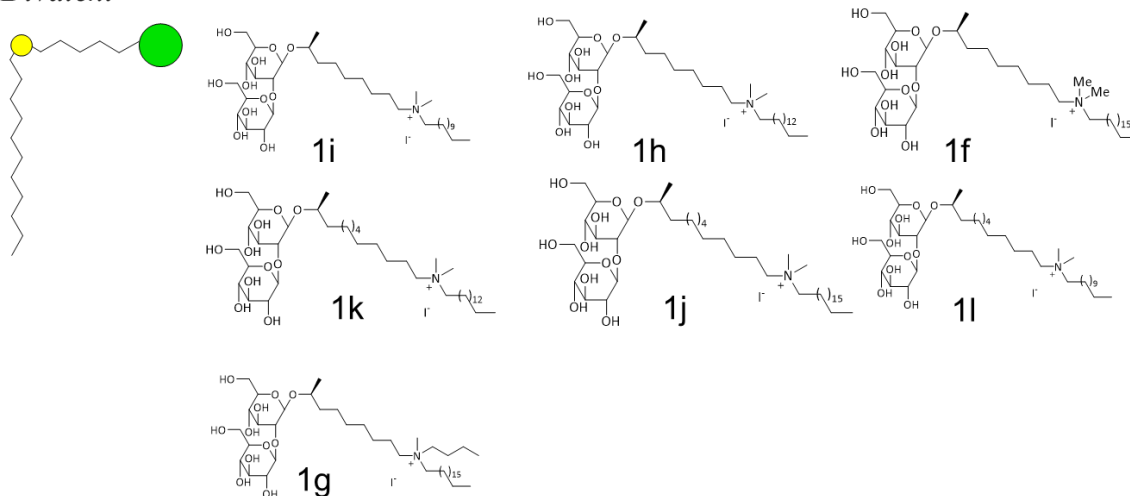
*General procedure for Et<sub>3</sub>N deprotection:* in a 25 mL flask, the protected quaternary ammonium sophorolipid (0.45 mmol, 1 eq) was dissolved in a methanol/water mixture (1:1) and 13 mL Et<sub>3</sub>N (0.90 mmol, 2 eq) was added. The mixture was stirred for 2 h at reflux temperature and concentrated under reduced pressure to yield the pure deprotected quaternary ammonium sophorolipid **1**.

### Quaternary ammonium derivatives of oleic and petroselinic acid SL

#### *Bolaform (asymmetric)*



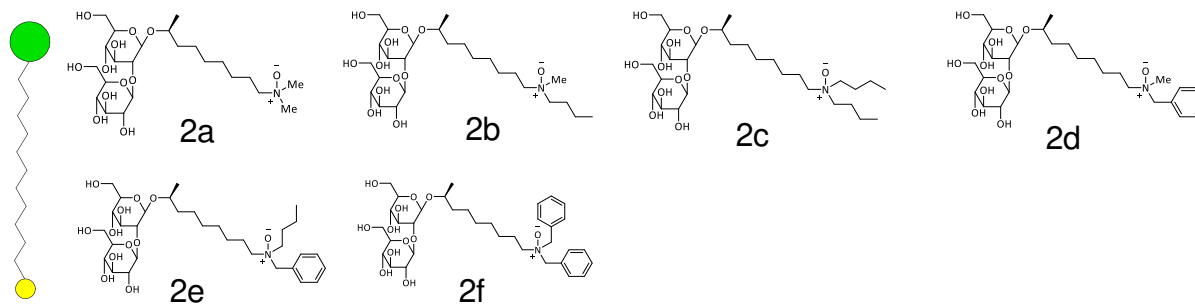
#### *Divalent*



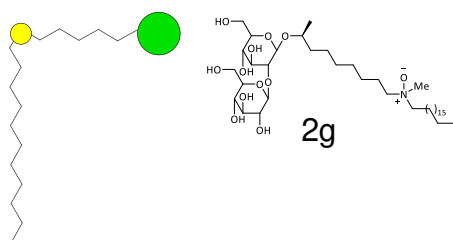
a)

## Amine oxides derivatives of oleic acid sophorolipids

### *Bolaform (asymmetric)*



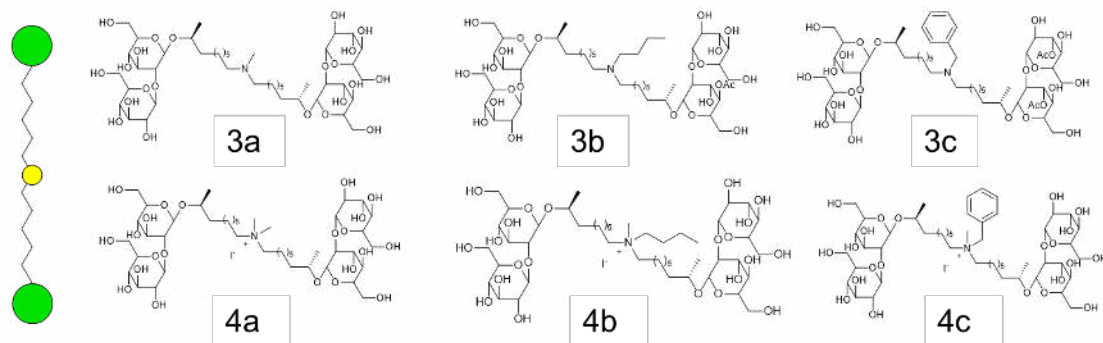
### *Divalent*



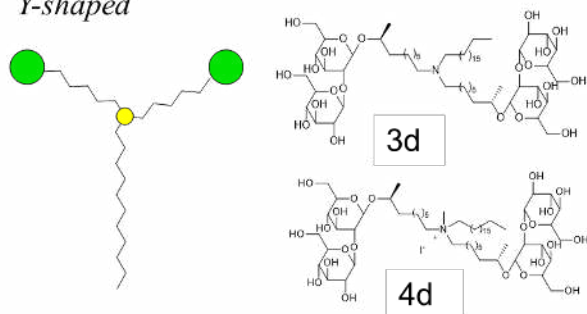
b)

## Quaternary ammonium derivatives of oleic acid SL

### *Bolaform (symmetrical) with three hydrophilic centers*



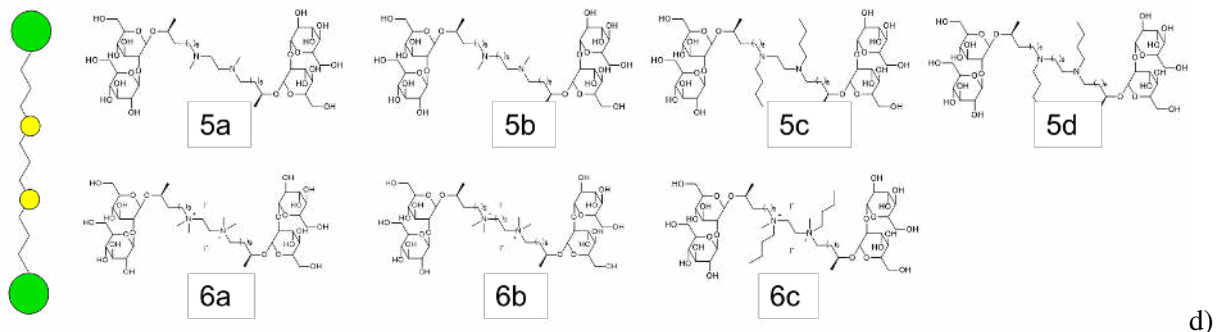
### *Y-shaped*



c)

## Quaternary ammonium derivatives of oleic and petroselinic acid SL

### *Bolaform (symmetrical) with four hydrophilic centers*



d)

**Figure 2 – Sphorolipid derivatives studied in this work.**

*Preparation method for solution analyzed by SAXS:* all solutions are freshly prepared in water at room temperature prior to analysis. For all compounds in a concentration series, the samples were first solubilized at a concentration of 100 mg/mL and then diluted by half in a concentration series up to 0.78 mg/mL. All concentrations (100, 50, 25, 12.5, 6.25, 3.13, 1.56 and 0.78 mg/mL) were analyzed at a total volume of 100  $\mu$ L in a 96-well plate. For all compounds in a pH series, the samples were first solubilized at a concentration of 10 mg/mL and were then adjusted to a pH of 11, 9, 7, 5 and 3 by addition of concentrated HCl or NaOH solutions (generally 0.1 M, 0.5 M or 1 M).

*Critical micelle concentration (cmc) experiments:* cmc experiments have been performed using A zetasizer Nano ZS instrument (Malvern-Panalytical) working at 633 nm in the forward scattering configuration ( $13^\circ$ ). All experiments have been collected at constant shutter open in the “no shutter” configuration. The scattered intensity is collected in terms of kilo counts per second. Experimentally, a concentrated solution (generally  $> 5$  mg/mL) of the given compound is analyzed and further diluted until the scattered intensity is close to zero. Data are linearly fitted before and after the inflection point and the cmc is determined at the intersection of the fit lines. Raw data are shown in Figure S 1 and cmc values are given in Table S 1.

*Cryogenic Transmission Electron Microscopy (Cryo-TEM).* These experiments were carried out on an FEI Tecnai 120 twin microscope operating at 120 kV equipped with a Gatan Orius CCD numeric camera. The sample holder was a Gatan Cryoholder (Gatan 626DH, Gatan). Digital

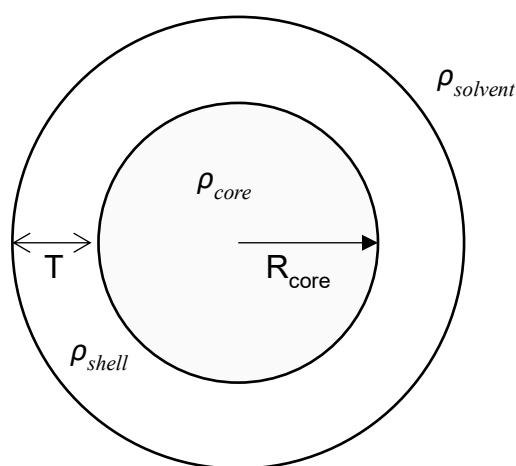
Micrograph software was used for image acquisition. Cryofixation was done on a homemade cryofixation device. The solutions were deposited on a glow-discharged holey carbon coated TEM copper grid (Quantifoil R2/2, Germany). Excess solution was removed and the grid was immediately plunged into liquid ethane at  $-180\text{ }^{\circ}\text{C}$  before transferring them into liquid nitrogen. All grids were kept at liquid nitrogen temperature throughout all experimentation.

*Small Angle X-ray Scattering (SAXS):* SAXS experiments are performed at  $25^{\circ}\text{C}$ , immediately after sample preparation, on the BioSAXS BM29 beamline at the ESRF synchrotron facility (Grenoble, France) using 12.5 KeV energy and a sample-to-detector distance of 2.867 m, imposed by the beamline standard configuration. The energy is calibrated by measuring the  $L_{\text{I}}$  and  $L_{\text{III}}$  edges of platinum and the sample-to-detector distance is determined using silver behenate ( $d_{\text{ref}} = 58.38\text{ \AA}$ ).<sup>47,48</sup> For this experiment, we employ the automatic sample changer for liquids using the 96-well plates and about 100  $\mu\text{L}$  of each sample.<sup>49</sup> The liquid sample is automatically loaded into a 1.8 mm quartz glass capillary and ten acquisitions of 1 s each are taken as the sample passes the beam. Individual frames are manually controlled for systematic changes and averaged for better statistics. Eventual changes can be either due to intrinsic sample heterogeneity or radiation damage. The signal of the Pilatus 1M 2D detector, used to record the data, is integrated azimuthally with PyFAI to obtain the scattered intensity  $I(q)$  vs.  $q$  spectrum ( $q = 4\pi \sin \theta / \lambda$ , where  $2\theta$  is the scattering angle) after masking systematically wrong pixels and the beam stop shadow.<sup>50</sup> Absolute intensity units were determined by measuring the scattering signal of water ( $0.0163\text{ cm}^{-1}$ ).

*Analysis of the SAXS data.* The general scattering function is defined as  $I(q) \propto (\Delta\rho)^2 P(q)S(q)$ , where  $\rho$  is scattering length density involved in the contrast between the solvent and the object,  $P(q)$  is the form factor and  $S(q)$  is the structure factor.<sup>51</sup> In this work, we have used two main approaches to analyze the SAXS data. In the absence of any hypothesis on the morphology of the self-assembled form of the sophorolipid derivatives, we use a combination of model-independent

Guinier  $\{I(q) = I(0)e^{\left[\frac{-q^2 R_g^2}{3}\right]}, qR_g < 1, R_g = \text{radius of gyration}\}$  and power law  $\{I(q) \propto q^{-m}, m = \text{dimensionless exponent} > 0\}$  analyses for, respectively, the mid- $q$  and low- $q$  range.

When a reasonable hypothesis on the morphology can be formulated, we fit the data using the form factor of a sphere combined with a structure factor, whenever the data suggest us so. In general, a uniform sphere form factor quickly reveals to be useless for all the systems tested in this work, as it cannot fit the data. On the basis of previous studies using small angle scattering techniques to study sophorolipids dispersions in water,<sup>11,52</sup> we successfully employ a core-shell sphere model (Figure 3) to build the form factor. Considering the charged nature of the class of sophorolipids used in this work, and in agreement with previous data in the literature,<sup>11,53</sup> we use a Hayter-MSA structure factor,  $S(q)$ ,<sup>54,55</sup> which describes a repulsive screened Coulombic intermicellar interaction potential.



**Figure 3 – Core-shell sphere model at the basis of the form factor used to fit the SAXS data in this work.**

Both the Guinier analysis and fitting have been performed with the 3.1.2 version of the SasView software, available free of charge on the developer’s website. For the Guinier analysis, the  $R_g$  is the only adjustable parameter. The core-shell spherical model in combination with a screened repulsive potential has a large number of variables (>10), defined in Table 1. To minimize the risk of error, we fix as many variables as possible (details in Table 1), in relationship to our systems (e.g., salt concentration, volume fraction, etc...): by doing so, we reduce the number of variables to four ( $R_{core}$ ,  $T$ ,  $\rho_{shell}$ ,  $Q$ ). The model is used to fit the data of compound **1f** (this choice is arbitrary), for which the model is refined and best fit parameters are retained. We iterate the results to fit the data corresponding to those compounds, which have the closest chemical structure: the fixed terms are consequently adapted while the fit only concerns the single parameter reflecting the structural change in the molecule. Finally, the other parameters are fitted for refinement. *Example:* the parameters obtained from the fit of **1f** (chain length: C18) are used to fit compound

**1h** (chain length: C15). Since this compound has a shorter chain length,  $R_{core}$  is the only variable employed to fit the data corresponding to **1h**, while  $T$ ,  $\rho_{shell}$  and  $Q$  are only fitted after  $R_{core}$  for refinement. When the chain length between two compounds is constant,  $T$  and  $\rho_{shell}$  are fitted before  $R_{core}$  and  $Q$ . The use of the present model can be performed only on those systems for which the exact salt concentration is known, which excludes all pH-dependent experiments performed in this work and which are treated using the Guinier analysis. A small polydispersity (0.1) has been employed for the fits. Uncertainty in the fitted parameters  $R_{core}$  and  $T$  is estimated to about 15 %.

**Table 1 – List of the fitting parameters used with the core-shell sphere form factor  $P(q)$  (Figure 3) combined with the screened Coulomb potential structure factor  $S(q)$ .**

Degree of freedom	$P(q)$ * $S(q)$	Parameter	Unit	Value	Description
Fixed variable	$P(q)$	$\rho_{core}$	nm <sup>-2</sup>	8.0 – 8.7*10 <sup>-4</sup>	Scattering Length Density of the core <sup>§</sup>
		$\rho_{solvent}$	nm <sup>-2</sup>	9.4*10 <sup>-4</sup>	Scattering Length Density of the solvent
	$S(q)$	$\epsilon$	none	78.3	Dielectric constant of water at 25°C <sup>56</sup>
		[S]	M	0.10 ± 0.02	Salt concentration. In this work, it refers to the concentration of the counterion, I. Counterion is assumed to be monovalent in the fit
		$T$	K	300	Temperature
		$\varphi$	none	0.09-0.10	Volume fraction <sup>°</sup>
		$R_{HS}$	nm	$(R_{core} + T)$	Hard-sphere radius
Fitted variable	$P(q)$	Background	cm <sup>-1</sup>		
		Scale factor	none		
		$R_{core}$	nm		Radius of the core
		$T$	nm		Thickness of the shell
		$\rho_{shell}$	nm <sup>-2</sup>		Scattering Length Density of the shell



	$S(q)$	$Q$	$e$		Charge of the micelle, scaled to charge of the electron
--	--------	-----	-----	--	---

\* The exact value depends on the  $M_W$  and concentration of the compound studied; § Considering the complexity of the compounds explored in this work, we allow  $\rho_{core}$  to vary in a narrow range, corresponding to a reasonable set of values expected for an aliphatic chain; ° The exact value depends on the concentration of a given compound

*Aggregation number.* The structural parameters that characterize the micellar aggregates and that are obtained from the modelling of the SAXS data are used to estimate the aggregation number through the expression  $N_{agg} = \frac{V_{mic}}{V_{molec}}$ , where  $V_{mic}$  is the micellar volume and  $V_{molec}$  is the molecular volume of a single lipid. In conventional surfactant systems, including glycolipids, where by *conventional* we mean a hydrophilic headgroup connected to a lipophilic chain,  $V_{molec}$  is generally associated to the volume of the alkyl chain, while  $V_{mic}$  represents the volume of the hydrophobic core.<sup>57</sup> The sophorolipids derivatives studied in this work cannot be considered as *conventional*, because they are constituted by two hydrophilic headgroups, different in nature and charge (bulky neutral sophorose, small charged ammonium), connected by a  $C_x$  ( $x= 8, 11$ ) spacer. Given such complexity, and considering that long ( $> 10$ -12 carbon atoms) spacers were shown to integrate the core of gemini surfactant micelles,<sup>37</sup> we prefer not to associate the hydrophobic region to a specific part of the molecule. Under these circumstances,  $V_{mic} = \frac{4}{3}\pi(R_{core} + T)^3$ , which corresponds to the volume of the entire core-shell sphere (hydrophobic core and hydrophilic regions) used to model the micelles;  $V_{molec}$  will then correspond to the molecular volume of the entire molecular entity, including the hydration shell and counterion. Considering again the complexity of the molecular structures investigated here, we cannot use a simple expression to obtain the molecular volume (e.g, the standard Tanford formula),<sup>58</sup> but we have to deconstruct the molecule into smaller parts and estimate their individual molecular volume, as proposed by Luzzati.<sup>59</sup>

$V_{molec}$  for each submolecular group used in this work are given in Table 2, while the number of each submolecular group for selected compounds is detailed in Table S 1. For the hydration of sophorose and the counterion, we make the assumption that the micelle-water palisade is uniform.<sup>59</sup> We also make the following approximations: 1) reported values for the hydration number,  $N_H$ , of I<sup>-</sup> in water varies between 6<sup>60,61</sup> and 9<sup>61</sup>, although 6 seems the most suitable one and for this reason we employ  $N_H(I^-)= 6$ . If  $N_H(I^-)= 9$  is used instead,  $N_{agg}$  decreases by 2 units.

2)  $N_H$  is not known exactly for sophorose, although values for glucose and disaccharides are reported. Unfortunately,  $N_H$  values for sugars in general vary quite a lot and for glucose  $N_H$  is given to vary between 10 and 30.<sup>62</sup> We have already reported a detailed discussion on which  $N_H$  would best suit to sophorose<sup>52</sup> and we concluded that values between 16 and 22 in the presence of salt are not outrageous. For this work we use  $N_H=20$ , although one should consider an error of at least  $\pm 5$  on the final aggregation number. 3) The volumes estimated for the aliphatic region (CH, CH<sub>2</sub>, CH<sub>3</sub>) are based on molten paraffin,<sup>59</sup> which was found to be a reasonable approximation,<sup>63</sup> nonetheless, slight discrepancies can be found in the literature. For instance, values as low as 0.0021 nm<sup>3</sup> for CH<sub>2</sub> are reported for glycolipids.<sup>63</sup> The molecules studied in this work have a large number of CH<sub>2</sub> groups and small variations on the volume of CH<sub>2</sub> can induce non-negligible variations of  $N_{agg}$ . For instance, if  $V_{molec}(\text{CH}_2)=0.021\text{ nm}^3$ , one should expect  $N_{agg}$  to increase by about 7 units. 4) We neglect the contribution of nitrogen to the entire  $V_{molec}$  for the nitrogen center, since all compounds studied here only have one or two nitrogen atoms. This approximation is reasonable and will not affect  $N_{agg}$ ; for instance,  $V_{molec}$  for N(CH<sub>3</sub>)<sub>3</sub> in alkylammonium salts is in the order of 0.6 nm<sup>3</sup>.<sup>59</sup> 5) It is known that the hydration number of CH<sub>2</sub> and CH<sub>3</sub> are respectively 0.999 and 0.586<sup>64</sup> and up to 3 CH<sub>2</sub> groups from the polar head can be hydrated.<sup>65-67</sup> In this work, considering the uncertain boundaries of the headgroup and actual hydration numbers of the C<sub>x</sub>(x=8, 11) spacer, we prefer not to consider the hydration of CH<sub>2</sub> groups. However, in the hypothesis that CH<sub>3</sub> groups are hydrated and at least two CH<sub>2</sub> groups are also hydrated on each side of the nitrogen atom and next to the glycosidic bond, one expects  $N_{agg}$  to decrease by at least 7 units.

The uncertainty on the estimation of  $V_{molec}$  as well as the consequent impact on the value of  $N_{agg}$  are reported in Table 2. All in all, the values of  $N_{agg}$  provided throughout this work are based on the  $V_{molec}$  data in Table 2, while the maxima and minima are respectively calculated by adding all positive and negative uncertainties given in Table 2.

Two final remarks: 1) x-rays are generally poorly sensitive to water and counterions in the hydrophilic shell; if the hydration shell of sophorose and counterion are not included in the calculation, the values of  $N_{agg}$  are about 1.6 times higher than the values reported here. In this case,  $N_{agg}$  should be considered as an upper limit; 2) if only the hydrophobic core of the micelle and the aliphatic chain of the lipids are used to evaluate  $N_{agg}$  (as commented above),  $N_{agg}$  is

contained between 10 and 30. These values are too low to be realistic, thus making our initial assumption of using the Luzzati approach more plausible.

**Table 2 – Values of the molecular volume,  $V_{molec}$ , used to calculate  $V_{molec}$  for the molecules given in Figure 2.**

$V_H$ = hydration volume,  $N_H$ = hydration number

Moiety	$V_{molec}/\text{nm}^3$	Limits	Uncertainty on $N_{agg}$	Ref
Sophorose	0.384 <sup>a</sup>			63,68
CH <sub>2</sub>	0.027 <sup>b</sup>	Lowest: 0.021 nm <sup>3</sup>	+7	59
CH <sub>3</sub>	0.054 <sup>b</sup>			
I <sup>-</sup>	0.037 <sup>c</sup>			69
H <sub>2</sub> O	0.030 <sup>d</sup>			
$V_H$ (sophorose)	0.598 <sup>e</sup>	16 < $N_H$ < 22 0.48 < $V_H/\text{nm}^3$ < 0.66	±5	52
$V_H$ (I <sup>-</sup> )	0.179 <sup>f</sup>	Highest $N_H$ : 9 Highest $V_H$ : 0.269 nm <sup>3</sup>	-2	60
$V_H$ (CH <sub>2</sub> , CH <sub>3</sub> )	0.214 <sup>g</sup>		-7 <sup>h</sup>	64-67

a) Taken as  $2 \cdot V_{molec}(\text{glucose}) = 0.192 \text{ nm}^3$ . As a comparison, the  $V_{molec}$  of maltose, a common disaccharide, was estimated to  $0.382 \text{ nm}^3$ ;

b) Known from molten paraffin and successfully employed to estimate  $V_{molec}$  of alkylammonium surfactants in a micellar environment;

c) Obtained from the radius of I<sup>-</sup> anion;

d) Calculated on the basis of the molar volume of  $18 \text{ cm}^3/\text{mol}$ ;

e)  $V_H(\text{sophorose}) = V_{molec}(\text{H}_2\text{O}) \cdot N_H(\text{sophorose})$ ;  $N_H(\text{sophorose}) = 20$ ;

f)  $V_H(\text{I}^-) = V_{molec}(\text{H}_2\text{O}) \cdot N_H(\text{I}^-)$ ;  $N_H(\text{I}^-) = 6$

g) For samples **1** and **2**,  $V_H(\text{CH}_2, \text{CH}_3)$  is calculated as follows:  $V_H(\text{CH}_2, \text{CH}_3) = V_{molec}(\text{H}_2\text{O}) \cdot N_H(\text{CH}_2) \cdot 0.999 + V_{molec}(\text{H}_2\text{O}) \cdot N_H(\text{CH}_3) \cdot 0.586$ ;  $N_H(\text{CH}_2) = 6$ ,  $N_H(\text{CH}_3) = 2$ . For samples **3**, **5** and **6**, the number of hydrated CH<sub>2</sub> and CH<sub>3</sub> could be estimated to be in the range between 10 and 12 altogether.

h) This value is estimated for samples **1** and **2**. Estimation for samples **3**, **5** and **6** provides an uncertainty on  $N_{agg} \sim -3$ .

## Results and Discussion

*Quaternary ammonium derivatives of oleic acid SL.*

In this work, we have studied a series of quaternary ammonium sophorolipid derivatives bearing a C8 spacer between the ammonium and sophorose headgroups. They are bolaform compounds bearing a positive charge on the nitrogen, which has various degrees of substitution (Figure 4), with increasing headgroup size: trimethyl (**1a**), tributyl (**1b**), methyl-dibutyl (**1c**), methyl-butyl-benzyl (**1d**), methyl-dibenzyl (**1e**), dimethyl-octadecyl (**1f**) and methyl-butyl-octadecyl (**1g**). The progression from **1a** to **1e** accounts for a small (trimethyl) to a much bulkier (methyl-dibenzyl) headgroup, while **1f** and **1g** mean to show the influence of a long aliphatic chain for two headgroup substitutions (dimethyl, **1f**, and methyl-butyl, **1g**).

The SAXS data presented in Figure 4 show that samples from **1a** through **1e** (C= 100 mg/mL) have a similar profile characterized by a low- $q$  scattering and a mostly flat, poorly-defined, signal above  $\sim 0.4 \text{ nm}^{-1}$ . The same trend is confirmed for a broader range of concentrations, varying from 0.78 mg/mL to 100 mg/mL (Figure S 2), although in most cases no signal is detected below 6 mg/mL. One can interpret the signal as the coexistence of small objects having a radius of gyration,  $R_g$ , varying between 0.3 nm and 1.1 nm (Table 1), and assembled structures larger than 120 nm. In the hypothesis of spherical objects, the radius is about  $1.3 * R_g$  ( $R = \sqrt{\frac{5}{3}} R_g$ ). For compounds **1a** through **1c**,  $R$  is centered around 0.4 nm, while compounds **1d** and **1e** have larger radii, ranging between 0.7 nm and 1.4 nm (Table 3a). According to the Tanford formula,<sup>58</sup> defining the length,  $L$ , of an aliphatic chain ( $L = 1.54 + 1.265 * n$ ,  $n$  being the number of  $\text{CH}_2$  groups), one expects the size of **1a** through **1e** to be at least 1.1 nm (distance from nitrogen to CH). One should add 1 nm to account for sophorose and at least 0.5 nm to consider the size of the largest headgroups (butyl, benzyl). The overall size of **1a** through **1e** should then be expected to be in the order of 2 nm, if one assumes an elongated configuration of sophorose with respect to the aliphatic chain. Bolaform surfactants are known to form smaller micelles than classical surfactants,<sup>41,70</sup> and this is commonly explained by the fact that bolaform surfactants interpenetrate and bend to keep the hydrophilic headgroups towards the micelle outer boundary. We have ourselves described the possible orientations of acidic C18:1 sophorolipids within ellipsoidal micelles in a previous work.<sup>52</sup> Nonetheless, even under these assumptions, the micellar radius  $R$  estimated for compounds **1a** through **1c** roughly corresponds to  $\frac{1}{4}$  of the length of the molecule, which either means that these compounds do not actually form micelles or, if they do, the lipid is in an uncommon configuration. The first hypothesis is probably more reasonable as the calculated molecular volume of **1a** (data in

Table 2 are used) is actually larger than the volume of a sphere having a radius of 0.4 nm. On the contrary, compounds **1d** and **1e** have a value of the radius, which is compatible with an interpenetrated and/or U-shaped molecular configuration in the micelle. A qualitative consideration on the basis of molecular volumes compared to the volume of a sphere having a radius between 1 nm and 1.4 nm suggests aggregation numbers between 5 and 10, which are compatible with previous works on small quaternary ammonium bolaamphiphiles.<sup>70</sup>

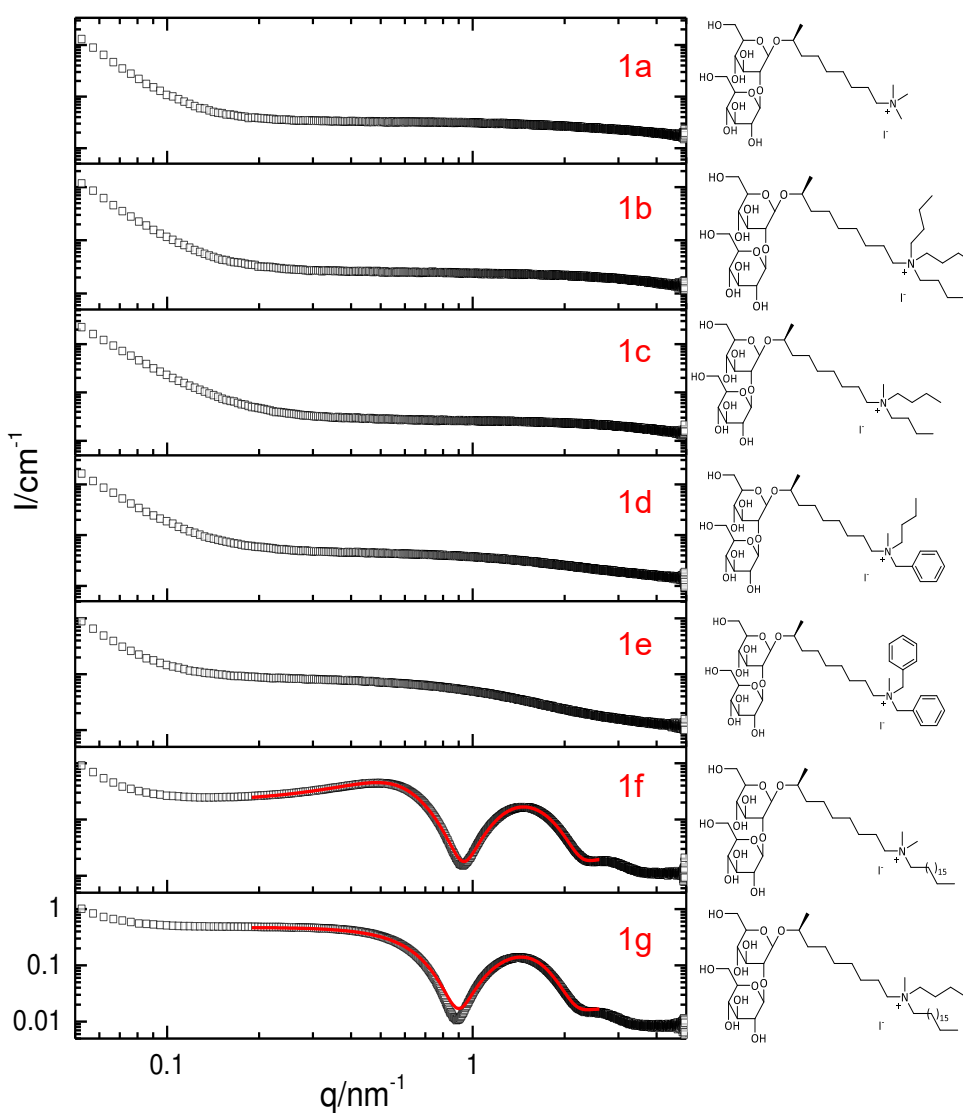


Figure 4 – SAXS profiles of bolaform asymmetric (**1a** through **1e**) and divalent (**1f**, **1g**) sophorolipid quaternary

**ammonium derivatives (C= 100 mg/mL). The fits for **1f** and **1g** have been obtained using the core-shell sphere form factor combined with a screened Coulomb potential structure factor.**

The bulkiness of the headgroup does not seem to play a significant role on the overall morphology, unless mono- (**1d**) and dibenzyl (**1e**) groups are introduced. The micellar size of these compounds increases by a factor two or more with respect to the methyl and butyl system (**1a-1c**). The benzyl group could then either promote molecular packing through  $\pi$ - $\pi$  stacking, or it could be solubilized at the interior of the micellar aggregate, thus increasing its size. Solubilization of an hydrophobic group on the polar head of surfactants has been discussed before for alkylammonium compounds with bulky headgroups,<sup>45</sup> and in that case, aryl groups (e.g., trimethylbenzene) increased the size of quaternary ammonium surfactants micelles.<sup>71</sup> Finally, the origin of the low- $q$  scattering signal is quite unclear: it could either be associated to minor amounts of an unidentified type of aggregation, as it was shown for the nanoplatelets/micelles coexistence in the case of acidic oleic-acid sophorolipids at high pH,<sup>72</sup> or to a superstructure composed of the individual objects described above. Further work will be needed to explore these hypotheses.

Very differently from the samples treated above, **1f** and **1g** have a well-defined signal characterized by a broad hump centered around  $0.5 \text{ nm}^{-1}$  and two oscillations above  $0.8 \text{ nm}^{-1}$  in a broad concentration range (Figure S 2), and particularly highlighted at C= 100 mg/mL in Figure 4. A similar signal is common for micelles with repulsive interactions. In this case, a model-dependent analysis could be employed to determine the micelles structural characteristics. For the form factor, the simplest assumption is to suppose a homogeneous sphere, while for the structure factor, the positive charge of the lipids suggests the presence of a screened Coulomb interaction potential. However, any attempt to fit the data, especially the broad oscillation above  $0.5 \text{ nm}^{-1}$ , using a homogeneous spherical form factor does not lead to satisfactory results. In fact, the chemical structure of **1f** and **1g** suggests an inhomogeneous electron density distribution along the molecule, and consequently, within a micelle. For this reason, we have employed a core-shell spherical (Figure 3) form factor, which has been used before to study the micellar structure of acidic sophorolipids.<sup>11,52,53</sup> The core-shell model accounts for sophorose, the more hydrophilic part of sophorolipids, and for the aliphatic chain, the hydrophobic region. In the case of **1f** and **1g**, such a distinction is less marked, because the C8 spacer between the ammonium and sophorose makes the identification of the hydrophilic shell less clear. The typical fit is superimposed to the

corresponding SAXS curves in Figure 4 ( $C= 100$  mg/mL) and the structural values are provided in Table 3b.

Fit of the concentration-dependent profiles for **1f** are shown in Figure S 2. One should note that a spherical model with polydispersity seems enough to model the SAXS profiles of **1f** and **1g**, while this was not the case for standard oleic acid acidic sophorolipids, which require a core-shell ellipsoid model instead.<sup>52</sup> The best fit values show that both **1f** and **1g** have a comparable  $R_{core}$  (respectively 1.7 nm and 1.6 nm) but a different shell thickness and charge. **1g** has a larger hydrophilic shell ( $T= 2.2$  nm) and a charge,  $Q$ , close to zero. The aggregation number is slightly larger for **1g** than **1f** and both are in the order of 100, a value which is compatible both with glycosylated surfactants with long tails<sup>57</sup> and in the order of magnitude of alkylammonium surfactants having a comparable chain length.<sup>73</sup> Finally, analysis of the fits for the concentration-dependent data for **1f** (Figure S 2) shows an expected decrease in the aggregation number, which varies from 93 at 100 mg/mL to 75 at 6.25 mg/mL. Interestingly, the entire dataset can be fitted using the same spherical shape, with no need to use anisotropic morphologies (e.g., ellipsoid of revolution).

**Table 3 – Structural parameter values extracted from the SAXS profiles in Figure 4. a) Radius of gyration,  $R_g$ , obtained through the Guinier approximation and equivalent sphere radius,  $R$ , calculated for samples 1a through 1e. b) Values obtained for 1f and 1g using the core-shell sphere form factor combined with a screened Coulomb potential structure factor. All parameters are explained in Table 1 and Table 2**

Sample	$R_g/\text{nm}$	$R/\text{nm}$
1a	0.4	0.5
1b	0.3	0.4
1c	0.3	0.4
1d	0.7	0.9
1e	1.1	1.4

a)

Sample	$R_{core}/\text{nm}$	$T/\text{nm}$	$Q/e$	$\rho/10^{-4} \text{ nm}^{-2}$		$[I]/M$	$\varphi$	$V_m/\text{nm}^3$	$N_{agg}$	$N_{agg} (\text{max, min})$
				$\rho_{shell}$	$\rho_{core}$					
1f	1.7	1.9	6.0	9.94	8.30	0.11	0.09	2.14	92	104,78
1g	1.6	2.2	0.4	9.80	8.30	0.11	0.10	2.22	104	116,90

b)

The experiments above demonstrate that:

1) when the modified quaternary ammonium sophorolipids stay in the bolaform configuration with a short (C8) spacer,  $R_g$  is smaller than about 1 nm, thus indicating the presence of small micelles with aggregation numbers below 10. The literature on bolaform surfactants is quite consistent on the fact that short-chain bolaamphiphiles are water-soluble at room temperature and form micelles, the diameter of which is equal to the molecular length, which involves molecular bending and intramicellar crossing.<sup>41,42,74</sup> Interestingly, the self-assembly process is even doubtful for some specific compounds (**1a-1c**,  $R_g < 0.5$ ). Introduction of a long (C18) aliphatic tail on the nitrogen atom transforms the bolaform into a divalent surfactant (two headgroups separated by a spacer with asymmetric chain lengths)<sup>39</sup> and it drives the formation of well-defined micelles through a standard hydrophobic interaction.<sup>33,34,43,58</sup> Critical micelle concentration (cmc) values are contained between 0.5 mg/mL and about 1 mg/mL for samples **1e** and **1g** (Table S 1), and their micellar self assembly is consistent with these values. On the contrary, although we do not dispose of the cmc values for samples **1a-1e**, their concentration-dependent small angle scattering (Figure S 2) shows a mild, if no, evolution of the scattering profile, indicating either high ( $> \sim 10$  mg/mL) cmc values, or even no cmc at all. Whatever the case, this behavior has little practical interest for applications.

2) The substitution of chemical groups on the nitrogen atom plays a minor role in the bolaform systems (**1a** through **1c**), except for a slight increase in  $R_g$ , probably due to the solubilisation of the arene moiety (**1d**, **1e**). For the divalent derivatives, substitution of a methyl (**1f**) with a more hydrophobic butyl (**1g**) group screens the intermicellar repulsive interactions, which is sign of a reduced charge density on the micelles. This effect has been observed both for linear<sup>45</sup> and gemini<sup>75</sup> cationic surfactants with bulky headgroups (tripropyl, tributyl). However, despite some analogies with divalent surfactant systems, one should not overinterpret the analogies between the lipids studied here and more classical cationic surfactants studied in the literature. The complexity of the headgroup in **1f** and **1g** is much harder to rationalize with respect to the broad amount of existing work.<sup>36,37,39,45,76</sup> For instance, the micellar charge for **1f** is only 6e, which is between two and three times smaller than the charge of cationic surfactant micelles:<sup>76</sup> the bulky sophorose-C8-ammonium headgroup has certainly an important screening effect of the charge.

3) The thickness of the hydrophilic region in **1f** and **1g** is of the order of 2 nm, a value which can only be explained if the length of hydrophilic sophorose adds to the extended size of the

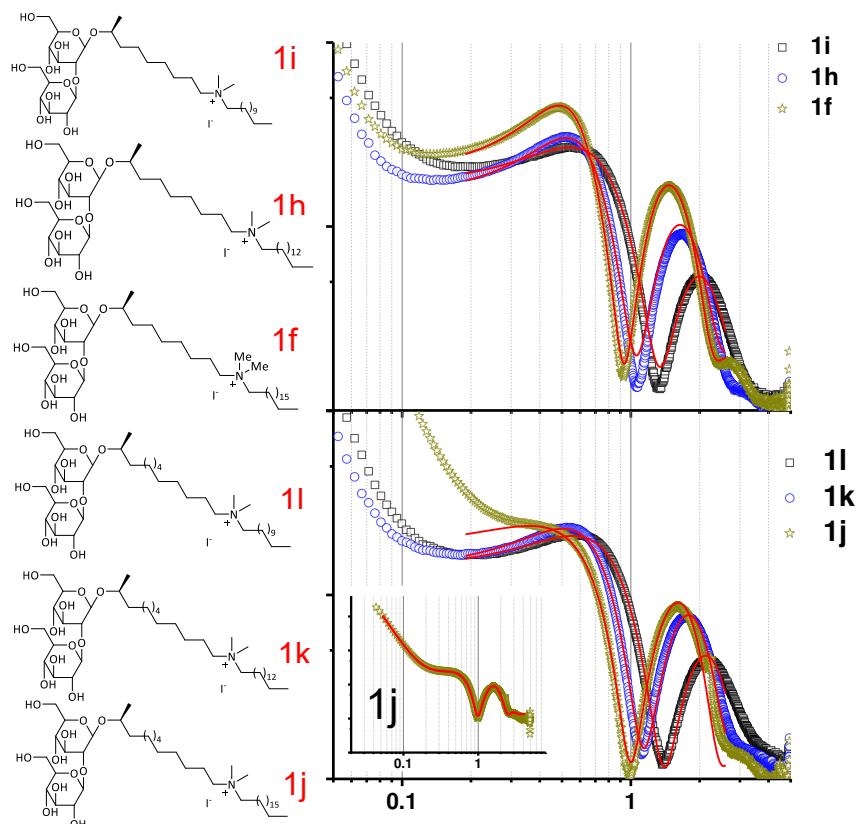


C8 spacer. In this sense, our hypothesis to use a core-shell containing a single shell composed of sophorose+C8 does not seem to be outrageous.

These same conclusions can be applied to a set of amine oxide derivatives (**2a-2g**) of C18:1 sophorolipids, the data of which are shown in Figure S 3 at pH 5 and pH 11 and in a broader pH range in Figure S 4. This set of compounds has the exact features of the quaternary ammonium derivatives described above ( $R_g$  corresponding to pH 5 are given in Table S 2), except for the presence of an amine oxide headgroup instead of the quaternary ammonium. When the compound is in the bolaform configuration (C8 spacer), either no or small aggregates are formed. In the presence of a C18 alkyl chain, the typical signal of micelles undergoing repulsive interaction is detected. One can reasonably conclude that the nature of the substitution on the headgroup leads the self-assembly of this class of sophorolipid derivative.

*Quaternary ammonium derivatives of SL: effect of spacer and chain length.*

In the previous section we have shown that the octadecyl substitution on the ammonium group changes the structure of the lipid from bolaform to divalent and it drives the formation of micelles. In the following, we show the influence of the length of the alkyl chain (C18, C15, C12, respectively in samples **1f**, **1h**, **1i** and in samples **1j**, **1k**, **1l**) but also of the length of the spacer (C8, C11) between sophorose and the ammonium headgroup for the same alkyl chain variation (C8 in samples **1f**, **1h**, **1i** and C11 in samples **1j**, **1k**, **1l**). The SAXS data are reported in Figure 5 for  $C=100$  mg/mL and in Figure S 5 for lower concentrations. At a first glance, all samples display the typical scattering profile of micelles with repulsive interactions, similarly to **1f**, and for this reason the same core-shell sphere model with screened Coulomb repulsive potential has been used to fit the SAXS curves (fits are shown in Figure 5). The fit has been contained above  $q=0.2$  nm<sup>-1</sup>, so to leave out the low- $q$  scattering region, indicating attractive interactions. Nonetheless, it can be easily shown that the entire curve can be fitted if an additional power law term is added to the general scattering equation  $I(q)$ . This is illustrated for sample **1j** in the inset of Figure 5 but will not be employed any further, due to the limited number of points in the low- $q$  region.



**Figure 5 – SAXS profiles of divalent sophorolipid quaternary ammonium derivatives ( $C=100$  mg/mL). The fits have been obtained using the core-shell sphere form factor combined with a screened Coulomb potential structure factor.**

For both the C8 (**1f**, **1h**, **1i**) and C11 (**1j**, **1k**, **1l**) systems, Figure 5 shows the systematic shift of the SAXS profiles towards the high- $q$  region when the size of the alkyl chain decreases. High- $q$  shift is expected with decreasing micellar size, a fact that is confirmed by the modelling data in Table 4.  $R_{core}$  decreases from 1.7 nm to 1.0 nm in the C8-spacer series and from 1.6 nm to 1.0 nm in the C11-spacer series, confirming that the alkyl chain drives the self-assembly and the size of the chain is directly connected to the micellar core size. Interestingly, the shell thickness is practically constant in the 1.6 nm – 1.9 nm range throughout all samples, despite the difference in the spacer length. This fact suggests that the sophorose-alkyl spacer does not assume an elongated configuration, which would induce a thicker shell for the C11 samples (**1j**, **1k**, **1l**), but it rather fills the shell volume homogeneously, with the alkyl spacer assuming a flexible configuration. If similar data on analogous divalent surfactants lack in the literature, we must outline that this behavior is

rather different than what has been reported in dimeric surfactants, where long spacers (above 10 carbon atoms) become part of the micellar core.<sup>37</sup> This hypothesis seems to be confirmed by the constant scattering length density (SLD) of the shell,  $\rho_{\text{shell}}$ , which is practically constant at around  $9.9 \cdot 10^{-4} \text{ nm}^{-2}$  for all samples. This value is consistent with an average SAXS SLD expected from hydrated sophorose ( $> 10 \cdot 10^{-4} \text{ nm}^{-2}$ )<sup>52</sup> and an aliphatic chain (aliphatic chains with densities in the range of  $0.8 \text{ g/cm}^3$  are expected to have  $\text{SLD} < 9 \cdot 10^{-4} \text{ nm}^{-2}$ ).

If the charge,  $Q$ , is particularly affected by the change in chain length and spacer, deeper considerations should be done.  $Q$  seems to be overall higher for the C8 than for the C11 series, however, the calculated micellar surface charge density ( $\frac{Q}{S} \equiv \frac{Q}{4\pi(R_{\text{core}}+T)^2}$ ) is very close for both systems, as it lies between 0.02 and 0.03  $\text{e/nm}^2$ .  $Q$  has a well-defined trend in the C8 samples, increasing with increasing length of the alkyl chain; the corresponding charge density varies as follows: 0.020, 0.031 and 0.037  $\text{e/nm}^2$  for, respectively, C12, C15 and C18 samples (**1i** through **1f**). Interestingly, the same trend has been described for alkyltrimethylammonium surfactants by Zana.<sup>76</sup> For the C11 samples, if the trend is respected for the C12 (**1l**) and C15 (**1k**) samples (respectively 0.021 and 0.022  $\text{charges/nm}^2$ ), the charge is strangely low for C18 (**1j**). To explain this failure, one should look at the **1j** SAXS profile in Figure 5: **1j** has the strongest low- $q$  scattering signal, which masks the broad hump due to the structure factor, and the charge is undoubtedly underestimated in the model that we have employed. Modelling the entire profile (inset in Figure 5) unfortunately does not produce different parameters, including the charge. Finally, despite the uncertainty (please, refer to the comments on the calculation of  $N_{\text{agg}}$  in the materials and method section),  $N_{\text{agg}}$  increases with increasing alkyl chain length for both C8 and C11 spacers, as expected for both alkylammonium<sup>76</sup> (decreasing of surface area per headgroup with number of carbons in the chain) and glycosylated<sup>57</sup> surfactants. In light of this, the surface per surfactant molecule, estimated as  $S/N_{\text{agg}}$ , decreases with increasing the chain length, as expected, and it ranges at around  $2 \text{ nm}^2/\text{molecule}$ , a value which is consistent with the bulky dimension of the sophorose- $\text{C}_x$  ( $x=8, 11$ )-ammonium moiety. Interestingly, the surface area per molecule is slightly larger for C11 (2.36 - 2.01  $\text{nm}^2/\text{molecule}$ ) than for C8 samples (2.10 - 1.81  $\text{nm}^2/\text{molecule}$ ). This shows that, although the thickness of the hydrophilic region is comparable for the short and long spacers, the C11 molecules occupy a larger surface area, as one would expect.

**Table 4 - Structural parameter values extracted from the SAXS profiles in Figure 5. Values are obtained using the core-shell sphere form factor combined with a screened Coulomb potential structure factor. Sp= length of the alkyl spacer between sophorose and nitrogen; C= length of the alkyl chain on nitrogen. All other parameters are explained in Table 1 and Table 2**

Sp	C	Sample	$R_{core}/nm$	$T/nm$	$Q/e$	$\rho/10^{-4} nm^{-2}$		$[I]/M$	$\phi$	$V_m/nm^3$	$N_{agg}$	$N_{agg} (max,min)$
						$\rho_{shell}$	$\rho_{core}$					
C8	C18	1f	1.7	1.9	6.0	9.94	8.30	0.11	0.09	2.14	92	104,78
	C15	1h	1.5	1.8	4.2	9.92	8.28	0.12	0.10	2.06	73	85,59
	C12	1i	1.0	1.8	2.0	9.93	8.00	0.12	0.10	1.98	48	60,34
C11	C18	1j	1.6	1.7	1.5*	9.87	8.44	0.12	0.10	2.22	70	82,56
	C15	1k	1.4	1.7	2.7	9.94	8.25	0.12	0.10	2.14	55	67,41
	C12	1l	1.0	1.6	1.8	9.96	8.00	0.12	0.10	2.06	37	49,23

\* This value is underestimated due to the strong scattering at low q-values of the corresponding SAXS profile.

The strong scattering masks part of the structure factor between  $0.1 nm^{-1}$  and  $0.8 nm^{-1}$

*Quaternary ammonium derivatives of SL: symmetrical bolaamphiphiles with three hydrophilic centers (sophorose-nitrogen-sophorose).*

The structural complexity of the sophorolipids derivatives has been brought one-step further by replacing one of the substituent on the nitrogen atom by an alkyl sophorose group. The resulting set of compounds is composed of a symmetrical bolaform sophorolipid containing a nitrogen atom in the middle, either in the form of tertiary amine (**3a**, **3b**, **3c**, **3d**) or in the form of ammonium (**4a**, **4b**, **4c**, **4d**). In the amines, methyl (**3a**), butyl (**3b**), benzyl (**3c**) and octadecyl (**3d**) are introduced as side groups, which become dimethyl (**4a**), methyl butyl (**4b**), methyl benzyl (**4c**) and methyl octadecyl (**4d**) in the ammonium derivatives. We have evaluated the self-assembly properties of this set of compounds as a function of pH (from pH 3 to pH 11 for **3a**, **3b**, **3c**, **3d** at C= 10 mg/mL) and concentration (**4a**, **4b**, **4c**, **4d**). The entire set of pH and concentration data are shown in Figure S 6, while a representative set of experiments at pH 3 and pH 11 for **3a**, **3b**, **3c**, **3d** (C= 10 mg/mL) and C= 12.5 mg/mL for **4a**, **4b**, **4c**, **4d** is shown in Figure 6.

Tertiary amines generally have a pKa around 10 and for this reason, Figure 6 shows the SAXS data collected at a pH lower and higher than the supposed pKa, where one expects the amine to be respectively protonated and neutral. For the methyl derivative (**3a**), the SAXS profile is pH-independent and it closely resembles to the asymmetric bolaform SL shown in Figure 4, with a  $R_g$  in the order of 1 nm (Table 5a), a size which is not compatible with a potential aggregation of this compound in water, at least not in a micellar form. The low- $q$  scattering also suggests the presence

of larger aggregates. Similar results are obtained for charged **3b** (pH 3) sample and for the corresponding quaternary ammonium derivatives (**4a**, **4b**). In particular, the scattering profile of **4a** and **4b** is quite atypical and it strongly suggests the coexistence of two types of self-assembled objects of very distinct size. Any attempt to use simple numerical models (sphere, cylinder, flexible cylinders, core-shell, etc...) that could fit the data was useless, although we could perform a satisfactory fit using a combination of small cylinders and flexible “infinitely” long cylinders, as shown in Figure S 7. Being aware of the arbitrary fitting strategy in this case, we conclude that SAXS alone cannot dissipate the doubts about the actual concomitant structures of samples **4a**, **4b**, but it shows the complexity of the self-assembly of this set of molecules.

All in all, samples **3a**, **3b**, **4a**, **4b** confirm that this class of bolaform amphiphiles with small substituents on the nitrogen atom are not good amphiphiles; exception is made for the butylamine (**3b**, pH 11) derivative, for which the qualitative size of the aggregate shows  $R_g \sim 2$  nm, confirmed by fitting the data using a core-shell sphere model (numerical data in Table S 3). In this case, the number of molecules in the aggregate can be estimated to about 10, which is compatible with a micelle composed of interpenetrated molecules, as discussed above. Similar conclusions are drawn from the benzylamine (**3c**, pH 11) sample: in its neutral form this compound forms small aggregates composed of less than 20 molecules (Table S 3), in the hypothesis that a core-shell model best suits the description of the corresponding SAXS data (Figure 6).

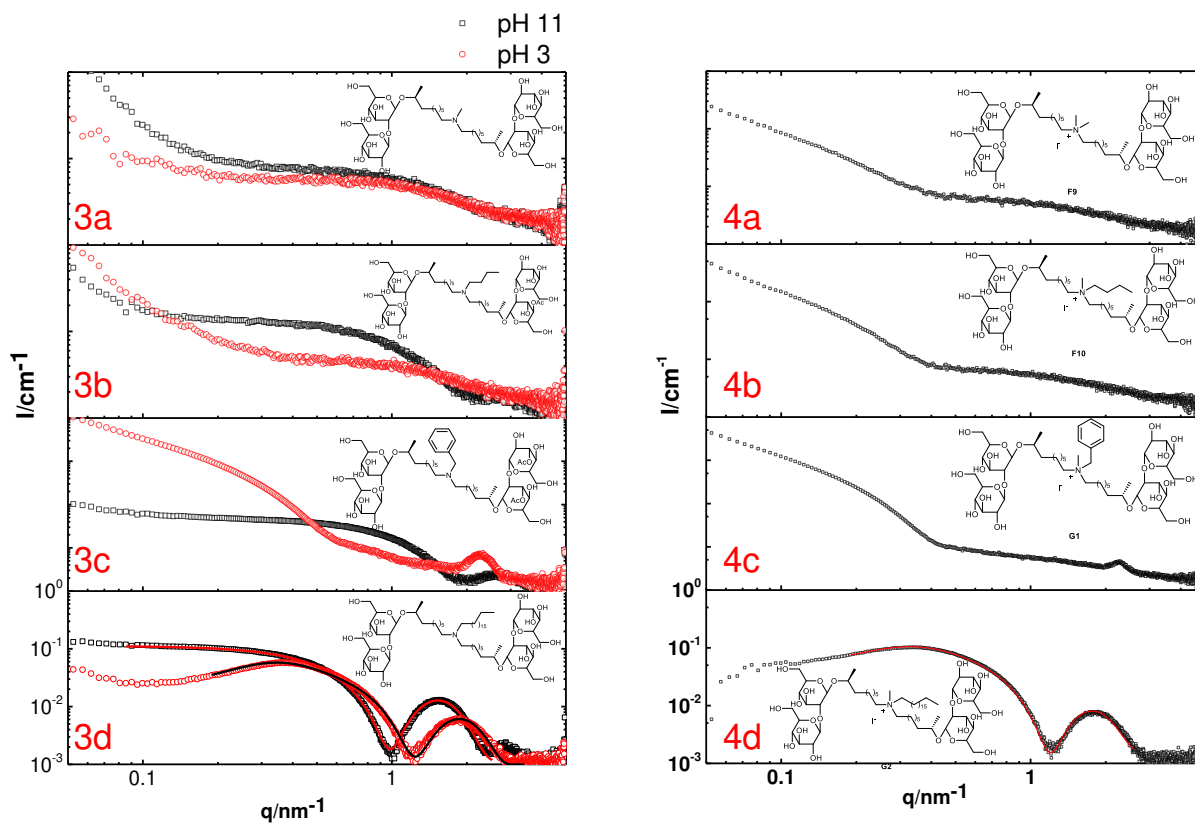
The charged benzyl derivative, both in its benzyl ammonium (**3c**, pH 3) and methyl benzyl ammonium (**4c**) forms provides a SAXS profile, which is not compatible with a micellar system, but rather with a more complex morphology.  $I(q)$  varies with  $q$  following a power law with an exponent contained between -1.6 and -1.8 (Table 5a) and a peak centered at  $2.3 \text{ nm}^{-1}$ . The value of the exponent in the log-log representation of  $I(q)$  is commonly associated to specific, although simple, morphologies (-1: cylinders; -2: flat surface; -4: sharp interface).<sup>51</sup> For exponents having non-integer values, the interpretation is less straightforward, as they can be associated to fractals,<sup>77</sup> semiflexible system including wormlike micelles,<sup>38,78</sup> or to polydispersity. The peak in the region of  $2 \text{ nm}^{-1}$  of samples **3c** and **4c** is not uncommon in sophorolipid systems, as it was observed for fibrillary but also lamellar and vesicular morphologies.<sup>15,18,19</sup> An exponent value of -2 was found in both lamellar and fibrillary sophorolipid aggregates,<sup>15,18</sup> and half-integer values were even observed for polydisperse fibrillary systems.<sup>79</sup> Considering the previous data collected by us on both fibrillary and lamellar assemblies of sophorolipids,<sup>15,79</sup> the SAXS profiles of **3c** and **4c** are

safely attributed to polydisperse fibers, in which the lipid forms a semi-crystalline arrangement within the plane of the fiber.<sup>15,79</sup> Bilayer lamellar systems are, on the contrary, characterized by two well-defined, sharp, diffraction peaks with lamellar order.<sup>18</sup> In the present case, we only observe a small, broad, peak, the actual interpretation of which is matter of debate (liquid crystalline vs. crystalline packing) in the literature for self-assembled fibers composed of low molecular weight compounds.<sup>80</sup>

*Quaternary ammonium derivatives of SL: Y-shaped surfactants with three hydrophilic centers.*

The octadecyl amine/ammonium (**3d**, pH 11/pH 3) and methyl ammonium (**4d**) derivatives show the same profiles of core-shell micelles under repulsive interactions (Figure 6). Addition of the octadecyl chain to the symmetric bolaamphiphiles has the same effect as observed on the asymmetric ones: the long alkyl chain enhances the hydrophobic component in the free energy of micellization,<sup>33,34</sup> driving the formation of micelles. Samples **3d** and **4d** cannot be considered as divalent surfactants as some samples in the **1**-series and sample **2g** in Figure 2, but rather atypical Y-shape surfactants, having a small hydrophilic center (the nitrogen atom) and two bulky headgroups composed of alkyl sophorose. Interestingly, the ammonium derivative (**3d**) at pH 3 and methyl ammonium (**4d**) in water provide the same SAXS profiles, both characterized by the typical interaction hump centered at about  $0.5 \text{ nm}^{-1}$  (Figure 6), indicative of repulsive interactions due to the positive charges on the ammonium center. This is confirmed by the fact that the hump is absent in the profile of the amine sample at pH 11. The same core-shell sphere model with screened Coulomb potential has been used to fit the data (fits are shown in Figure 6) and it confirms that **4d** and **3d** at pH 11 are composed of charged micelles having a comparable  $R_{core}$  ( $\sim 1.1 \text{ nm}$ ) and a hydrophilic shell with thickness in the order of  $2 \text{ nm}$  (Table 5b). The aggregation number is below 40 and in the same order for both systems, considering the uncertainty.  $N_{agg}$  is not less than half than what was found for the single sophorose divalent lipids above, which is not surprising, considering that the molecular weight of the Y-shaped lipids is about 1.5 larger than for the divalent ones. The surface-to-molecule ratio increases by a factor 1.6 when going from the divalent ( $\sim 2 \text{ nm}^2/\text{molecule}$ ) to the Y-shaped lipids ( $\sim 3.3 \text{ nm}^2/\text{molecule}$ ), as one would expect from a compound with a much bulkier headgroup. Interestingly, the thickness of the shell is practically the same between the Y-shaped (Table 5b) and divalent (Table 4) lipids, confirming that the bulkier headgroup in the Y-shaped occupies larger volume. In the absence of the positive charge (**3d**, pH

11),  $N_{agg}$  is almost doubled (60), which is consistent with a denser molecular packing per unit area (surface-to-molecule ratio is about  $2.5 \text{ nm}^2/\text{molecule}$  in the amine against  $\sim 3.3 \text{ nm}^2/\text{molecule}$  for the ammonium derivative). This observation is also consistent with the loss in the electrostatic repulsion for the amine derivatives.



**Figure 6 – SAXS profiles of bolaform symmetric with three hydrophilic centers (3a through 3c, 4a through 4c) and Y-shaped (3d, 4dg) sophorolipid quaternary ammonium derivatives. Concentration for samples 3: 10 mg/mL; concentration for samples 4: 12.5 mg/mL. The fit for 3d and 4d have been obtained using the core-shell sphere form factor combined with a screened Coulomb potential structure factor.**

In conclusion for this section, we can argue that the self-assembly in water of both charged and uncharged symmetrical bolaform sophorolipids is limited for small-sized substituents (methyl butyl) but it becomes relevant for the benzyl group in the ammonium form only (ammonium and methyl ammonium). These specific conditions seem to promote fibrillation, the most plausible morphology on the basis of the SAXS signal, a fact which could only be explained by the promotion of intermolecular  $\pi$ - $\pi$  stacking.<sup>81</sup> On the contrary, a long alkyl chain induces the formation of a Y-

surfactant and it drives the formation of well-defined micellar aggregates, where the double alkyl sophorose moieties are contained in the hydrophilic shell. Interestingly, the micellar charge, which is still small relatively to alkyltrimethylammonium surfactants, can nonetheless be modulated with pH for the amino derivative, which nicely demonstrates the double hydrophilic role of the both the ammonium and sophorose moieties. Simple alkyl amines do not have an amphiphilic behavior, in contrast to the more hydrophilic ammonium derivative. On the contrary, the self-assembly behavior of compound **3d** at both pH 3 and pH 11, respectively in its ammonium and amine configurations, shows that micelles form even for the amino derivative, thus demonstrating the strong hydrophilic role of the two sophorose moieties. The influence of the third hydrophilic center is also demonstrated at low pH for **3d** and **4d** compounds: the positive charge on the ammonium headgroup introduces repulsive interactions between the micelles.

**Table 5 - Structural parameter values extracted from the SAXS profiles in Figure 6. a) Radius of gyration,  $R_g$ , obtained through the Guinier approximation. b) Values are obtained using the core-shell sphere form factor combined with a screened Coulomb potential structure factor. All parameters are explained in Table 1 and Table 2**

**Table 2**

Sample	pH	$R_g/\text{nm}$	Power law exponent	Peak position/ $\text{nm}^{-1}$
3a	11	1.2		
	3	1.2		
3b	11	1.9		
	3	1.2		
3c	11	2.3		
	3		-1.6	2.3
4a		1.2		
4b		1.2		
4c			-1.8	2.3

a)

Sample	$R_{core}/\text{nm}$	$T/\text{nm}$	$Q/e$	$\rho/10^{-4} \text{ nm}^{-2}$		$[I]/M$	$\varphi$	$V_m/\text{nm}^3$	$N_{agg}$	$N_{agg} (\text{max,min})$
				$\rho_{shell}$	$\rho_{core}$					
4d	1.1	2.0	3.9	9.86	8.34	0.009	0.011	3.38	40	52,26
3d (pH11)	1.3	2.3	0.0	9.76	8.00	0	0.009	3.00	62	74,48
3d (pH3)	1.2	1.6	5.2	9.78	8.70	0.008 <sup>§</sup>	0.009	3.21	31	43,17

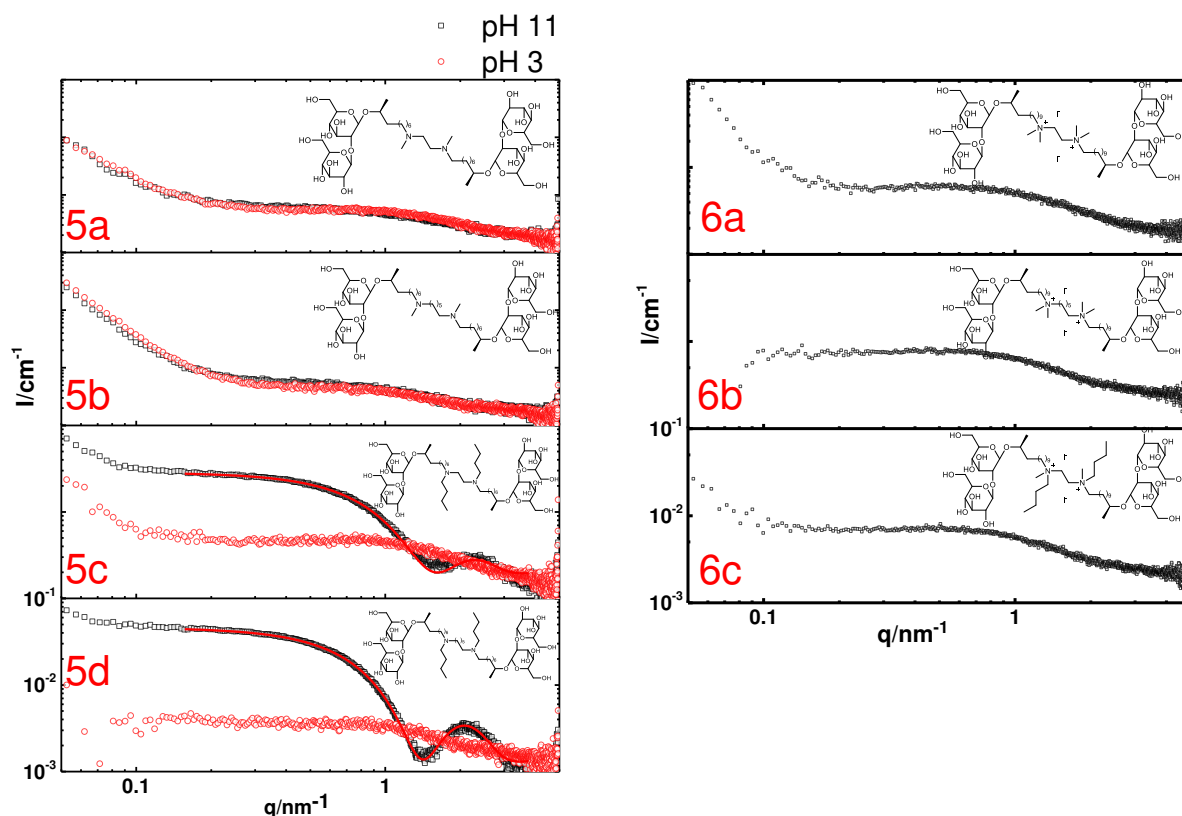
b)

<sup>§</sup> This sample contain  $\text{Cl}^-$  instead of  $\text{I}^-$ .  $V_m$  is calculated using the  $\text{Cl}^-$  ionic radius (0.181 nm)<sup>82</sup> and a hydration number of 6<sup>83</sup>



*Quaternary diammonium derivatives of SL: symmetrical bolaamphiphiles with four hydrophilic centers.*

Bolaform amphiphiles are generally characterized by a hydrophobic linker connecting two hydrophilic headgroups. Compounds **5a-5d** and **6a-6c** (Figure 7) are atypical, as they are characterized by two intra-alkyl amine/ammonium centers, unpredictably affecting the hydrophilic/hydrophobic balance. In this set of samples, we want to test the amphiphilic character for two inter-nitrogen spacers (C2 against C6) for the methyl and butyl derivatives.



**Figure 7** – SAXS profiles of bolaform symmetric quaternary ammonium sophorolipid derivatives with four hydrophilic centers. Concentration for samples 5: 10 mg/mL; concentration for samples 6: 12.5 mg/mL. The fit for 5c and 5d have been obtained using a core-shell sphere form factor (no structure factor used) and the corresponding structural parameters are provided in Table S 4.

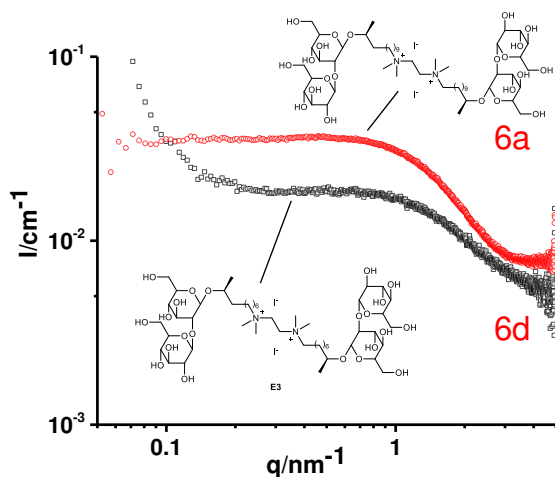
The typical SAXS profiles for **5a** and **5d** (C= 10 mg/mL, pH 11 and pH 3 in Figure 7, and pH 3, pH 5, pH 7, pH 9, pH 11 in Figure S 8) and for **6a** through **6c** (C= 12.5 mg/mL and Figure S 9) show a scattering pattern of low intensity, as observed previously throughout this work.

Similarly, the estimated  $R_g$  are below 0.8 nm (Table 6), which, compared to the size of the molecules, expected to be in the order of 4 nm, do not make this set of samples compatible with micellar aggregates. The only exception is observed for samples **5c** and **5d** at pH 11, having different spacers (respectively C2 and C6) but the same butyl substitution. At pH 11, the compounds are in the amine form, less hydrophilic than the ammonium derivatives (**5a**, **5b**, **6a-6c**). Under these conditions,  $R_g$  becomes reasonably high (above 2 nm, Table 6) and a core-shell model (fits are given in Figure 7) suggests the presence of a small core ( $R_{core}=0.7$  nm) compared to the shell ( $\sim 2$  nm).  $N_{agg}$  is evaluated to be below 30 (Table S 4). The small  $R_{core}$  is equivalent to an alkyl chain containing between 5 and 6 CH<sub>2</sub> groups, a fact which could only be explained by a strong hydration of the spacer between the sophorose groups in **5c** and **5d**. If the butyl groups increase the hydrophobic character of the core, thus driving micelles formation, the difference in the inter-nitrogen spacer (C2 vs. C6) only seems to impact shell size (respectively 1.8 nm vs 2.0 nm, Table S 4), as one can see by the slight, but noticeable, difference in the oscillation profile of the SAXS curves above 1 nm<sup>-1</sup> (Figure 7). One should note that the butyl groups are not hydrophobic enough to counterbalance the ammonium positive charge at pH < 8, which contribute, as expected, to reduce the hydrophobic character of the entire sophorose-to-sophorose spacer. This is also confirmed by the poorly-contrasted scattering profile of the methyl butyl ammonium derivative (**6c**). These hypotheses are confirmed by the difference in the scattering profile between **6a** and **6d** (Figure 8): the longer nitrogen-to-sophorose spacer (C11) of **6a** compared to **6d** (C8 spacer) enhances the hydrophobic character of the entire sophorose-to-sophorose spacer, reducing the hydration and adding more contrast (the signal of **6a** has a higher intensity). In summary, bolaform systems having more hydrophilic centers lose their amphiphilic character unless the bulky groups and longer spacers are introduced, but even in this case, the amphiphilic character is quite mild and small interpenetrated micelles with relatively low aggregation numbers are formed.

**Table 6 - Structural parameter values extracted from the SAXS profiles in Figure 7. a) Radius of gyration,  $R_g$ , obtained through the Guinier approximation.**

Sample	pH	$R_g$ /nm
5a	11	0.7
	3	0.6
5b	11	0.8
	3	0.7
5c	11	1.9
	3	0.6

5d	11	2.4
	3	0.6
6a		0.7
6b		0.8
6c		0.9

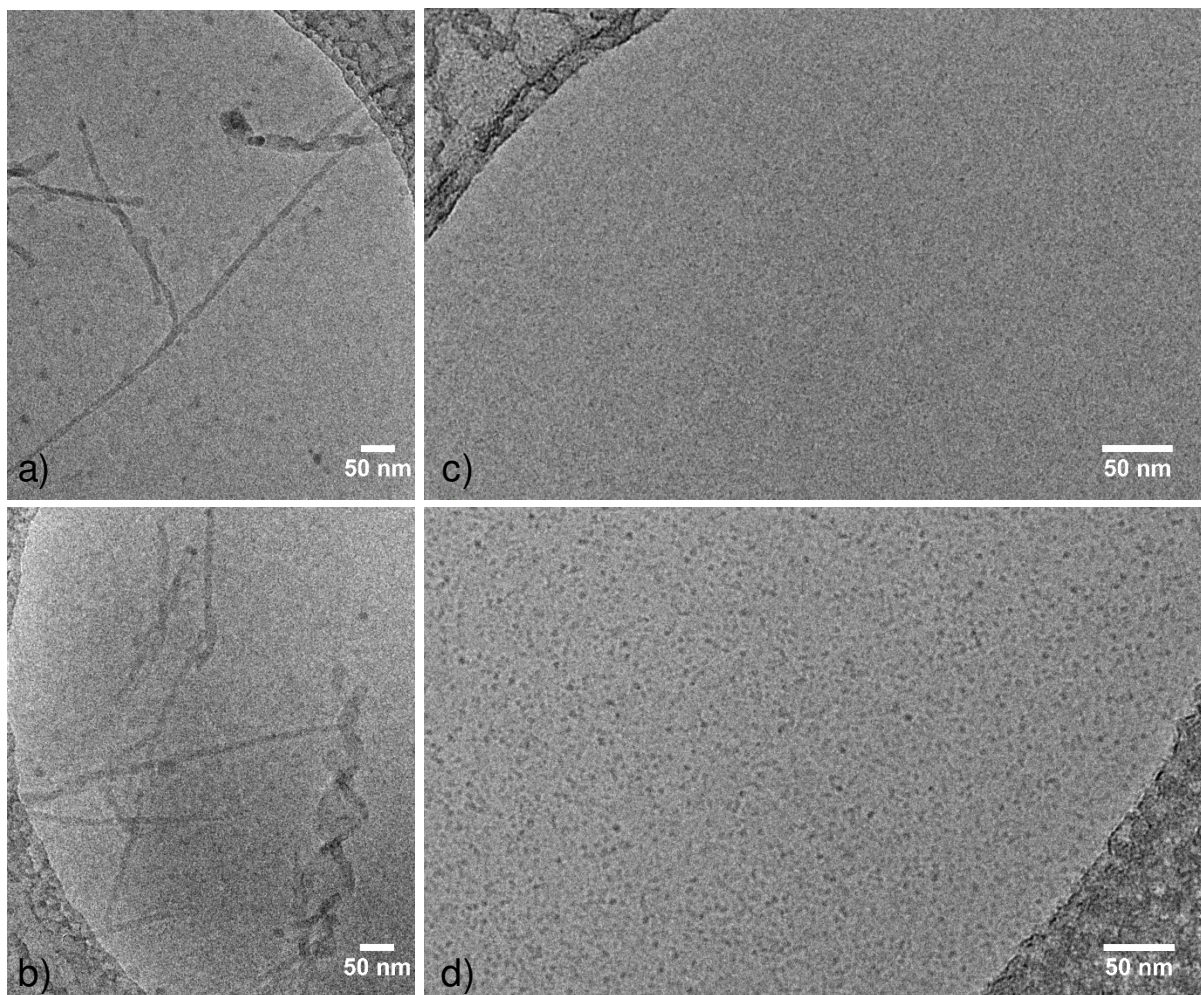


**Figure 8 – SAXS profiles recorded for sample 6a (C11 nitrogen-to-sophorose spacer) and 6d (C8 nitrogen-to-sophorose spacer). Both samples are bolaform symmetric quaternary ammonium sophorolipid derivatives with four hydrophilic centers**

The entire set of compounds studied in this work was initially conceived to develop new biobased antimicrobial and transfection agents.<sup>29,31,84</sup> We have recently shown that it exists a good correlation between the antimicrobial and transfection efficiency and the self-assembly properties of the compounds,<sup>31</sup> and the present work provides a more general outlook on the latter aspect of this new family of sophorolipid derivatives. On another level, bolaamphiphiles have certainly been investigated for years,<sup>42</sup> but the molecular design is generally the same: two symmetric or asymmetric hydrophilic headgroups are separated by a hydrophobic spacer. In this work, we go beyond this construction and we show a set of data collected on atypically molecules like hydrophobic spacers containing three or four hydrophilic centers, Y-shaped or divalent bolaamphiphiles with two different headgroups separated by a long  $C_x$  ( $x=8,11$ ) spacer.

All in all, the entire set of SAXS data can be summarized into three main scattering profiles depicting: 1) well-defined micelles of total radius above 3 nm, generally characterized by repulsive interactions (typical profile of sample 1f, Figure 4) and with aggregation numbers above 50; 2)

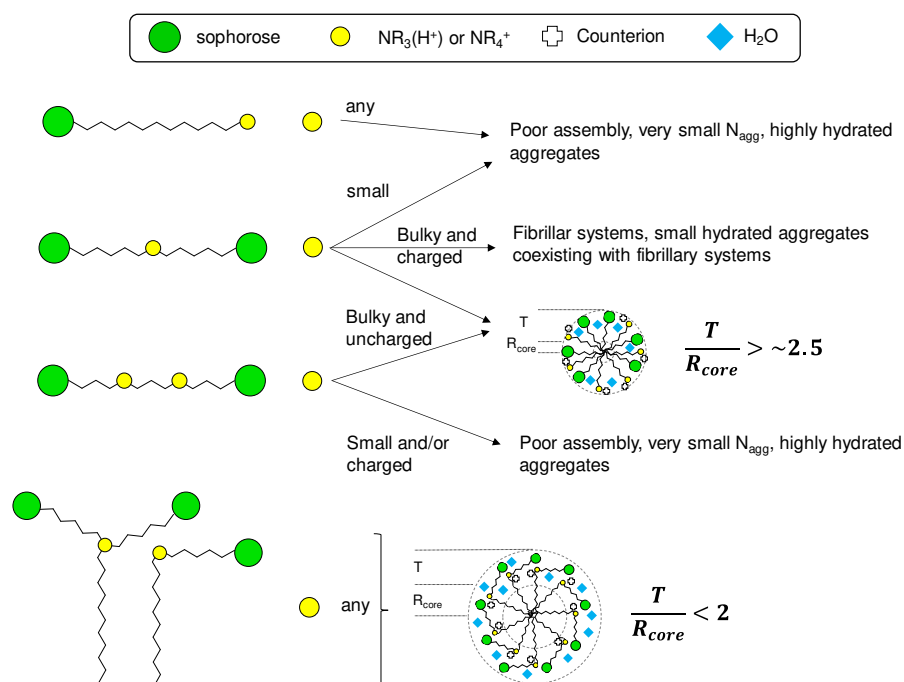
poorly-defined aggregates with total radius generally well below 1.5 nm and small aggregation numbers (typical profiles are shown for samples 3a-c at pH 11, Figure 6); 3) semi-crystalline fibrillar aggregates (typical profile is shown for sample 3c at pH 3, Figure 6). To confirm these attributions, we have run cryo-TEM experiments on three representative samples. Figure 9a,b show, indeed, the presence of twisted fibers for sample 3c pH 3. These images confirm that the typical SAXS pattern (3c pH 3, Figure 6) is characteristic of these morphologies, as hypothesized above on the basis of the similarities between the SAXS patterns in Figure 6 and the ones recorded on stearic acid derivatives of sophorolipids.<sup>15,79</sup> Figure 9d, recorded on sample 1f, shows the presence of spheroidal micelles of typical size between 3 nm and 4 nm, in agreement with the corresponding SAXS analysis (sample 1f, Figure 4 and Table 4). Finally, in Figure 9c, recorded on sample 3c pH 11, one can hardly observe the presence of micellar aggregates (dark grey spheres) of size below 2 nm, even if one should keep in mind that the resolution of our microscope under cryogenic conditions reveals to be limited to provide a clear-cut observation of such small aggregates. For this reason, cryo-TEM of aggregates of even smaller sizes (e.g., samples 1a-1d) is practically useless. Nonetheless, if compared to the image in Figure 9d, taken at the same magnification (40000x), one can safely state that sample 3c pH 11 produces micellar aggregates of smaller size than sample 1f, in agreement with the corresponding SAXS analysis (Table 5). Finally, cryo-TEM of micellar samples 1f and 3c pH 11 also shows the spurious presence of much larger aggregates of undefined shape (not shown). These are clearly part of the sample but their rarity on the TEM grid does not allow a clear-cut attribution. Nonetheless, their presence is in total agreement with the intense low-q scattering signal characterizing their corresponding SAXS data and, in fact, characterizing the SAXS patterns of almost all samples studied here. The spurious (less than about 5% in term of number of scattering objects) presence of large-scale (micron) aggregates coexisting with micellar solution of sophoro and glucolipids is not uncommon; it was reported before<sup>13,18,19,72</sup> and, according to the type of compound, we identified either nanoscale platelets (in sophorolipids)<sup>72</sup> or bilayers (in glucolipids).<sup>18,19</sup>



**Figure 9 – Cryo-TEM images recorded on samples a-b) 3c pH 3 (C= 10 mg/mL; corresponding SAXS pattern and description are respectively given in Figure 6 and Table 5), c) 3c pH 11 (C= 10 mg/mL, Figure 6 and Table 5) and d) 1f (C= 25 mg/mL, corresponding SAXS pattern and description are respectively given in Figure 4 and Table 4)**

Figure 10 summarizes the global impact of nitrogen substitutions in sophorolipid derivatives. All C8 asymmetric (**1a-1e** and **2a-2f**), as well as symmetric derivatives with three (**3a**, **3b**, **4a**, **4b**) or four (**5a-5d**, **6a-6c**) hydrophilic centers, mainly in the ammonium (charged) form, containing aliphatic substitutions on the nitrogen atom(s) have poor, if no, self-assembling capacity. The aggregates are generally smaller than the size of the molecule itself and the aggregation numbers lay below 10-20 molecules. We suppose that hydrophobic interactions are not important enough to decrease the overall free energy of micellization.<sup>34,58</sup> Uncharged symmetric derivatives with either two or three hydrophilic centers bearing bulky substituents have,

on the contrary, the ability to form well-defined micellar aggregates (e.g., **3b-3c** or **5c-5d** at basic pH). In this case, the thickness of the hydrophilic shell is larger than the core radius, indicating that an important portion of the molecule, composed of sophorose and part of the spacer, lay in the hydrophilic region. When these same compounds are charged, they have the tendency to form fibrillar systems, the structure and mechanism of formation of which needs further clarification. Finally, Y-shaped (**3d**, **4d**) and divalent (**1f-1l**, **2g**) derivatives have more classical micellization properties with aggregation number above 50; these are attributed to the enhanced hydrophobic interactions carried by the long aliphatic chain. Interestingly, the size of the hydrophilic shell corresponds to the typical size of sophorose-C<sub>x</sub>-ammonium indicating that the effective headgroup is composed of the charged ammonium headgroup, the spacer and sophorose. The data acquired on **3d** at acidic and basic pH as well as on the amine oxide **2g** sample actually indicate that the charge of the ammonium headgroup is not necessary to the purpose of micellization.



**Figure 10 – Summary of the self-assembly properties of the set of bolaform compounds studied in this work: asymmetric, symmetric with three hydrophilic centers, symmetric with four hydrophilic centers. Divalent and Y-shaped surfactants. The former have poor self-assembly properties; in the presence of an amine center and bulky substituents, symmetric bolaform compounds form small hydrated micellar aggregates with aggregation number below 20. Divalent and Y-shaped surfactants have good self-assembly properties: well-defined micellar aggregates with aggregation numbers above 50 are formed.**



## **Conclusion**

We have studied the solution self-assembly properties of a broad family of amino-derivatives of sophorolipid biosurfactant: asymmetric (sophorose-ammonium and sophorose-amine oxide) and symmetric (sophorose-sophorose bearing three and four hydrophilic centers) bolaamphiphiles are studied as a function of the substitution on the nitrogen atom. Many of these compounds have shown poor tendency to self-assemble into micelles, which, in the best case scenario and only for the most hydrophobic substituents (butyl, aryl), are stable and characterized by small ( $< 20$ ) aggregation numbers. The compounds are most likely to adopt an interpenetrated configuration within the micelles, characterized by a small hydrophobic core. Meanwhile, divalent and Y-shaped sophorolipid amino derivatives, composed of an alkylammonium hydrophobic backbone and a complex alkylsophorose hydrophilic headgroup, systematically self-assemble into micelles of aggregation numbers contained between 30 and 100, as a function of the size of the hydrophobic chain and spacers between sophorose and nitrogen. Modelling of SAXS data shows that the size of the aliphatic chain has an expected influence on the total micellar size, where larger micelle have longer chains, with a total micellar radius (core radius + shell thickness) contained between 2.5 nm and 4 nm. These experiments also highlight that the hydrophilic shell is most likely composed of the complex headgroup constituted by sophorose, nitrogen and the corresponding aliphatic spacer, the length of which has little influence on the shell thickness but it rather impacts the aggregation number (the shorter the spacer, the higher the aggregation number). In few cases, and out of any prediction, we observe an atypical SAXS signal, which we attribute to the formation of fibrillary, semi-crystalline, aggregates.

Unmodified acidic sophorolipids are known to form micelles in water and these experiments clearly show that modification of acidic sophorolipids does not necessarily improve their aggregation behavior in water, unless a long aliphatic chain is introduced to bring the hydrophobic interaction, necessary to drive the formation of well-defined micelles. This piece of information is to be taken into account in future development of amino derived sophorolipid derivatives with potential application as transfection or antimicrobial agents.

## **Acknowledgements**

This work received financial support by the European Synchrotron Radiation Facility (ESRF), Grenoble, France, under the experiment number MX1821. This work benefited from the use of the SasView application, originally developed under NSF award DMR-0520547. SasView contains code developed with funding from the European Union's Horizon 2020 research and innovation programme under the SINE2020 project, grant agreement No 654000. The research leading to these results has also received funding from the Long Term Structural Methusalem funding by the Flemish Government (grant number BOF09/01M00409). We thank Dr. Gervaise Mosser (LCMCP, Sorbonne Université, Paris, France) for her assistance on the cryo-TEM experiments.

### References:

- (1) Brown, R. C.; Brown, T. R. *Biorenewable Resources: Engineering New Products from Agriculture: Second Edition*; 2014. <https://doi.org/10.1002/9781118524985>.
- (2) Metzger, J. O.; Eissen, M. Concepts on the Contribution of Chemistry to a Sustainable Development. Renewable Raw Materials. *Comptes Rendus Chim.* **2004**. <https://doi.org/10.1016/j.crci.2003.12.003>.
- (3) Clark, J.; Deswarte, F. *Introduction to Chemicals from Biomass: Second Edition*; 2015. <https://doi.org/10.1002/9781118714478>.
- (4) Vandermeulen, V.; Van der Steen, M.; Stevens, C. V.; Van Huylenbroeck, G. Industry Expectations Regarding the Transition toward a Biobased Economy. *Biofuels, Bioprod. Biorefining* **2012**. <https://doi.org/10.1002/bbb.1333>.
- (5) Van Bogaert, I. N. A.; Saerens, K.; De Muynck, C.; Develter, D.; Soetaert, W.; Vandamme, E. J. Microbial Production and Application of Sophorolipids. *Appl. Microbiol. Biotechnol.* **2007**, *76* (1), 23–34. <https://doi.org/10.1007/s00253-007-0988-7>.
- (6) Van Bogaert, I. N. A.; Zhang, J.; Soetaert, W. Microbial Synthesis of Sophorolipids. *Process Biochem.* **2011**, *46* (4), 821–833. <https://doi.org/10.1016/j.procbio.2011.01.010>.
- (7) Dhasaiyan, P.; Le Griel, P.; Roelants, S.; Redant, E.; Van Bogaert, I. N. A.; Prevost, S.; Prasad, B. L. V; Baccile, N. Micelles versus Ribbons: How Congeners Drive the Self-Assembly of Acidic Sophorolipid Biosurfactants. *ChemPhysChem* **2017**, *18*, 643–652. <https://doi.org/10.1002/cphc.201601323>.
- (8) Davila, A. M.; Marchal, R.; Vandecasteele, J. P. Sophorose Lipid Production from Lipidic Precursors: Predictive Evaluation of Industrial Substrates. *J. Ind. Microbiol.* **1994**.



<https://doi.org/10.1007/BF01569757>.

- (9) Pekin, G.; Vardar-Sukan, F.; Kosaric, N. Production of Sophorolipids from *Candida Bombicola* ATCC 22214 Using Turkish Corn Oil and Honey. *Eng. Life Sci.* **2005**.  
<https://doi.org/10.1002/elsc.200520086>.
- (10) Zhou, S.; Xu, C.; Wang, J.; Gao, W.; Akhverdiyeva, R.; Shah, V.; Gross, R. Supramolecular Assemblies of a Naturally Derived Sophorolipid. *Langmuir* **2004**, *20* (19), 7926–7932. <https://doi.org/10.1021/la048590s>.
- (11) Penfold, J.; Chen, M.; Thomas, R. K.; Dong, C.; Smyth, T. J. P.; Perfumo, A.; Marchant, R.; Banat, I. M.; Stevenson, P.; Parry, A.; et al. Solution Self-Assembly of the Sophorolipid Biosurfactant and Its Mixture with Anionic Surfactant Sodium Dodecyl Benzene Sulfonate. *Langmuir* **2011**, *27* (14), 8867–8877.  
<https://doi.org/10.1021/la201661y>.
- (12) Baccile, N.; Nassif, N.; Malfatti, L.; Van Bogaert, I. N. a.; Soetaert, W.; Pehau-Arnaudet, G.; Babonneau, F. Sophorolipids: A Yeast-Derived Glycolipid as Greener Structure Directing Agents for Self-Assembled Nanomaterials. *Green Chem.* **2010**, *12* (9), 1564–1567. <https://doi.org/10.1039/c0gc00163e>.
- (13) Baccile, N.; Babonneau, F.; Jestin, J.; P ehau-Arnaudet, G.; Van Bogaert, I. Unusual, PH-Induced, Self-Assembly of Sophorolipid Biosurfactants. *ACS Nano* **2012**, *6* (6), 4763–4776. <https://doi.org/10.1021/nn204911k>.
- (14) Dhasaiyan, P.; Banerjee, A.; Visaveliya, N.; Prasad, B. L. V. Influence of the Sophorolipid Molecular Geometry on Their Self-Assembled Structures. *Chem. Asian J.* **2013**, *8* (2), 369–372. <https://doi.org/10.1002/asia.201200935>.
- (15) Cuvier, A.-S.; Berton, J.; Stevens, C. V; Fadda, G. C.; Babonneau, F.; Van Bogaert, I. N. a; Soetaert, W.; Pehau-Arnaudet, G.; Baccile, N. PH-Triggered Formation of Nanoribbons from Yeast-Derived Glycolipid Biosurfactants. *Soft Matter* **2014**, *10* (22), 3950–3959.  
<https://doi.org/10.1039/c4sm00111g>.
- (16) Dhasaiyan, P.; Pandey, P. R.; Visaveliya, N.; Roy, S.; Prasad, B. L. V. Vesicle Structures from Bolaamphiphilic Biosurfactants: Experimental and Molecular Dynamics Simulation Studies on the Effect of Unsaturation on Sophorolipid Self-Assemblies. *Chem. - A Eur. J.* **2014**, *20* (21), 6246–6250. <https://doi.org/10.1002/chem.201304719>.
- (17) Dhasaiyan, P.; Pr evost, S.; Baccile, N.; Prasad, B. L. V. PH- and Time-Resolved in-Situ

- SAXS Study of Self-Assembled Twisted Ribbons Formed by Elaidic Acid Sophorolipids. *Langmuir* **2018**, *34*, 2121–2131. <https://doi.org/10.1021/acs.langmuir.7b03164>.
- (18) Baccile, N.; Selmane, M.; Le Griel, P.; Prévost, S.; Perez, J.; Stevens, C. V.; Delbeke, E.; Zibek, S.; Guenther, M.; Soetaert, W.; et al. PH-Driven Self-Assembly of Acidic Microbial Glycolipids. *Langmuir* **2016**, *32* (25), 6343–6359. <https://doi.org/10.1021/acs.langmuir.6b00488>.
- (19) Baccile, N.; Cuvier, A.-S.; Prévost, S.; Stevens, C. V.; Delbeke, E.; Berton, J.; Soetaert, W.; Van Bogaert, I. N. A.; Roelants, S. Self-Assembly Mechanism of PH-Responsive Glycolipids: Micelles, Fibers, Vesicles, and Bilayers. *Langmuir* **2016**, *32* (42), 10881–10894. <https://doi.org/10.1021/acs.langmuir.6b02337>.
- (20) Desai, J. D.; Banat, I. M. Microbial Production of Surfactants and Their Commercial Potential. *Microbiol. Mol. Biol. Rev.* **1997**, *61* (1), 47–64. [https://doi.org/10.1016/S0140-6701\(97\)84559-6](https://doi.org/10.1016/S0140-6701(97)84559-6).
- (21) Lang, S. Biological Amphiphiles (Microbial Biosurfactants). *Curr. Opin. Colloid Interface Sci.* **2002**, *7* (1–2), 12–20. [https://doi.org/10.1016/S1359-0294\(02\)00007-9](https://doi.org/10.1016/S1359-0294(02)00007-9).
- (22) Koh, A.; Linhardt, R. J.; Gross, R. Effect of Sophorolipid N-Alkyl Ester Chain Length on Its Interfacial Properties at the Almond Oil-Water Interface. *Langmuir* **2016**, *32* (22), 5562–5572. <https://doi.org/10.1021/acs.langmuir.6b01008>.
- (23) Koh, A.; Gross, R. A Versatile Family of Sophorolipid Esters: Engineering Surfactant Structure for Stabilization of Lemon Oil-Water Interfaces. *Colloids Surfaces A Physicochem. Eng. Asp.* **2016**. <https://doi.org/10.1016/j.colsurfa.2016.07.089>.
- (24) Koh, A.; Gross, R. Molecular Editing of Sophorolipids by Esterification of Lipid Moieties: Effects on Interfacial Properties at Paraffin and Synthetic Crude Oil-Water Interfaces. *Colloids Surfaces A Physicochem. Eng. Asp.* **2016**. <https://doi.org/10.1016/j.colsurfa.2016.07.084>.
- (25) Meng, J.; Zhang, P.; Wang, S. Recent Progress in Biointerfaces with Controlled Bacterial Adhesion by Using Chemical and Physical Methods. *Chem. - An Asian J.* **2014**, *9* (8), 2004–2016. <https://doi.org/10.1002/asia.201402200>.
- (26) Zini, E.; Gazzano, M.; Scandola, M.; Wallner, S. R.; Gross, R. A. Glycolipid Biomaterials: Solid-State Properties of a Poly(Sophorolipid). *Macromolecules* **2008**. <https://doi.org/10.1021/ma800496f>.

- (27) Delbeke, E. I. P.; Movsisyan, M.; Van Geem, K. M.; Stevens, C. V. Chemical and Enzymatic Modification of Sophorolipids. *Green Chem.* **2016**, *18* (October), 76–104. <https://doi.org/10.1039/C5GC02187A>.
- (28) Peng, Y.; Totsingan, F.; Meier, M. A. R.; Steinmann, M.; Wurm, F.; Koh, A.; Gross, R. A. Sophorolipids: Expanding Structural Diversity by Ring-Opening Cross-Metathesis. *Eur. J. Lipid Sci. Technol.* **2015**, *117*, 217–228. <https://doi.org/10.1002/ejlt.201400466>.
- (29) Delbeke, E. I. P.; Roman, B. I.; Marin, G. B.; Van Geem, K. M.; Stevens, C. V. A New Class of Antimicrobial Biosurfactants: Quaternary Ammonium Sophorolipids. *Green Chem.* **2015**, *17* (6), 3373–3377. <https://doi.org/10.1039/C5GC00120J>.
- (30) Delbeke, E. I. P.; Roelants, S. L. K. W.; Matthijs, N.; Everaert, B.; Soetaert, W.; Coenye, T.; Van Geem, K. M.; Stevens, C. V. Sophorolipid Amine Oxide Production by a Combination of Fermentation Scale-up and Chemical Modification. *Ind. Eng. Chem. Res.* **2016**, *55* (27), 7273–7281. <https://doi.org/10.1021/acs.iecr.6b00629>.
- (31) Delbeke, E. I. P.; Everaert, J.; Lozach, O.; Gall, T. Le; Berchel, M.; Montier, T.; Jaffrès, P.-A.; Rigole, P.; Coenye, T.; Brennich, M.; et al. Synthesis and Biological Evaluation of Bolaamphiphilic Sophorolipids. *ACS Sustain. Chem. Eng.* **2018**, *6*, 8992–9005.
- (32) Delbeke, E. I. P.; Everaert, J.; Uitterhaegen, E.; Verweire, S.; Verlee, A.; Talou, T.; Soetaert, W.; Van Bogaert, I. N. A.; Stevens, C. V. Petroselinic Acid Purification and Its Use for the Fermentation of New Sophorolipids. *AMB Express* **2016**, *6* (1), 28. <https://doi.org/10.1186/s13568-016-0199-7>.
- (33) Israelachvili, J. N.; Mitchell, D. J.; Ninham, B. W. Theory of Self-Assembly of Hydrocarbon Amphiphiles into Micelles and Bilayers. *J. Chem. Soc. Faraday Trans. 2* **1976**, *72*, 1525. <https://doi.org/10.1039/f29767201525>.
- (34) Nagarajan, R. Molecular Packing Parameter and Surfactant Self-Assembly: The Neglected Role of the Surfactant Tail Self-Assembly. *Langmuir* **2002**, *18*, 31–38.
- (35) Israelachvili, J. N.; Mitchell, D. J. A MODEL FOR THE PACKING OF LIPIDS IN BILAYER MEMBRANES. *Biochim. Biophys. Acta* **1975**, *389*, 13–19.
- (36) Zana, R. Dimeric (Gemini) Surfactants: Effect of the Spacer Group on the Association Behavior in Aqueous Solution. *J. Colloid Interface Sci.* **2002**, *248* (2), 203–220. <https://doi.org/10.1006/jcis.2001.8104>.
- (37) Zana, R. Gemini (Dimeric) Surfactants. *Curr. Opin. Colloid Interface Sci.* **1996**, *1* (5),

- 566–571. [https://doi.org/10.1016/S1359-0294\(96\)80093-8](https://doi.org/10.1016/S1359-0294(96)80093-8).
- (38) Menger, F. M.; Littau, C. A. Gemini Surfactants: A New Class of Self-Assembling Molecules. *J. Am. Chem. Soc.* **1993**, *115* (22), 10083–10090. <https://doi.org/10.1021/ja00075a025>.
- (39) Hagslatt, H.; Soderman, O.; Jonsson, B. Divalent Surfactants - Experimental Results and Theoretical Modeling of Surfactant Water Phase Equilibria. *Langmuir* **1994**, *10* (7), 2177–2187.
- (40) Hagslätt, H.; Fontell, K. Divalent Surfactants—The Influence of Headgroup Charge on the Phase Behavior of an Alkyldiamine Surfactant/Water System. *J. Colloid Interface Sci.* **1994**, *165*, 431–444.
- (41) Nagarajan, R. Self-Assembly of Bola Amphiphiles. *Chem. Eng. Commun.* **1987**, *55*, 251–273.
- (42) Fuhrhop, J.-H.; Wang, T. Bolaamphiphiles. *Chem. Rev.* **2004**, *104* (6), 2901–2938. <https://doi.org/10.1021/cr030602b>.
- (43) Nagarajan, R.; Ruckenstein, E. Theory of Surfactant Self-Assembly: A Predictive Molecular Thermodynamic Approach. *Langmuir* **1991**, *7* (12), 2934–2969. <https://doi.org/10.1021/la00060a012>.
- (44) Israelachvili, J. N.; Ninham, B. W.; Mitchell, D. J. Theory of Self-Assembly of Lipids and Vesicles. *Biochim. Biophys. Acta* **1977**, *470*, 185–201.
- (45) Buckingham, S. A.; Garvey, C. J.; Warr, G. G. Effect of Head-Group Size on Micellization and Phase Behavior in Quaternary Ammonium Surfactant Systems. *J. Phys. Chem.* **1993**, *97* (39), 10236–10244. <https://doi.org/10.1021/j100141a054>.
- (46) Delbeke, E. I. P.; Everaert, J.; Lozach, O.; Gall, T. Le; Berchel, M.; Montier, T.; Jaffrès, P.-A.; Rigole, P.; Coenye, T.; Brennich, M.; et al. Synthesis of Novel Lipid-Based Quaternary Ammonium Sphorolipids. *Submitted* **2019**.
- (47) [http://www.esrf.eu/home/UsersAndScience/Experiments/MX/About\\_our\\_beamlines/bm29.html](http://www.esrf.eu/home/UsersAndScience/Experiments/MX/About_our_beamlines/bm29.html). ESRF BM29  
[http://www.esrf.eu/home/UsersAndScience/Experiments/MX/About\\_our\\_beamlines/bm29.html](http://www.esrf.eu/home/UsersAndScience/Experiments/MX/About_our_beamlines/bm29.html).
- (48) Pernot, P.; Round, A.; Barrett, R.; De Maria Antolinos, A.; Gobbo, A.; Gordon, E.; Huet,

- J.; Kieffer, J.; Lentini, M.; Mattenet, M.; et al. Upgraded ESRF BM29 Beamline for SAXS on Macromolecules in Solution. *J. Synchrotron Radiat.* **2013**, *20* (4), 660–664. <https://doi.org/10.1107/S0909049513010431>.
- (49) Round, A.; Felisaz, F.; Fodinger, L.; Gobbo, A.; Huet, J.; Villard, C.; Blanchet, C. E.; Pernot, P.; McSweeney, S.; Roessle, M.; et al. BioSAXS Sample Changer: A Robotic Sample Changer for Rapid and Reliable High-Throughput X-Ray Solution Scattering Experiments. *Acta Crystallogr. Sect. D* **2015**, *71* (1), 67–75. <https://doi.org/10.1107/S1399004714026959>.
- (50) Ashiotis, G.; Deschildre, A.; Nawaz, Z.; Wright, J. P.; Karkoulis, D.; Picca, F. E.; Kieffer, J. The Fast Azimuthal Integration Python Library: PyFAI. *J. Appl. Crystallogr.* **2015**, *48* (2), 510–519. <https://doi.org/10.1107/S1600576715004306>.
- (51) Glatter, O.; Kratky, O. *Small Angle X-Ray Scattering*; Academic Press: London, 1982.
- (52) Manet, S.; Cuvier, A. S.; Valotteau, C.; Fadda, G. C.; Perez, J.; Karakas, E.; Abel, S.; Baccile, N. Structure of Bolaamphiphile Sophorolipid Micelles Characterized with SAXS, SANS, and MD Simulations. *J. Phys. Chem. B* **2015**, *119* (41), 13113–13133. <https://doi.org/10.1021/acs.jpcc.5b05374>.
- (53) Baccile, N.; Pedersen, J. S.; Pehau-Arnaudet, G.; Van Bogaert, I. N. a. Surface Charge of Acidic Sophorolipid Micelles: Effect of Base and Time. *Soft Matter* **2013**, *9* (19), 4911–4922. <https://doi.org/10.1039/c3sm50160d>.
- (54) Hansen, J.-P.; Hayter, J. B. A Rescaled MSA Structure Factor for Dilute Charged Colloidal Dispersions. *Mol. Phys.* **1982**, *46* (3), 651–656. <https://doi.org/10.1080/00268978200101471>.
- (55) Hayter, J. B.; Penfold, J. Determination of Micelle Structure and Charge by Neutron Small-Angle Scattering. *Colloid Polym. Sci.* **1983**, *261* (12), 1022–1030. <https://doi.org/10.1007/BF01421709>.
- (56) Malmberg, C. G.; Maryott, A. A. Dielectric Constant of Water from 0 to 100 C. *J. Res. Natl. Bur. Stand. (1934)*. **1956**, *56* (1), 1. <https://doi.org/10.6028/jres.056.001>.
- (57) Oliver, R. C.; Lipfert, J.; Fox, D. A.; Lo, R. H.; Doniach, S.; Columbus, L. Dependence of Micelle Size and Shape on Detergent Alkyl Chain Length and Head Group. *PLoS One* **2013**, *8* (5). <https://doi.org/10.1371/journal.pone.0062488>.
- (58) Tanford, C. *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*;

Wiley-Interscience, 1973.

- (59) Reiss-Husson, F.; Luzzati, V. The Structure of the Micellar Solutions of Some Amphiphilic Compounds in Pure Water as Determined by Absolute Small-Angle X-Ray Scattering Techniques. *J. Phys. Chem.* **1964**, *68* (12), 3504–3511. <https://doi.org/10.1021/j100794a011>.
- (60) Markovich, G.; Giniger, R.; Levin, M.; Cheshnovsky, O. Photoelectron Spectroscopy of Iodine Anion Solvated in Water Clusters. *J. Chem. Phys.* **1991**, *95* (12), 9416–9419. <https://doi.org/10.1063/1.461172>.
- (61) Karmakar, A.; Chandra, A. Water in Hydration Shell of an Iodide Ion: Structure and Dynamics of Solute-Water Hydrogen Bonds and Vibrational Spectral Diffusion from First-Principles Simulations. *J. Phys. Chem. B* **2015**, *119* (27), 8561–8572. <https://doi.org/10.1021/jp510714e>.
- (62) Shiraga, K.; Suzuki, T.; Kondo, N.; De Baerdemaeker, J.; Ogawa, Y. Quantitative Characterization of Hydration State and Destructuring Effect of Monosaccharides and Disaccharides on Water Hydrogen Bond Network. *Carbohydr. Res.* **2015**, *406*, 46–54. <https://doi.org/10.1016/j.carres.2015.01.002>.
- (63) Marsh, D. Molecular Volumes of Phospholipids and Glycolipids in Membranes. *Chem. Phys. Lipids* **2010**, *163* (7), 667–677. <https://doi.org/10.1016/j.chemphyslip.2010.06.005>.
- (64) Gliński, J.; Burakowski, A. Is the Hydration Number of a Non-Electrolyte Additive with Length and Constituents of the Solute Molecule? In *European Physical Journal: Special Topics*; 2008; Vol. 154, pp 275–279. <https://doi.org/10.1140/epjst/e2008-00560-y>.
- (65) Bales, B. L.; Zana, R. Characterization of Micelles of Quaternary Ammonium Surfactants as Reaction Media I: Dodecyltrimethylammonium Bromide and Chloride. *J. Phys. Chem. B* **2002**, *106* (8), 1926–1939. <https://doi.org/10.1021/jp013813y>.
- (66) Berr, S. S.; Caponetti, E.; Johnson, J. S.; Jones, R. R. M.; Magid, L. J. Small-Angle Neutron Scattering from Hexadecyltrimethylammonium Bromide Micelles in Aqueous Solutions. *J. Phys. Chem.* **1986**, *90* (22), 5766–5770. <https://doi.org/10.1021/j100280a059>.
- (67) MacKerell, A. D. Molecular Dynamics Simulation Analysis of a Sodium Dodecyl Sulfate Micelle in Aqueous Solution: Decreased Fluidity of the Micelle Hydrocarbon Interior. *J. Phys. Chem.* **1995**, *99* (7), 1846–1855. <https://doi.org/10.1021/j100007a011>.
- (68) Dey, P. C.; Motin, M. A.; Biswas, T. K.; Huque, E. M. Apparent Molar Volume and

- Viscosity Studies on Some Carbohydrates in Solutions. *Monatshefte fur Chemie* **2003**, *134* (6), 797–809. <https://doi.org/10.1007/s00706-002-0530-7>.
- (69) Shannon, R. D. Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Crystallogr. Sect. A* **1976**, *32* (5), 751–767. <https://doi.org/10.1107/S0567739476001551>.
- (70) Davey, T. W.; Ducker, W. A.; Hayman, A. R. Aggregation of W-Hydroxy Quaternary Ammonium Bolaform Surfactants. *Langmuir* **2000**, *16* (6), 2430–2435. <https://doi.org/http://dx.doi.org/10.1071/C98001>.
- (71) Ottaviani, M. F.; Moscatelli, A.; Desplandier-Giscard, D.; Di Renzo, F.; Kooyman, P. J.; Alonso, B.; Galarneau, A. Synthesis of Micelle-Templated Silicas from Cetyltrimethylammonium Bromide/1,3,5-Trimethylbenzene Micelles. *J. Phys. Chem. B* **2004**, *108* (32), 12123–12129. <https://doi.org/10.1021/jp049032p>.
- (72) Cuvier, A. S.; Babonneau, F.; Berton, J.; Stevens, C. V.; Fadda, G. C.; Péhau-Arnaudet, G.; Le Griel, P.; Prévost, S.; Perez, J.; Baccile, N. Nanoscale Platelet Formation by Monounsaturated and Saturated Sophorolipids under Basic PH Conditions. *Chem. - A Eur. J.* **2015**, *21* (52), 19265–19277. <https://doi.org/10.1002/chem.201502933>.
- (73) Pisárčik, M.; Devínsky, F.; Pupák, M. Determination of Micelle Aggregation Numbers of Alkyltrimethylammonium Bromide and Sodium Dodecyl Sulfate Surfactants Using Time-Resolved Fluorescence Quenching. *Open Chem.* **2015**, *13* (1), 922–931. <https://doi.org/10.1515/chem-2015-0103>.
- (74) Shimizu, T.; Masuda, M. Stereochemical Effect of Even - Odd Connecting Links on Supramolecular Assemblies Made of 1-Glucosamide Bolaamphiphiles. *J. Am. Chem. Soc.* **1997**, *119*, 2812–2818.
- (75) Zhang, Q.; Gao, Z.; Xu, F.; Tai, S. Effect of Hydrocarbon Structure of the Headgroup on the Thermodynamic Properties of Micellization of Cationic Gemini Surfactants: An Electrical Conductivity Study. *J. Colloid Interface Sci.* **2012**, *371* (1), 73–81. <https://doi.org/10.1016/j.jcis.2011.12.076>.
- (76) Zana, R. Ionization of Cationic Micelles: Effect of the Detergent Structure. *J. Colloid Interface Sci.* **1980**, *78* (2), 330–337. [https://doi.org/10.1016/0021-9797\(80\)90571-8](https://doi.org/10.1016/0021-9797(80)90571-8).
- (77) Teixeira, J. Small-Angle Scattering by Fractal Systems. *J. Appl. Crystallogr.* **1988**, *21* (6), 781–785. <https://doi.org/10.1107/S0021889888000263>.

- (78) Pedersen, J. S.; Schurtenberger, P. Scattering Functions of Semiflexible Polymers with and without Excluded Volume Effects. *Macromolecules* **1996**, *29* (23), 7602–7612.  
<https://doi.org/10.1021/ma9607630>.
- (79) Cuvier, A. S.; Babonneau, F.; Berton, J.; Stevens, C. V.; Fadda, G. C.; Genois, I.; Le Griel, P.; Péhau-Arnaudet, G.; Baccile, N. Synthesis of Uniform, Monodisperse, Sophorolipid Twisted Ribbons. *Chem. - An Asian J.* **2015**, *10* (11), 2419–2426.  
<https://doi.org/10.1002/asia.201500693>.
- (80) Barclay, T. G.; Constantopoulos, K.; Matisons, J. Nanotubes Self-Assembled from Amphiphilic Molecules via Helical Intermediates. *Chem. Rev.* **2014**, *114* (20), 10217–10291. <https://doi.org/10.1021/cr400085m>.
- (81) Estroff, L. a; Hamilton, A. D. Water Gelation by Small Organic Molecules Water Gelation by Small Organic Molecules. *Chem. Rev.* **2004**, *104* (February), 1201–1218.  
<https://doi.org/10.1021/cr0302049>.
- (82) Ifft, J. B.; Williams, A. The Partial Molar Volumes of Cesium and Chloride Ions in Solution as a Function of Concentration. *Biochim. Biophys. Acta* **1967**, *136*, 151–153.
- (83) Enderby, J. E. Ion Solvation via Neutron Scattering. *Chem. soc* **1995**, *24*, 159–168.
- (84) Delbeke, E. I. P.; Lozach, O.; Le Gall, T.; Berchel, M.; Montier, T.; Jaffrès, P.-A.; Van Geem, K. M.; Stevens, C. V. Evaluation of the Transfection Efficacies of Quaternary Ammonium Salts Prepared from Sophorolipids. *Org. Biomol. Chem.* **2016**, *14*, 3744–3751.  
<https://doi.org/10.1039/c6ob00241b>.



## Supporting Information

# Asymmetrical, symmetrical, divalent and Y-shaped (bola)amphiphiles: the relationship between molecular structure and self-assembly in amino derivatives of sophorolipid biosurfactants

**Niki Baccile,<sup>a,\*</sup>† Elisabeth I. P. Delbeke,<sup>b,c</sup>† Martha Brennich,<sup>d</sup> Chloé Seyrig,<sup>a</sup> Jonas Everaert<sup>b,e</sup> Sophie L. K. W. Roelants<sup>e,f</sup> Wim Soetaert<sup>e,f</sup> Inge N. A. Van Bogaert,<sup>g</sup> Kevin M. Van Geem,<sup>c</sup> Christian V. Stevens<sup>b,\*</sup>**

<sup>a</sup> Sorbonne Université, Centre National de la Recherche Scientifique, Laboratoire de Chimie de la Matière Condensée de Paris, LCMCP, F-75005 Paris, France.

<sup>b</sup> SynBioC, Department of Green Chemistry and Technology, Ghent University, Coupure Links 653, 9000 Ghent, Belgium.

<sup>c</sup> Laboratory for Chemical Technology (LCT), Department of Materials, Textiles and Chemical Engineering, Ghent University, Technologiepark 125, 9052 Ghent, Belgium.

<sup>d</sup> European Molecular Biology Laboratory, Synchrotron Crystallography Group, 71 Avenue des Martyrs, 38042 Grenoble, France

<sup>e</sup> InBio, Department of Biotechnology, Ghent University, Coupure Links 653, 9000 Ghent, Belgium.

<sup>f</sup> Bio Base Europe Pilot Plant (BBEU), Rodenhuisenkaai 1, 9042 Ghent (Desteldonk), Belgium.

<sup>g</sup> Centre for Synthetic Biology, Department of Biotechnology, Ghent University, Coupure Links 653, 9000 Ghent, Belgium.

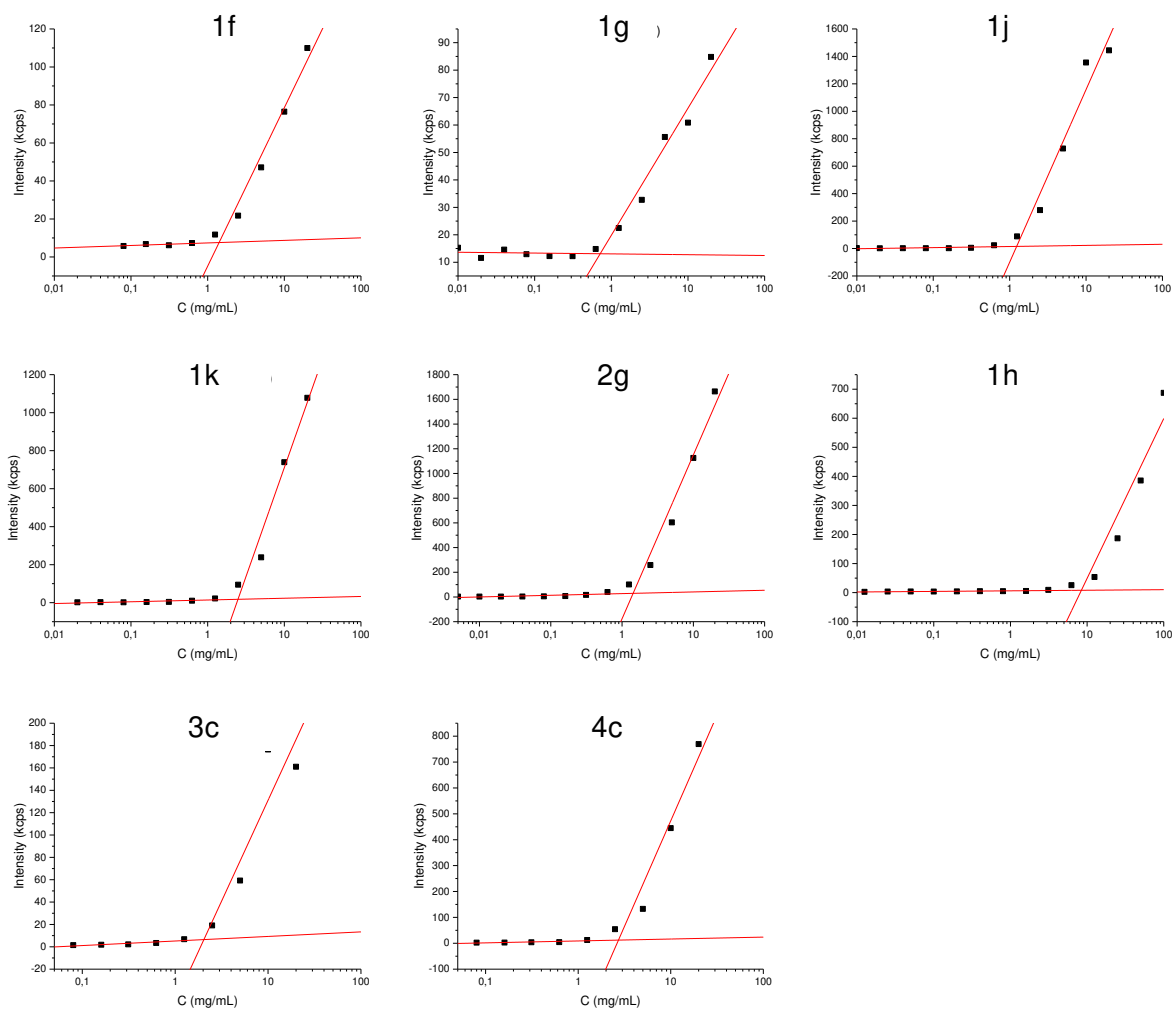
\* Corresponding authors: [niki.baccile@sorbonne-universite.fr](mailto:niki.baccile@sorbonne-universite.fr); [chris.stevens@UGent.be](mailto:chris.stevens@UGent.be)

† These authors equally contributed to this work

**Table S 1 – Decomposition of selected samples into submolecular groups for the calculation of the respective molecular volumes.  $V_{molec}$  of each subgroup is given in Table 2. The  $V_{molec}$  of each compound is indicated throughout the main text in Table 3, Table 4 and Table 5. Mw is in g/mol. CMC values are in mg/mL.**

Compound	Mw	CMC	N° sophorose	N° CH <sub>2</sub> <sup>§</sup>	N° CH <sub>3</sub>	N° CH <sup>§</sup>	N°I	N <sub>H</sub> (sophorose)	N <sub>H</sub> (I)
1g	933.4	0.7	1	27	4	1	1	20	6
1f	891.5	1.4	1	24	4	1	1	20	6
1i	807.9	-	1	18	4	1	1	20	6
1h	849.9	8.5	1	21	4	1	1	20	6
1l	849.9	-	1	21	4	1	1	20	6
1k	891.5	2.5	1	24	4	1	1	20	6
1j	933.5	1.2	1	27	4	1	1	20	6
2g	765.6	1.4	1	26	3	11	0	20	0
4c	1182.5	-	2	15	3	7	0	20	0
4d	1344.5	2.7	2	31	4	2	1	20	6
3b	1005.6	-	2	17	3	2	0	20	0
3c	1039.6	2.0	2	15	2	7	0	20	0
3d	1201.8	-	2	31	3	2	0	20	0
5c	1104.7	-	2	22	4	2	0	20	0
5d	1160.7	-	2	26	4	2	0	20	0
6a	1388.5	-	2	22	6	2	2	20	6
6d	1305.1	-	2	16	6	2	2	20	6

§ CH<sub>2</sub> and CH groups in sophorose are not included in this column



**Figure S 1 – Light scattering experiments as a function of concentration of the sophorolipid derivatives (Figure 2) in water measured at 25°C. The critical micelle concentration (cmc) is determined by the intersection of the linear fits.**

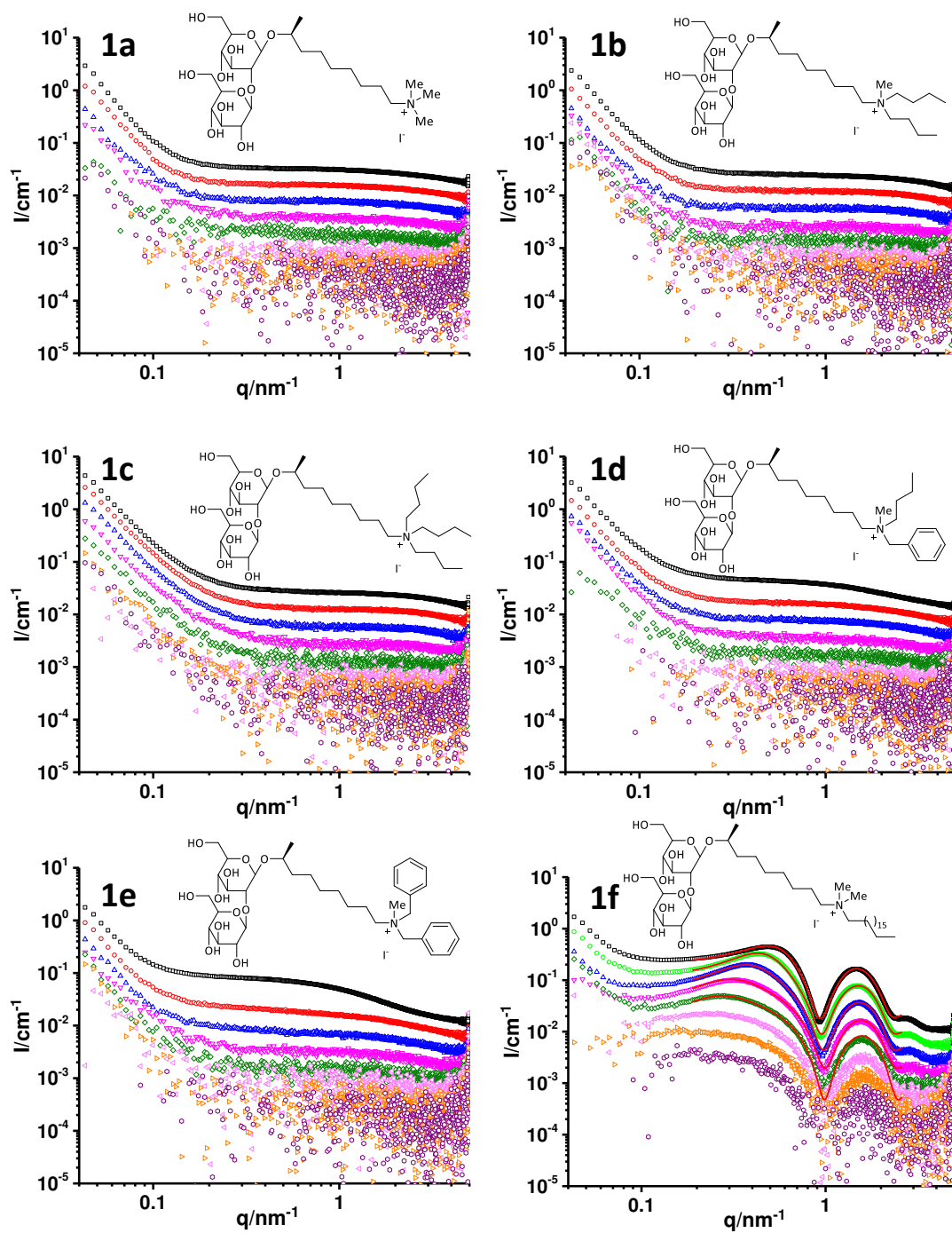
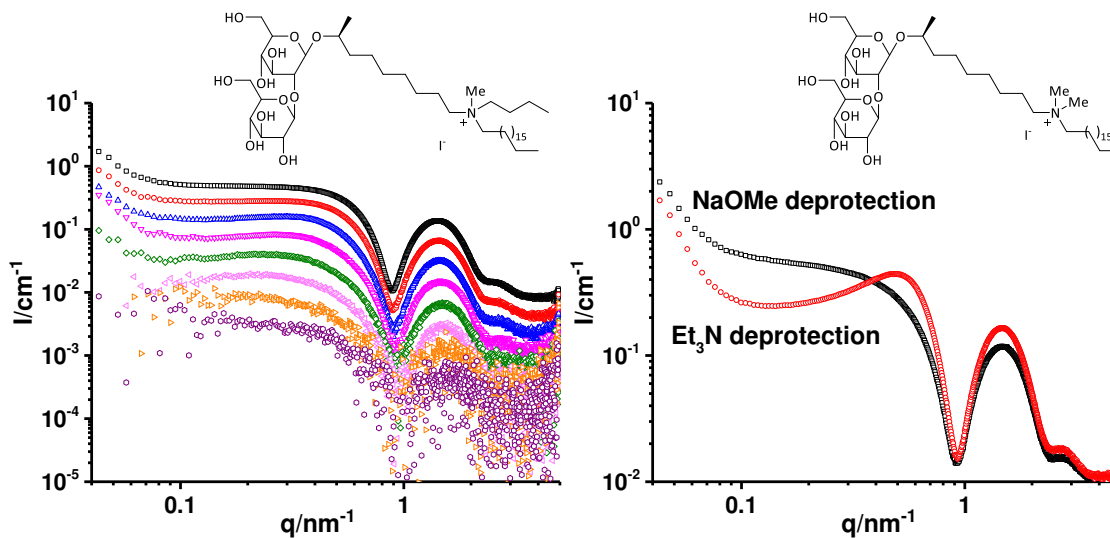


Figure S 2 – Concentration-dependent SAXS profiles for compounds 1a through 1g. For 1f, the typical SAXS profile recorded for  $C = 100$  mg/mL using two deprotecting strategies is also shown. The data show that use of NaOMe flattens the SAXS curve, which is a sign that positive charges of the micelle are screened by the negatively-charged  $^-OMe$  deprotecting group. All SAXS data for all other samples have then been recorded after  $Et_3N$  deprotection.



**Figure S 2 (continued) - Concentration-dependent SAXS profiles for compounds 1g. For 1f, the typical SAXS profile recorded for  $C = 100 \text{ mg/mL}$  using two deprotecting strategies is also shown. The data show that use of NaOMe flattens the SAXS curve, which is a sign that positive charges of the micelle are screened by the negatively-charged  $\text{OMe}^-$  deprotecting group. All SAXS data for all other samples have then been recorded after  $\text{Et}_3\text{N}$  deprotection.**

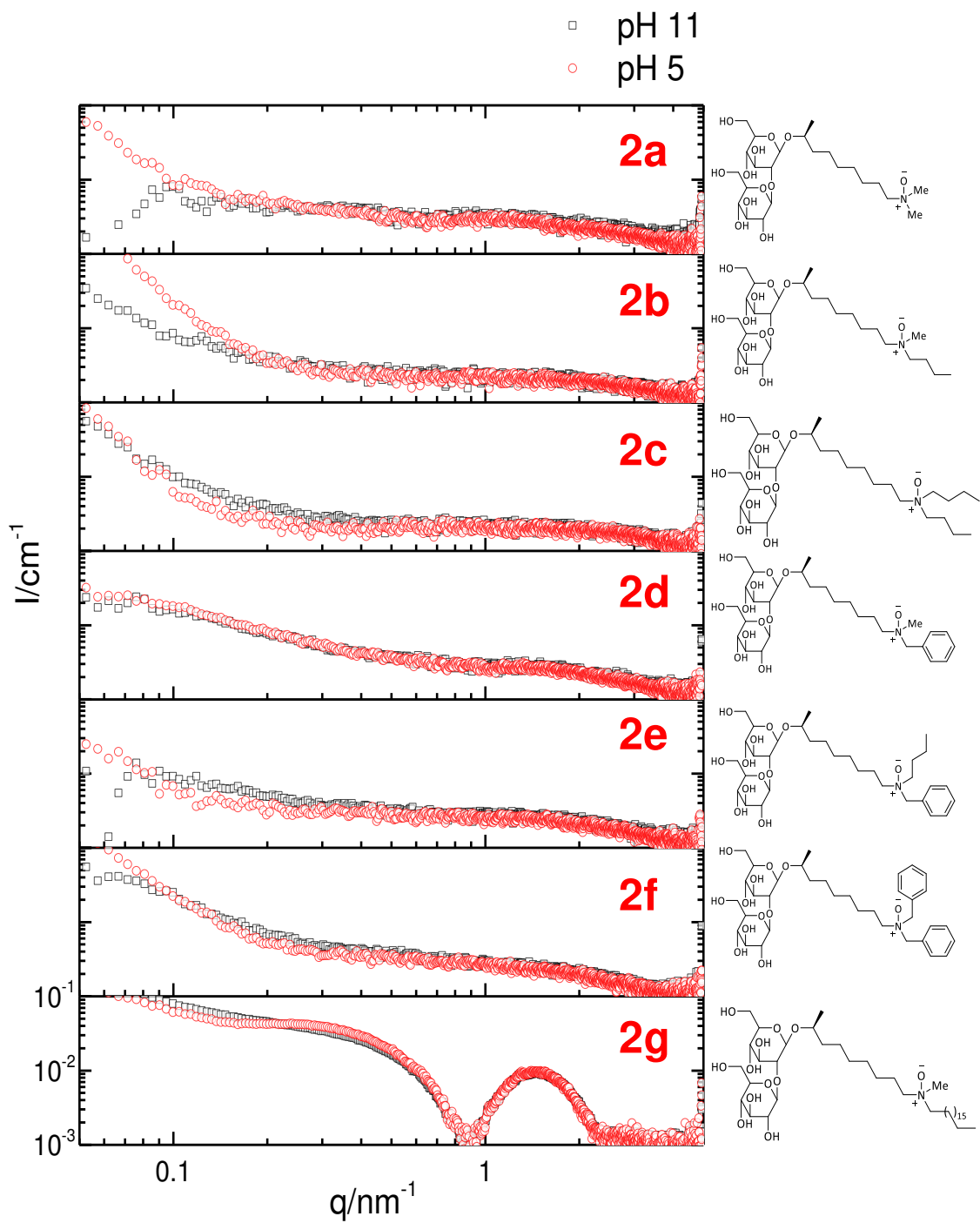


Figure S 3 - Concentration-dependent SAXS profiles for amine oxide sophorolipid derivatives (2a through 2g).

**Table S 2 - Structural parameter values extracted from the SAXS profiles in Figure S 3: radius of gyration,  $R_g$ , is obtained through the Guinier approximation**

<b>Sample</b>	<b><math>R_g/\text{nm}</math></b>
2a	0.4
2b	0.3
2c	0.3
2d	0.4
2e	0.4
2f	0.5
2g	3.9

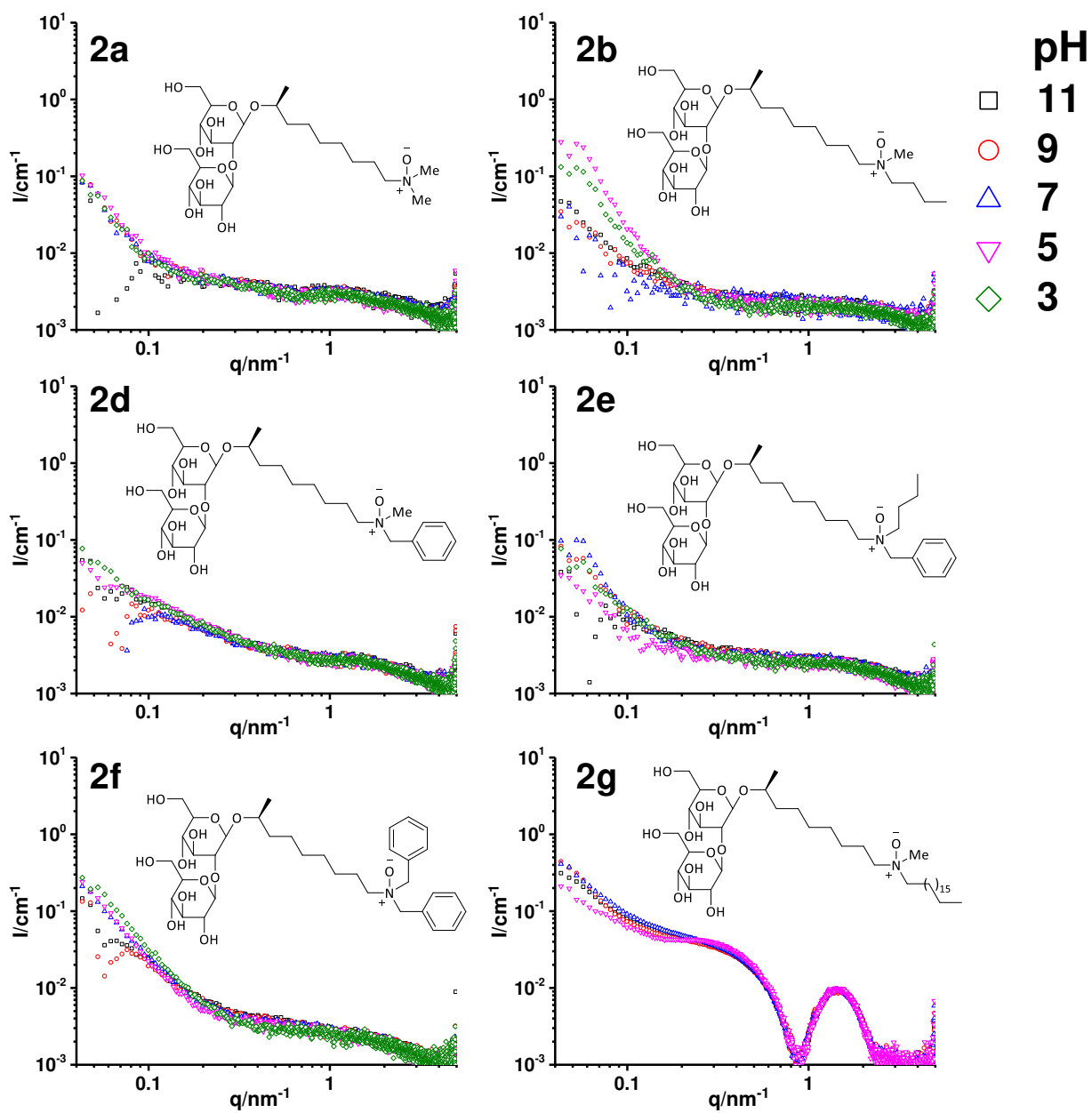


Figure S 4 – pH-dependent SAXS profiles for amine oxide derivatives 2a through 2g recorded at C= 10 mg/mL



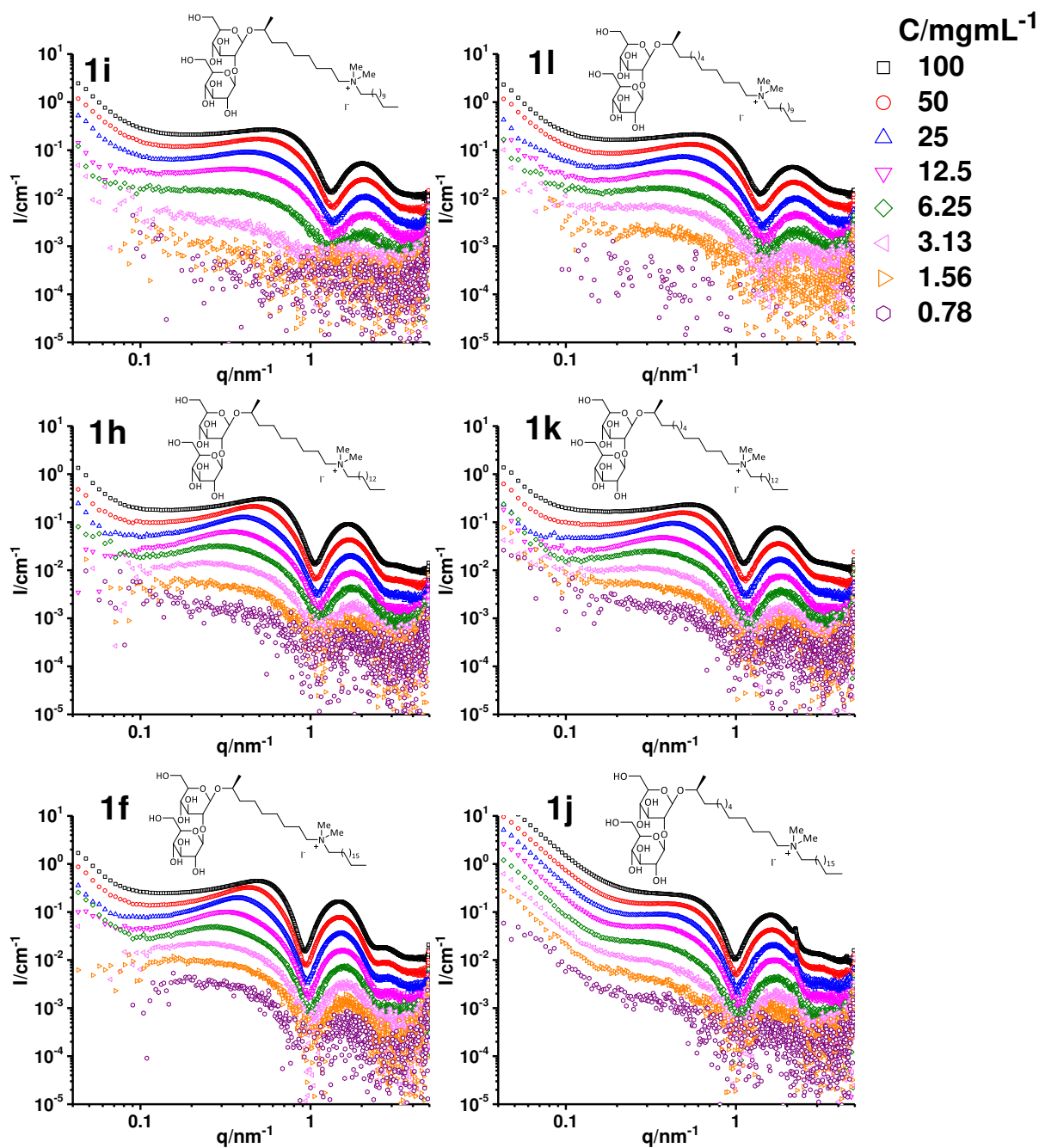


Figure S 5 – Concentration-dependent SAXS profiles for divalent quaternary ammonium sophorolipids 1f and 1h through 1j.

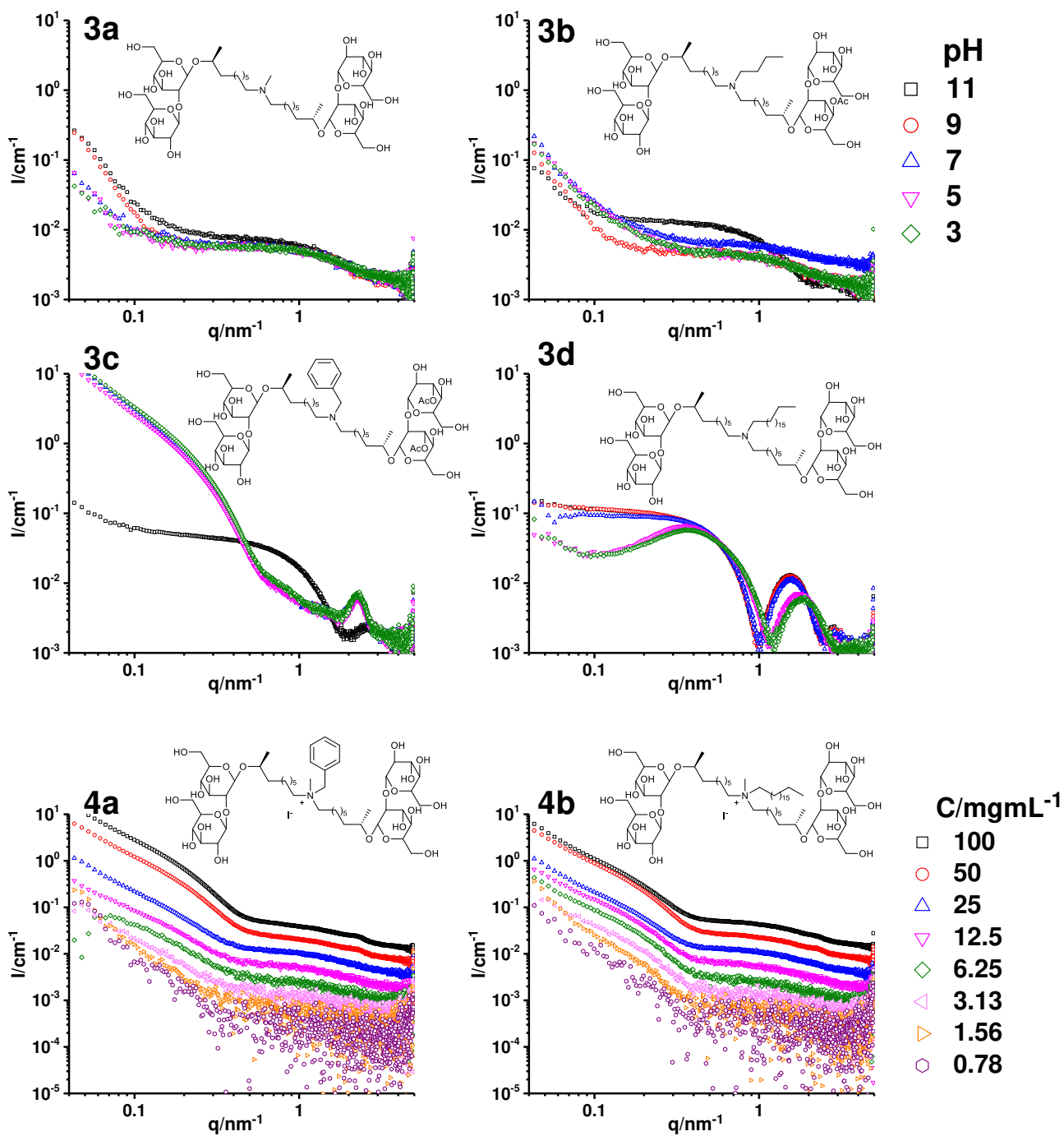


Figure S 6 - pH-dependent SAXS profiles for compounds 3a through 3d ( $C = 10 \text{ mg/mL}$ ) and concentration-dependent SAXS profiles for compounds 4a and 4b

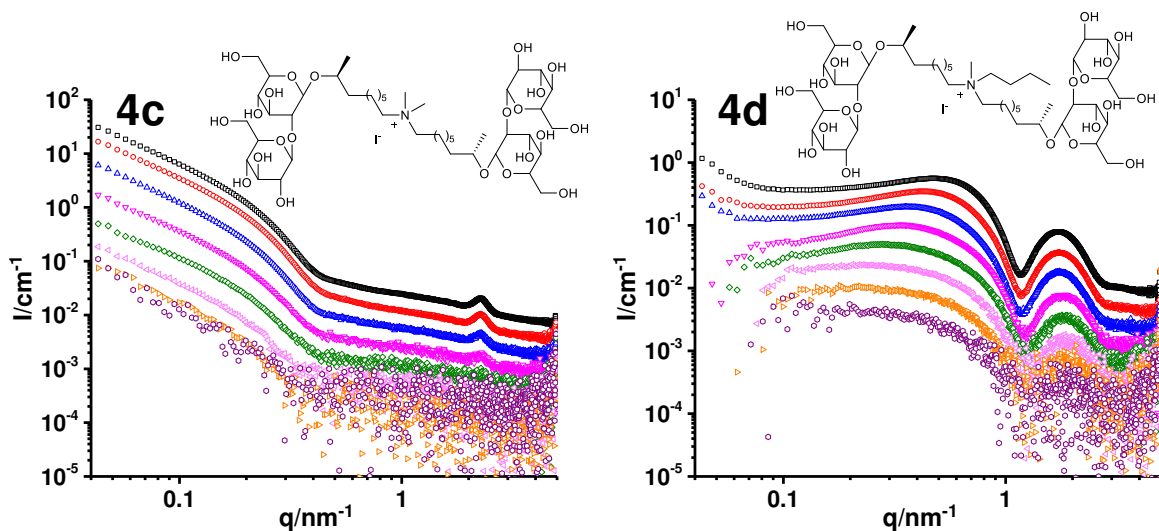


Figure S 6 (continued) - Concentration-dependent SAXS profiles for compounds 4c and 4d

Table S 3 - Structural parameter values extracted from the SAXS profiles in Figure S 6. Values are obtained using a core-shell sphere form factor (no structure factor employed). All parameters are explained in Table 1 and Table 2 in the main text.

Sample	$R_{core}/nm$	$T/nm$	$\rho/10^{-4} nm^{-2}$		$V_m/nm^3$	$N_{agg}$	$N_{agg}$ (max,min)
			$\rho_{shell}$	$\rho_{core}$			
3b (pH11)	0.4	1.4	9.64	8.70	2.63	10	22,0
3c (pH11)	0.6	1.5	9.80	8.70	2.52	16	28,6

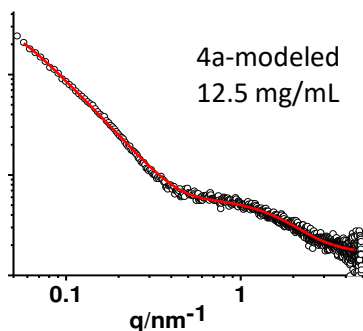


Figure S 7 – Full-scale fit for sample 4a ( $C= 12.5$  mg/mL). The fit has been performed using a model composed of the sum of a cylinder and flexible cylinder form factors. The fit is merely meant to be indicative of the complexity of the signal.

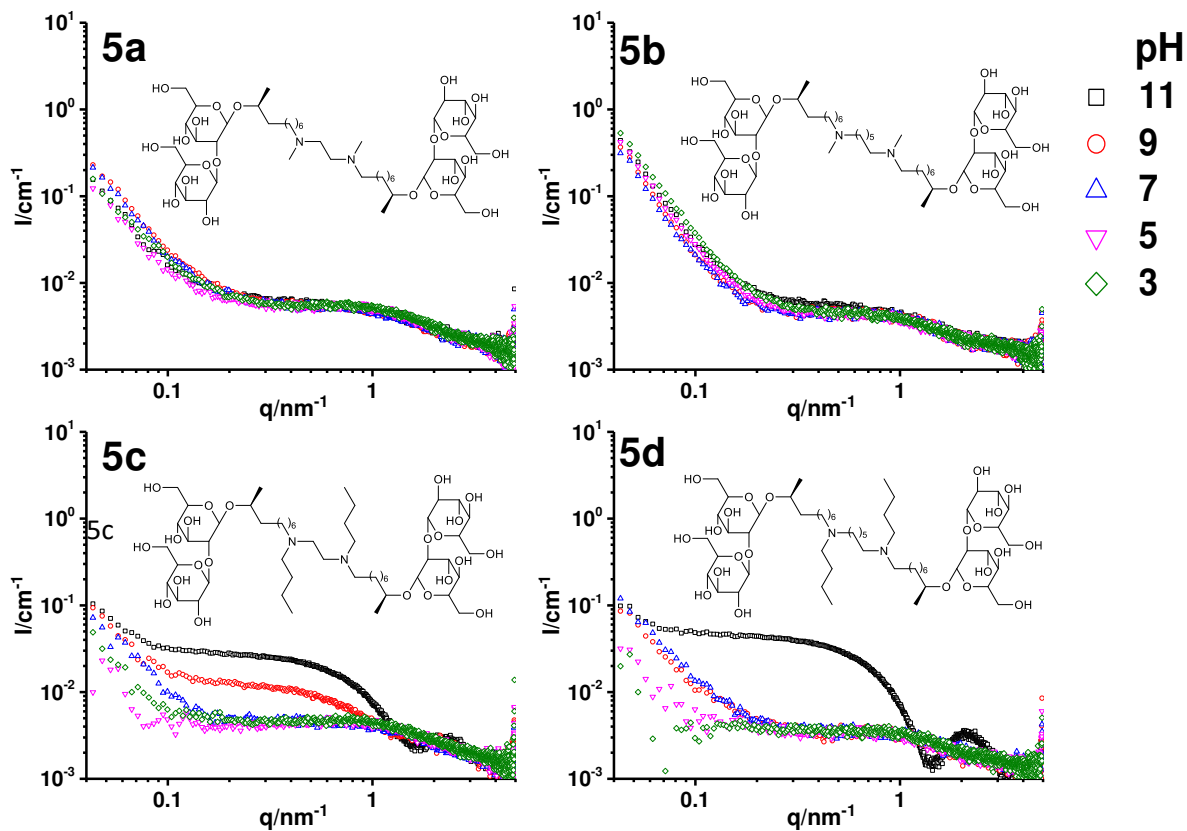


Figure S 8 – pH-dependent SAXS profiles for compounds 5a through 5d ( $C = 10 \text{ mg/mL}$ )

Table S 4 - Structural parameter values extracted from the SAXS profiles in Figure S 8. Values are obtained using a core-shell sphere form factor (no structure factor employed). All parameters are explained in Table 1 and Table 2 in the main text.

Sample	$R_{core}/\text{nm}$	$T/\text{nm}$	$\rho/10^{-4} \text{ nm}^{-2}$		$V_m/\text{nm}^3$	$N_{agg}$	$N_{agg}$ (max,min)
			$\rho_{shell}$	$\rho_{core}$			
5c (pH11)	0.7	1.8	9.63	8.70	2.82	22	34,12
5d (pH11)	0.7	2.0	9.67	8.00	2.92	28	40,18

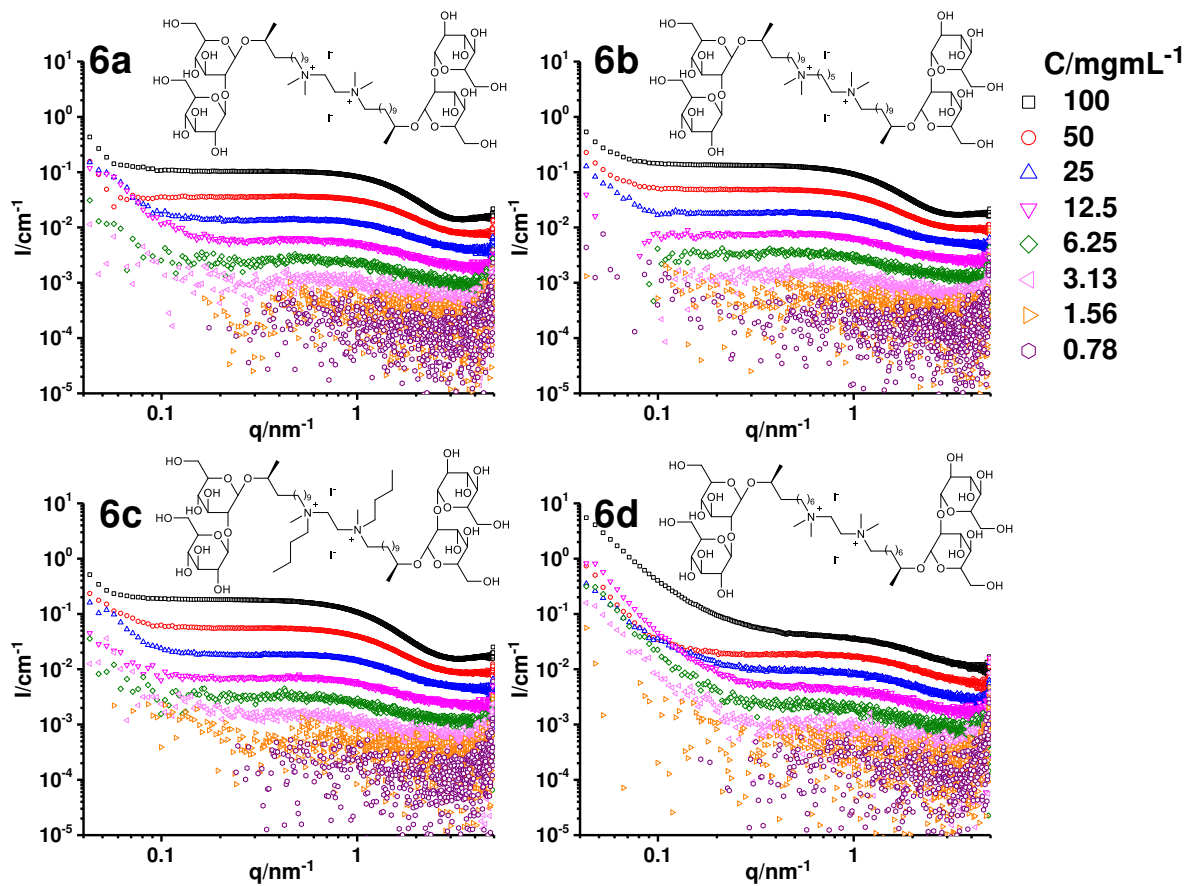


Figure S 9 - Concentration-dependent SAXS profiles for compounds 6a through 6d