



Micelles vs. Ribbons: How Congeners Drive the Self-Assembly of Acidic Sophorolipid Biosurfactants

Prabhu Dhasaiyan, Patrick Le Griel, Sophie Roelants, Emile Redant, Inge N. A. Van Bogaert, Sylvain Prevost, B. L. V. Prasad, Niki Baccile

► To cite this version:

Prabhu Dhasaiyan, Patrick Le Griel, Sophie Roelants, Emile Redant, Inge N. A. Van Bogaert, et al..
Micelles vs. Ribbons: How Congeners Drive the Self-Assembly of Acidic Sophorolipid Biosurfactants.
ChemPhysChem, 2017, 18 (6), pp.643-652 10.1002/cphc.201601323 . hal-01465879

HAL Id: hal-01465879

<https://hal.sorbonne-universite.fr/hal-01465879>

Submitted on 13 Feb 2017

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.

IMPORTANT NOTE : Please be aware that slight modifications occurring after Proof correction may occur between this version of the manuscript and the version on the Publisher's website-----

Micelles vs. ribbons: how congeners drive the self-assembly of acidic sophorolipid biosurfactants

Prabhu Dhasaiyan,¹ Patrick Le Griel,² Sophie Roelants,^{3,4} Emile Redant,³ Inge N. A. Van Bogaert,⁴ Sylvain Prevost,⁵ B. L. V. Prasad,^{*,1} and Niki Baccile^{*,2}

¹ Physical and Materials Chemistry Division, CSIR - National Chemical Laboratory, Dr. Homi Bhabha road, Pune, Maharashtra, India, Fax: (+91) 20-25902636; Tel: +912025902013.
pl.bhagavatula@ncl.res.in

² Sorbonne Universités, UPMC Univ Paris 06, CNRS, Collège de France, Laboratoire de Chimie de la Matière Condensée de Paris, UMR 7574, 4, Place Jussieu, 75005 Paris, France.
Fax: 00 33 1 44 27 15 04, Tel: 00 33 144 27 15 44. niki.baccile@upmc.fr

³ Bio Base Europe Pilot Plant, Rodenhuzekaai 1, 9042 Ghent, Belgium

⁴ InBio, Department of Biochemical and Microbial Technology, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000, Ghent, Belgium

⁵ ESRF - The European Synchrotron, High Brilliance Beam line ID02, 38043 Grenoble, France

Abstract

Sophorolipids (SLs), a class of microbially derived biosurfactant are reported to have different self-assembled structures under the same conditions - either micelles or giant ribbons -by different research groups. Here we explored the reasons for such contradicting results and we attribute these differences to the role of specific congeners which are present in minute quantities. We show that a sample majorly composed of oleic acid (C18:1) sophorolipid in the presence of congeners bearing stearic acid (C18:0) or linoleic acid (C18:2) above as low as 0.5 % results in the formation of micelles that are stable over long periods of time. On the other hand, the presence of only 10-15 % of congeners bearing a stearic acid chain yields fibrillar structures instead of micelles. To study the responsible mechanisms, oleic acid SLs devoid of any other congeners were prepared. Very interestingly, this sample self-assembles

into both micelles and fibers depending on minute modification in self-assembling conditions. The findings are supported by Light Scattering (LS), Small Angle X-ray Scattering (SAXS), Transmission Electron Microscopy under cryogenic conditions (cryo-TEM), High-Pressure Liquid chromatography (HPLC) and Nuclear Magnetic Resonance spectroscopy (NMR).

Introduction

Uprising concerns for global sustainability have promoted a worldwide interest towards the chemical, enzymatic and microbial conversion of natural resources to produce not only fuels, but also chemical compounds and materials. [1] This process has largely increased in the past half century after major political decisions on chemicals regulation in the USA [2] and Europe [3] and since the development of new disciplines such as green chemistry. The use of biobased feedstock to produce chemicals is certainly not a novelty, but in the recent years new trends that try to bypass old paradigms such as development of cheap, straightforward and clean technologies for lignocellulose separation and conversion into high valuable chemicals or large scale production of biodegradable biodetergents have emerged. In this quest for biobased valuable goods, the degree of purity of many biobased chemicals is a major common problem that limits the development of a viable, so-called, biobased economy. Carbohydrates and lignin are known to be the richest sources of chemicals: ethanol, furans, glycerol, lactic acid, succinic acid, levulinic acid can be obtained, among others, from carbohydrates, while phenols, aryl acids, solvents, bio-oil and much more are obtained from lignin. [4,5] If the conversion methods are multifold (hydrolysis, oxidation, pyrolysis, catalysis, enzymatic and even microbial), the selectivity can sometimes be very low. This is the case for some furans, furfural and hydroxymethylfurfural, commonly obtained from 5- and 6-carbon sugars and for which the selectivity is known to be quite low (45% for furfural from xylose using a zirconia catalyst). If recent developments have increased it to 90% using ionic liquids, high selectivity and conversion rate using cheap sustainable approaches are still demanded. Lactic acid is produced from expensive glucose using microbial fermentation. The yield is quite good, 90%, but the process is known to contain impurities needing major efforts (e.g., water splitting electro dialysis) to remove them.[4]

Surfactants are also highly demanded chemicals and they can be obtained directly from plant extracts, e.g., flavonoids from soy by-products, [1] from vegetable oils or from microbial conversion of sugars and vegetable oils. In all cases, biobased surfactants are never constituted of pure products. Rhamnolipids, sophorolipids, cellobioselipids, mannosylerythritol lipids are some of the most studied, and commercially viable, biosurfactants available. In all cases, the crude extract contains between at least two and four major compounds, [6 -12] as this is the case for rhamnolipids, which can contain mono- and di-rhamnose congeners, [6] for sophorolipids, generally composed of the open acidic and closed lactonic congeners, [11,12] or cellobioselipids and mannosylerythritol lipids, which

can contain at least three main congeners each. [9] The same phenomenon was also observed in the field of pharmacy. Chondroitin sulfate is a common impurity associated with heparin, a natural carbohydrate based drug. [13] However, the presence of few abundant congeners, that can efficiently be separated using column chromatography, represents the best case scenario. In reality, real raw mixtures actually contain several minor components varying in the chain length, number of unsaturation in the hydrophobic moiety and degree of acetylation. This has been discussed for rhamnolipids, [14] cellobioselipids, [9] but is most well-known for sophorolipids. [11,12,14] In fact, one can actually consider it to be a general assumption for most substances. Trying to separate single lipid components to obtain an impurity-free compound is technically difficult, often due to the similar polarity that most congeners have in common, and, for this reason, economically not viable for large scale production. Even for fundamental studies, extreme purification becomes a true question, because that can only occur on limited amounts, thus reducing the final yield to a point which is not enough for a complete study of the compound properties. Chen et al. [15] have recently performed a deep physico-chemical study on mono- (R1) and di-rhamnolipids (R2) and their controlled mixture; however, they only separated R1 from R2, the main components of the raw microbial fermentation broth, but not from the minor impurities known to be present in a R1/R2 batch. [14] This case is not isolated. A sophorolipid raw extract contains a complex mixture of acetylated acidic and lactonic congeners and it would be very difficult to efficiently separate individual components. However, a simple hydrolysis step under alkaline conditions, followed by acidification and easy recovery without time-consuming column chromatography [16] can be performed on this particular compound to obtain the acidic congener only at a good degree of purity in a relatively straightforward way. However, even in this favourable case, residual amounts of congeners (di-unsaturated, fully saturated, C16 fatty acids sophorolipids derivatives) [11,12] are not removed.

To which extent is it then necessary to purify biobased chemicals in general, and biosurfactants in particular? This is an old question which has been previously set even for important petrochemical-derived surfactants. For instance, it is well-known that the critical micellar concentration (cmc), [17] minimum of the surface tension, [18,19] phase boundaries [20] and even adsorption behaviour onto surfaces [18] can be influenced by the presence of batch impurities in sodium alkyl sulfate surfactants. If the impact of dodecyl alcohol, a by-product derived from the autocatalytic reaction of sodium dodecyl sulfate (SDS), on cmc and surface tension (a minimum in the surface tension is observed in commercial SDS but no minimum in recrystallized SDS) has been often discussed,[17,18,20] the influence of

significant amounts (28.7 %) of C14 congener has been shown to significantly impact the cmc (by a factor 3); nevertheless, minor amounts of C14 did not show any influence. Considerations on SDS can be translated to biobased products and, given the fact that extreme purification is a tedious and expensive process for both petrochemical amphiphiles and, more specifically, for biobased products because of the broader, natural, chemical dispersity of the latter, we believe that the question above is of paramount importance. However, the answer to it depends on what one is looking for and available data are contrasting. The physico-chemical behaviour of R1 and R2 rhamnolipids at the air-water interface and in water has shown that parameters like the critical micelle concentration (CMC) only varies between 0.2 mM and 0.3 mM at any R1/R2 ratio while the micellar aggregation number only varies between 30 and 50 at any concentration below 60 mM for any R1/R2 ratio. However, above 40 mM and at roughly 50-60% of R1 content, a lamellar $L\alpha$ phase, attributed to R1, is more stable than the micellar $L1$ phase, attributed to R2. The acidic form of sophorolipids has been reported to form both micelles and giant ribbons under similar acidic conditions, a fact which has not been explained so far. In terms of applications, the use of biosurfactants is highly dependent on the control of their solution self-assembly. For detergency, antimicrobial or environmental remediation purposes, which are typical uses for biosurfactants, [14] one would prefer to work with a stable micellar solution, but if one targets biomedical applications like wound healing, one prefers to work with stable bilayers or fibrillar aggregates, instead.

In this work, we want to address the question of tolerable purity for sophorolipids (SLs), an important, scaled-up, biosurfactant with actual use in cosmetics and detergency and potential use as antimicrobial and antitumor agent. [21,22] Considering their potential medical and pharmaceutical use, this question is of utmost importance, as shown by the case of heparin. Similarly, reducing purification steps (and costs) without losing the main physico-chemical properties can be a strong advantage in terms of applications and economic viability in the cleaning and cosmetics fields, as shown by the rhamnolipids example above. Several literature reports since 2004 describe a different self-assembly behavior for exactly the same non-acetylated open acidic form of sophorolipids bearing a C18:1 (oleic acid) fatty acid with a subterminal glycosidic bond (Figure 1). In all cases, the compound is obtained from the classical acidic/lactonic mixture after the typical hydrolysis/acidification step. In 2004, Zhou et al. [23] have reported the first supramolecular self-assembly behavior of acidic sophorolipids. They have observed formation of giant ribbons at acidic pH and micellar structures at higher pH. Several years later, Dhasaiyan et al.²⁴ reported a comparative pH dependant self-assembling behavior of oleic, elaidic and stearic acid SLs. This study also

concluded the formation of ribbon-like structures with oleic acid based sophorolipids, similar to Zhou et al. The same report also demonstrated that the stearic acid derived SL self-assembles into sheet like structures. However, these data were contradicted by several concomitant works by other research teams showing that the C18:1 derivative of acidic sophorolipids form micelles in the entire pH range from acidic to basic, the main difference being in the micellar size and electrostatic charge repulsion. [25-28] Effect of base, time and counterion on the same system was also reported, with a conclusion that, in all cases, micelles are always found and they are stable over time. Similar data were also reported by Penfold et al. [29] It is of course not possible that the same compound behaves differently under the same physico-chemical conditions and we believe that such a difference originates from the presence of minor, although under-evaluated, impurities.

To prove this point, we adopt a practical approach and we study different batches containing a majoritary, but variable, content of the C18:1 form of acidic sophorolipids and prepared by multiple research groups worldwide (CSIR–NCL, Pune, India; LCMCP, UPMC, Paris, France; InBio, Ghent University, Ghent, Belgium). In fact, the representation of acidic sophorolipids (Figure 1), classically found in all works dealing with this molecule, is an approximation with respect to the real composition of this sample. The actual sample is generally naturally polluted by one or more congeners, that include the oleic acid (C18:1) derivative with a terminal glycosidic-oleic acid bond, the stearic (C18:0) and the linoleic (C18:2) acid derivatives along with the C16 derivatives. [11,12] We will then connect the exact molecular composition of each sophorolipid batch sample in our possession to its self-assembly properties in water and compare it with a series of 100% pure batches of the compound given in Figure 1. In particular, we show that ^1H NMR can be qualitatively used to highlight the differences between the samples but High Pressure Liquid Chromatography (HPLC) is necessary to put in evidence the exact composition and relative amounts. By mean of Light Scattering (LS), Small Angle X-ray Scattering (SAXS) and Transmission Electron Microscopy under cryogenic conditions (cryo-TEM) we compare the self-assembled structures of each sample and we show that the same giant ribbons [23,24] are only observed in the sample containing more than 15-20 wt% of the stearic acid sophorolipid derivative, known to form nanoscale ribbon fibers by its own. [30] All other batches, poorer in this congener, form micelles, irrespective of their exact congeners impurities. This hypothesis is validated by a specific study on a 0-to-100 mixture of the stearic acid sophorolipid congener in a micelles-forming SL solution. Finally, we also show how the method of preparation can influence the final structure.

This work clearly shows that among all impurities contaminating a majoritary C18:1 acidic sophorolipids (subterminal glycosidic bond) batch, only one strongly affects its self-assembly and that can occur at relative content as low as 15 w%. We show here that a 100% pure product is not an absolute necessity, however, one must reduce as much as possible that specific impurity that strongly perturbs the overall sample behaviour.

Material and methods

The material and method section is reported in the Supporting Information.

Results and Discussion

SL nomenclature

As biosynthesized sophorolipids are a complex mixture of various compounds, lactonic, acidic and both having various degrees of acetylation. The glycosidic bond generally occurs on the subterminal position of the fatty acid but terminal glycosilation also occurs; the chain length is commonly C18 bearing one single *cis* double bond on C9 (oleic acid) but C16 and presence of two *cis* double bonds may also occur. Considering this broad variety of compounds, the general terms “sophorolipids”, or even “acidic/lactonic sophorolipids” is misleading and imprecise. The hydrolysis of a sophorolipids mixture reduces the number of congeners, as it eliminates all acetylated and lactonic derivatives; in many works, hydrolyzed sophorolipids are broadly referred to “acidic sophorolipids” or “oleic acid sophorolipids (OASL)”, because oleic acid constitutes the main fatty acid backbone. Unfortunately, this nomenclature is not precise enough as it does distinguish between the terminal and subterminal position of glycosilation. In this work, we will use a more specific nomenclature to identify a given congener: aSL-CXX:Y_z, where aSL stands for acidic sophorolipids, CXX:Y, with $16 < XX < 18$, $0 < Y < 2$, refers to the number of carbon atoms (CXX) and unsaturation degree (Y) of the fatty acid, and $z = s$ or t , being the position of the glycosidic bond with respect to the fatty acid, subterminal (s) or terminal (t). The position of the *cis* double bond is set on the 9th carbon atom. For simplicity and unless otherwise stated, we name the sophorolipid mixtures by its major constituent, aSL-C18:1s (Figure 1) and we address to them as “(sample) batches”, in opposition to the fully purified molecules that we could obtain for this specific study, and which will be addressed to as aSL-C18:1s “controls”.

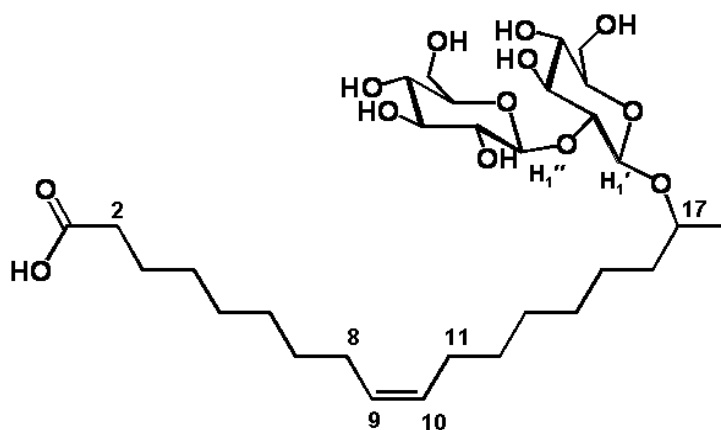


Figure 1 - The chemical structure of acidic sub terminal C18:1 sophorolipid, aSL-C18:1s. This compound constitutes the major congener of a classical acidic sophorolipid batch. This compound may also be referred to as oleic acid sophorolipids (OASL) in the literature. H_1' and H_1'' refer to the anomeric \underline{CH} respectively located on the first and second glucose with respect to the C17.

Compositional sample characterization

We have initiated our investigation by carefully looking at the solution state NMR spectra of aSL-C18:1s batches obtained from different institutions. The samples that are included in this study are tagged as follows: Batch 1 is prepared at CSIR-NCL in Pune, India; Batch 2 is prepared at Inbio at Ghent University, Belgium; Batch 3 and Batch 4 were obtained at UPMC in Paris, France by hydrolyzing and purifying SLs purchased from Soliance, France; Batch 5 is prepared at Bio Base Europe Pilot Plant (BBEPP, Ghent). Batch 5 is then further purified using Flash Chromatography at BBEPP into three additionally highly homogenized samples, used as references in this work and tagged Control 1, Control 2 and Control 3. The typical synthesis scheme valid for all samples is shown in Figure S1 in the ESI. However, specific variations in the synthesis conditions, substrate used, post-synthesis treatment and purification protocols have a major impact on the natural variability and relative congener content in the final samples. From the solution state 1H NMR (solvent MeOD- d_4) data presented in Figure S2a through Figure S2d in the ESI, one can find the typical signatures of aSL-C18:1s backbone, and in particular the $\underline{CH=CH}$ ($H_{9,10}$) at 5.37 ppm, the $\underline{CH_2}$ in α (H_2) to the COOH at 2.28 ppm, the anomeric \underline{CH} of sophorose (H_1' , H_1'') at 4.64, 4.45 ppm, for instance. From a qualitative point of view, all samples present the same spectral signature, which can be safely attributed to the sub-terminal form of aSL-C18:1s. However, Table S1 shows that significant differences in terms of the expected integrated values of selected peaks may occur, and the $H_{9,10}$ is particularly concerned. The NMR analysis brings out the fact that

none of the analyzed samples is 100 % pure, despite the purification process (e.g., column chromatography) operated on some of them (please refer to the materials and methods for more details on sample preparation). The NMR analysis also highlights that all other molecules (impurities, congeners) which are present in minor amounts are hardly detectable by ^1H NMR, as this technique finds its limits both in sensitivity and resolution for analogous compounds. As ^1H NMR is unable to tell the exact composition of the samples and in particular the presence and amount of congeners, all samples are analyzed by HPLC-ELSD, better suited to identify the presence of similar molecules with slight different polarities. This technique, coupled to LC-MS, is typically used to identify biosurfactant congeners. [30,31] The structures of the most common congeners present in the batches studied here are given in ESI Figure S3. Please note that the glycosidic bond for these compounds is given in its subterminal form, but the terminal form is also observed; this is stated by the use of the s or t subscript in the nomenclature. Typical HPLC-ELSD data are given in the ESI (Figure S4), while Figure 2a (numerical data are given in ESI, Table S2) shows the compositions of the samples produced for this study. The relative amount of aSL-C18:1s (Figure 1) represents more than 87.0 % of the total composition for Batch 2 through Batch 5 while it only represents about 77.0 % of Batch 1. The second most abundant compound is generally the terminal congener (aSL-C18:1t), representing 7.0 %, 4.9 %, 6.0 % and 4.4 % respectively in Batch 1, Batch 3, Batch 4 and Batch 5 samples.

aSL-C16:0s, aSL-C16:0 and aSL-C18:2s and aSL-C18:2t congeners are not systematically detected and their relative amount generally represents less than 1 % to 3 % of each sample. One can then safely say that all batches analyzed here contain more than 90 % of the terminal and sub terminal aSL-C18:1z (z= s, t) congeners, except for Batch 1, which only contains about 84.0 %. Interestingly, this batch contains about 13.0 % of the aSL-C18:0s, the fully saturated congener, while this contamination is only 5.2 % in Batch 2 and as low as 1.0 % and 2.1 % in Batch 4 and Batch 5 respectively. Three samples of high degree of purity have also been produced for control purposes only, as the extreme purification to one single congener is a time-consuming expensive process that would not be viable on large scale, unless a very high added value application is linked to it. Control 1 and Control 3 samples (Figure 2c) respectively contain 99.7 % and 96.7 % while Control 2 contains 100 % of aSL-C18:1s. The latter is characterized by one single peak after HPLC-ELSD analysis observed at 12.87 min (Figure S4d, green chromatogram) while Control 1 and Control 2 only have, respectively, 0.3 % of aSL-C18:2t and 3.3 % of aSL-C18:1t congeners, only detectable as small broad shoulders next to the aSL-C18:1s peak (inset in Figure S4d).

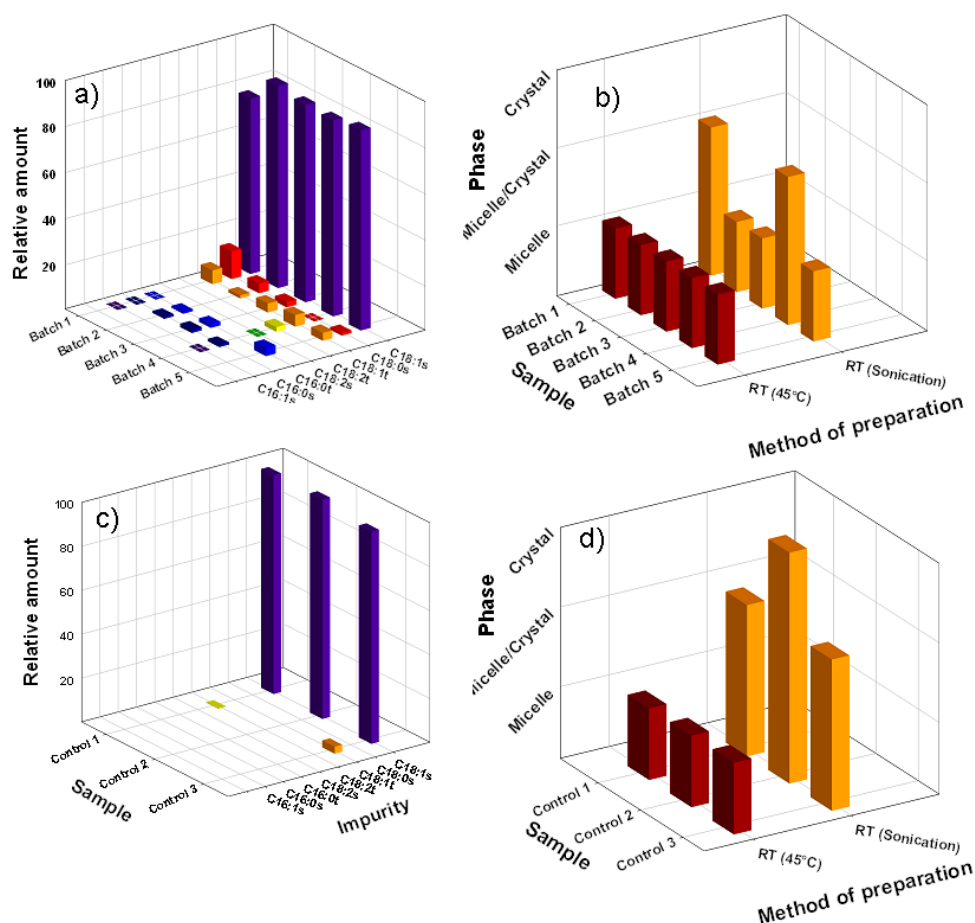


Figure 2 – (a) Distribution of sophorolipid aSL-CXX:Yz congeners ($16 < XX < 18$; $0 < Y < 2$; $z = s, t$) in the sample batches used in this study and (b) their phase behaviour in water (5 mg/mL) at room temperature ($RT = 22.0 \pm 1.0$ °C) after either a preliminary solubilization step at 45 °C or after simple sonication process. (c) aSL-C18:1s controls at high degree of purity and (d) their phase behavior.

Self-assembly behaviour

Once all sophorolipid sample batches are thoroughly characterized, each sample is dispersed in solution following two methods: 1) dispersion of the sample in water at RT followed by sonication; 2) fast sonication/heating of the sample in water at about 45°C for 30 s to 1 min so to obtain a clear solution before keeping the sample at RT for further studies. The pH is acidic, between 4 and 5, and the concentration is generally 5 mg/mL, however, analogous results are obtained in a wider concentration span, from at least 1 mg/mL to 50 mg/mL. Both methods have been chosen in function of previous literature studies. [24,25] The turbidity of the solutions is generally characterized by light scattering (LS) while the structure is analyzed by SAXS (Figure S5). Cryo-TEM and SEM are also employed

complementary to SAXS analysis and shown when necessary. Figure 2b shows the phase behaviour for all sample batches, the exact congener composition of which is given in Figure 2a. If most sample batches directly solubilize at RT (e.g., SAXS pattern of batch 5 sample at RT in Figure S5), for homogeneity purposes in terms of treatment, we present here a protocol including a gentle heating step, which systematically solubilizes all samples, which eventually form clear solutions after cooling at RT. SAXS [27] and SANS [25,29] have long been used to characterize the micellar structure of sophorolipids either in their pure acidic [25,27] or acidic/lactonic mixture.[29] Information on the methodology used to model SAXS/SANS data in sophorolipids-based systems are given in the Supporting Information although details can be found elsewhere [27,29,32] and will not be repeated here. All cited works agree on the fact that sophorolipids solutions are always composed of micelles, which can be approximatively modeled as ellipsoids of revolution with a hydrophilic shell (sophorose + H₂O) and a hydrophobic core (fatty acid), and the radius of which generally falls below 2 nm, thus suggesting that the lipid conformation is in its bent form. The shell thickness and core radius estimated from the SAXS analysis for the batch samples studied in this work are given in Table 1, which also reports the shell scattering length density (SLD). The hydrophobic core radius size is always comprised between 0.75 nm and 0.90 nm while the shell thickness is between 0.9 nm and 1.3 nm; the SLD, which provides information on the hydration of the sophorose shell, varies between 10.7 and 11.0 x 10⁻⁴ nm⁻². These values indicate that sophorose is hydrated as they are comprised between ~14.0 x 10⁻⁴ nm⁻² (sophorose) and 9.4 x 10⁻⁴ nm⁻² (water). To continue, we also report SAXS data obtained on the three pure control references, which also form micelles under the same conditions (Figure 2d and Figure S5). The radius, shell thickness and shell SLD values fall exactly in the same range as for the congener mixtures. Interestingly, the average values for all compounds (Batch 1-5 and Control 1-3) of the core radius is 0.81 ± 0.04 nm, the shell thickness is 1.13 ± 0.10 nm and shell SLD is 10.9 ± 0.2 x 10⁻⁴ nm⁻², thus showing the very strong homogeneity (relative errors comprised between 1 % and 10 %) in terms of equatorial structure of the micelles, despite the very broad impurity distribution in the samples. Finally, the average core aspect ratio of the ellipsoid is 5.0 ± 2.1 among all batch samples. The SAXS modeling results, both in terms of radii and aspect ratio are in agreement with previously published data. [25,27,29] All in all, these data clearly indicate that once sophorolipids are in the micellar form, the structure of the latter (radius, aspect ratio, shell thickness and shell SLD) is practically independent from the sophorolipid batch composition.

Table 1. Size of the micellar core radius (fatty acid), hydrophilic shell (hydrated sophorose) thickness and shell scattering length density (SLD) for all sophorolipid Batch and Control samples analyzed in this work. These values have been estimated by using a core-shell ellipsoid of revolution model to fit the SAXS data presented in Figure S5 in the ESI

	Relative amount SL C18:1s [%]	Core radius [nm]	Shell thickness [nm]	Shell SLD [$\times 10^{-4}$, nm ⁻²]
Batch 1	77.0	0.78	1.24	10.9
Batch 2	88.7	0.86	0.99	11.0
Batch 3	86.6	0.87	0.99	10.7
Batch 4	86.0	0.83	1.15	11.0
Batch 5	87.8	0.79	1.14	10.9
Control 1	99.7	0.79	1.26	10.7
Control 2	100.0	0.85	1.11	11.1
Control 3	96.7	0.75	1.20	10.7
Average	-	0.81 ± 0.04	1.13 ± 0.10	10.9 ± 0.2

If no preliminary solubilization step is adopted and samples are dispersed in solution at RT using a simple sonication process, the congener composition actually plays a strong role in the type of phase observed. LS data qualitatively illustrates this point on a selection of samples presented in ESI Figure S6: all solutions are clear except Batch 1, which is highly turbid, and Batch 4 which is only slightly turbid. The extra pure, Control 1 to Control 3, references are also water-insoluble under these conditions and present a comparable turbidity with Batch 1. Further analysis using SAXS confirms the structural difference with respect to the solubilization protocol. As an example, we report in Figure S5 the SAXS data for Control 2 (100 % pure aSL-C18:1s), Batch 5 (87.8 %) and Batch 1 (77.0 %) recorded at RT after a simple sonication step: Control 2 shows a pattern typical of a crystalline powder (diffraction peak at $q = 2.30 \text{ nm}^{-1}$), Batch 1 shows a mixed signal between a crystalline sample (diffraction peak at $q = 2.30 \text{ nm}^{-1}$) and micelles while Batch 5 shows a typical micellar pattern and no diffraction pattern.

The phase behavior for all samples (batches and pure references) analyzed at RT after a simple sonication step is summarized in Figure 2b, d. In general, if all batch samples practically form micellar solutions (including Batch 4, which presents a minor degree of crystalline inclusion), Batch 1 behaves sensibly differently than all other ones, being a frankly turbid solution. The SAXS signal of this sample shows the presence of a strong diffraction peak at 2.30 nm^{-1} , corresponding to a repeating distance of 2.70 nm, superimposed on the classical micellar signature, and a strong increase in the scattering intensity in the low- q

region. This feature is consistent with the presence of large structures with crystalline molecular order coexisting with micelles. Indeed, the same sample forms a clear solution and its SAXS pattern displays features corresponding to the formation of micelles when heated, as already discussed.

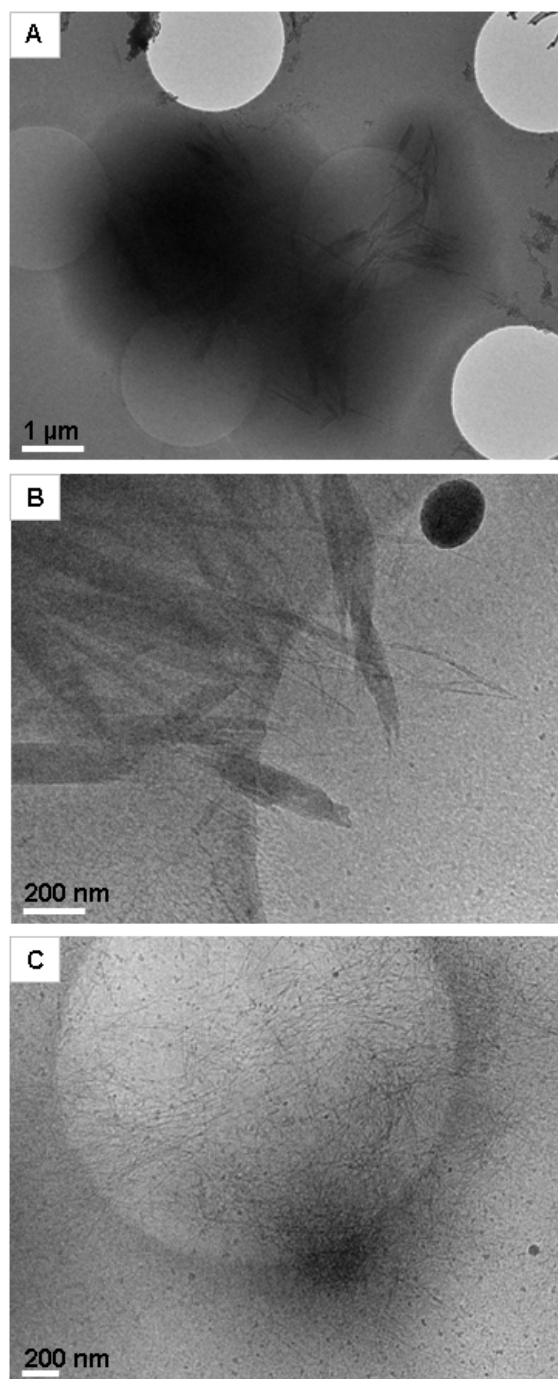


Figure 3. A-C) Series of cryo-TEM images of self-assembled structures obtained from Batch 1.

When Batch 1 turbid sample is imaged using cryo-TEM, aggregates of fibrillar

structures are clearly visible throughout the grid (Figure 3). In particular, the sample is constituted both by large ribbons of about 50 to 100 nm in cross-section (Figure 3 A, B) and thin fibers, the cross-section of which is about 10 nm (Figure 3C). Also, from Figure 3B one can see that the large ribbons are actually constituted by thin ones, which compose the end-tips of the larger structures. The formation of ribbon-like structure from Batch 1 is congruent with previous results published by some of the co-authors of the present communication, [24] even if SEM, a technique which works on a dried sample under vacuum, was used as a characterization technique. By comparing the SAXS results for this sample (Figure S5) and the cryo-TEM images, one can suppose that the diffraction peak at 2.30 nm^{-1} (2.70 nm) can be attributed to the intra-ribbon sophorolipid periodic arrangement. Interestingly, a very similar value (2.65 nm) was found for the inter-lipid distance measured in nanoscale ribbons formed from the saturated derivative of sophorolipids, aSL-C18:0s. [30] The pure Control 1- 3 samples seem to have a similar behaviour as Batch 1: they are not soluble in water at RT, unless gentle heating is applied, and the corresponding SAXS profile is also characterized by a strong diffraction peak at 2.30 nm^{-1} . The strong structural similarity is qualitatively confirmed by the SEM image obtained on Control 2 (Figure S7a) and showing the omnipresence of large platelets/fibers.

The data presented above ratify the fact that different batches of aSL-C18:1s self-assemble into at least two stable structures, micelles and fibers, even though the gross chemical composition is very similar. We also wish to state that Scanning Electron Microscopy (SEM) analysis confirmed that the solid state appearance of the as-prepared powdery sample batches from different sources before mixing with water is reasonably identical to one another in their morphological features and that no specific morphology can be identified (typical SEM images for selected samples are given in Figure S8 in the ESI). This poses a pivotal question: why do different batches of the same sample form different structures in solution and how do they relate to the behaviour of pure aSL-C18:1s?

Importance of secondary congeners, pH and time

While Batch 1 behaves more closely to the pure sophorolipid, it actually has the lowest aSL-C18:1s content among all samples (only 77 %), making the understanding of its behavior particularly puzzling. However, one should note from Figure 2a, the highest contamination of saturated sophorolipids (aSL-C18:0s) as main secondary congener (13 %), whereas all other sample batches only have a content of less than 5 % of aSL-C18:0s. We then make the hypothesis that the presence of moderate amounts of the saturated form of SLs

has a strong impact on the different self-assembly behavior of Batch 1. In particular, it was reported that aSL-C18:0s congener forms sheets [28] and nanoscale ribbons [30] at acidic pH according to the method employed to prepare the sample. In fact, the reported cryo-TEM structures of aSL-C18:0s [28,30] closely resemble the ones found in this study for Batch 1. Interestingly, one must also note that even if the previous work on the formation of giant sophorolipid ribbons [23] did not provide the exact composition of the sample used, the same authors have already shown the presence of non-negligible amounts (not quantified) of aSL-C18:0s congener in their works. [33] We believe that this could explain the giant ribbon formation in ref. 23 just as it suggests it for Batch 1.

The previous discussion poses the hypothesis that relatively small amounts ($> 15\%$) of the aSL-C18:0s congener are able to impact the self-assembly behaviour of a sample containing a majority of aSL-C18:1s sophorolipid. We have verified this hypothesis by controlling the mixture of aSL-C18:1s and aSL-C18:0s congeners (the synthesis and characterization of the latter was reported before. [30] To do so, we have used a stable micellar solution batch of SL, containing 88.7 % of aSL-C18:1s (Batch 2) at RT and mixed it with increasing amounts of aSL-C18:0s. The combined LS and cryo-TEM data presented in Figure S9 (commented in the ESI) show that minor amounts (at least 15/20 wt %) of aSL-C18:0s as congener in a majority of micelles-forming sample are responsible for the formation of a fibrous ribbon-based structure. If, at a first glance, the fact that minor amounts of aSL-C18:0s congener have such a strong influence on the collective behaviour of a micelles-forming batch, one could actually explain it from the fact that the pure form of aSL-C18:1s is able to form fibers having the same crystalline packing. As a general rule, it then seems that aSL-C18:1s sophorolipids have two stable phases, a micellar one and a fibrous crystalline one, and that the chemical structure of the congener impurity drives the self-assembly process towards one or the other phase. At the moment, we have found that a slight excess in aSL-C18:0s, a compound which forms twisted ribbons (in the neutral/acidic pH region) by itself, helps stabilizing the fiber structure of aSL-C18:1s. On the contrary, relative excess in other congeners (aSL-C18:1t, aSL-C18:2t, aSL-C18:2s, etc) seems to stabilize the micellar phase, even if, unfortunately, no data exist at the moment on the self-assembly properties of these molecules.

Next, we address one last question concerning the stability of the micelles and fibrous phases against time and pH. All micellar batch solutions formed at RT (45 °C) shown in Figure 2b are stable over time at RT: the solution never becomes turbid. The only exception is constituted by Batch 1, which undergoes a clear-to-turbid transition after about 30 minutes (at

RT and at rest), the final phase being the above-described fibrillar one. As far as Control 1-3 samples are concerned, we must highlight that Control 2 (100 % aSL-C18:1s) has a similar behaviour: the precipitation of a platelet/fibrous phase from the micellar solution occurs after about 20-30 min at RT at rest (Figure S7b). Interestingly, Control 1 and Control 3, which only differ from Control 2 for the presence of very low amounts of, respectively, aSL-C18:2t (< 0.5 %) and aSL-C18:1t (< 3.5 %) congeners, have a different kinetic behavior, as the stability of their micellar phase is much longer (days). Without going into deeper details, the qualitative analysis of the self-assembly behavior of Control 1 and Control 3 shows that even very low amounts (< 0.5 %) of a given congener are able to influence the formation of the crystalline fibrous phase. Finally, one can outline the role of temperature: the fibers found in those samples containing a majority of the monounsaturated form of SL (Batch 1 or Control 2) are crystalline and stable up to $45 \pm 5^\circ\text{C}$; it is also worth saying that in no case one can form fibers (e.g., by a heating/cooling step) in those batches, of which the micellar form is the most stable one. The complexity of aSL-C18:1s self-assembly as a function of pH, temperature, time and impurities seems to be analogous to what is known for other surfactants, as recalled in the introduction to this work, [17-20] and in particular on the impact of dodecyl alcohol impurities on SDS or the impact of the R1/R2 ratio on the CMC, area-per-molecule, L1/L α phase, in rhamnolipids. [15] Finally, the pH seems to play a similar role to what has been already described. Upon pH increase up to 10-11, all solutions, irrespective of their composition, are clear and stable over time, assuming the formation of a micellar phase. Upon pH decrease (as low as pH 2.5), both the micellar (Batch 2-5) and crystalline fiber (Batch 1, Control 1-3) structures do not undergo remarkable phase changes. More details on the pH-dependent self-assembly of sophorolipids are given elsewhere. [32] However, we have noticed that the kinetic of fiber formation is faster at low pH, as reported before. [23]

Possible explanations for this uncanny behavior could come from a fine thermodynamics analysis of the self-assembly, and a measure of the free energies of the micellar and crystalline phases, as routinely done for protein systems. [34] In a second hypothesis, one should also consider the possibility of bistability, a phenomenon which has been specifically described for natural glycolipids like gangliosides. Bistability refers to the different phase behaviour of a lipid in relationship with the dual conformation possibilities of their hydrophilic headgroups (e.g., parallel vs. perpendicular with respect to the membrane plane), and for which glycosidic groups are particularly affected, in particular under a temperature stimulus. [35] Although a deep understanding of the dual behaviour of pure sophorolipids is out of the scope of this work, one can easily formulate the following

hypothesis and have it tested using solution NMR, a classical technique which has been previously employed to evaluate the carbohydrate conformational dynamics of ganglioside GD1a micelles, [36] in combination with solid state NMR, used in the past to study the distribution of aSL-C18:0s molecules in twisted ribbons. [30] If relevant conformational changes occur in the sophorose headgroup, or elsewhere in the molecule, during the time-dependent micelles-to-fibers phase change, it could be possible to detect variations in the characteristics chemical shifts of the sophorolipid. We have then prepared a clear micellar solution containing Control 2 sophorolipid at 5 mg/mL in D₂O, the temporal behaviour of which was immediately analyzed at rest using solution ¹H NMR. Figure S10a shows the first (t= 0) and last (t= 1h30) spectra of the time-resolved series, whereas at t= 0 the solution is clear while at t= 1h30 the solution is highly turbid, a typical sign of massive fiber formation. While no specific shift could be observed for any of the sample resonances (3-(trimethylsilyl)propionic-2,2,3,3-d₄ acid sodium salt, TMSPd₄, was used as internal reference), one must nevertheless highlight the time-dependent loss in intensity for each peak, as shown in Figure S10b for few selected resonances. However, if one considers that solution NMR is sensitive to micelles but not to the fiber phase, these data seem to suggest that the amount of micelles-forming sophorolipids exponentially decays with time in favour of fiber formation; However, since no chemical shift is observed, one can also deduce that the micelle-to-fiber transition does not occur via a conformational change of neither sophorose nor oleic acids moieties. In other words, these data seem to tell that micelles loose matter over time but their structure does not change, a fact which recalls the role of reservoir role of micelles in the pH-induced micelle-to-fiber transition observed in the aSL-C18:0s system. [32]

Nonetheless, one could argue that chemical shifts are not sensitive enough to eventual conformational variations in sophorolipids. To dissipate any doubt, we have collected the Control 2 sample fibers from the colloidal solution soon after the time-resolved experiment. The fibers were centrifuged within a 2.5 mm zirconia rotor to which a small amount of TMSPd₄ was added as internal reference. The system was then studied with solid state ¹H under both static and MAS, as this technique is also sensitive to soft solid samples. Figure S10c compares the solution (black curve) and solid (blue curve) state ¹H NMR spectra, respectively representing the micellar and fiber phases. As one can see, variations in the chemical shift occur for all resonances except for the TMSPd₄ 0 ppm reference. The solution and solid state chemical shift values and their differences are reported in Table S3. The value of Δ , the difference between the chemical shifts in liquid and solid state, is centered in the

range between 0.02 and 0.04, although Δ as high as 0.10 can be measured for the 2.31 ppm resonance, which corresponds to the CH_2 group in position 2 with respect to the COOH, thus translating the important changes in the hydrogen bonding environment around the COOH group between the micelle and the fiber phases. Without getting into more details, these experiments show that chemical shift analysis is a valuable tool to have qualitative pieces of information on the local molecular rearrangement between the micellar and fiber phase and in this sense, one cannot exclude a conformational difference of the sophorolipid molecule between the micelle and the fiber. Nevertheless, a gradual chemical shift variation could not be observed during the micelle-to-fiber phase transition using solution NMR; this fact prevents a clear-cut answer on the actual nature and localization of the driving force that pulls sophorolipid molecules away from the micelles and drag them into crystalline fibers.

All in all, this communication settles the end of a debate in the literature concerning the effective self-assembly properties of acidic aSL-C18:1s sophorolipids in water in the neutral/acidic pH region. We show here that both the fibrous [23,24] and micellar [25,27,28] phases are possible to obtain and we give the conditions to obtain them. A classical acidic sophorolipid mixture obtained from either oleic acid or rapeseed oil using a standard microbial fermentation and subsequent purification process [16] will form a micellar solution in water. Micelles can form directly at RT and they always do form after gentle fast heating. The micellar structure and behaviour against pH has been described before. [25,27,28] The micellar size and structure is independent of the congener relative composition and the micelles are highly stable in time. Nevertheless, if the purification (or fermentation) process is responsible for an unwanted enrichment in the aSL-C18:0s sophorolipid congener (above 15 %-20 %), the fibrillar form will be most likely stabilized, as discussed before. [23,24] We then show here that an end-user of aSL-C18:1s sophorolipids will now be able to select the type of phase (micelles, fibers) by playing with the congener relative amount. We also show that it is not necessary to strive for a 100 % pure aSL-C18:1s sample as this one behaves very similarly to a micellar-forming batch enriched with aSL-C18:0s. All these options are summarized in Figure 4: a aSL-C18:1s sophorolipid batch will always contain a mixture of congeners. If an excess of the saturated aSL-C18:0s occurs, fibers will be formed. If the congener content is more heterogeneous (generally the terminal aSL-C18:1t is majoritarian), micelles will be formed. However, if to such a heterogeneous mixture one adds extra saturated aSL-C18:0s sophorolipids, fibers will be formed again. Finally, extreme purification will lead to a water-insoluble compound which can form, at RT, both micelles (kinetic phase) and fibers (thermodynamic phase) according to the method of preparation.

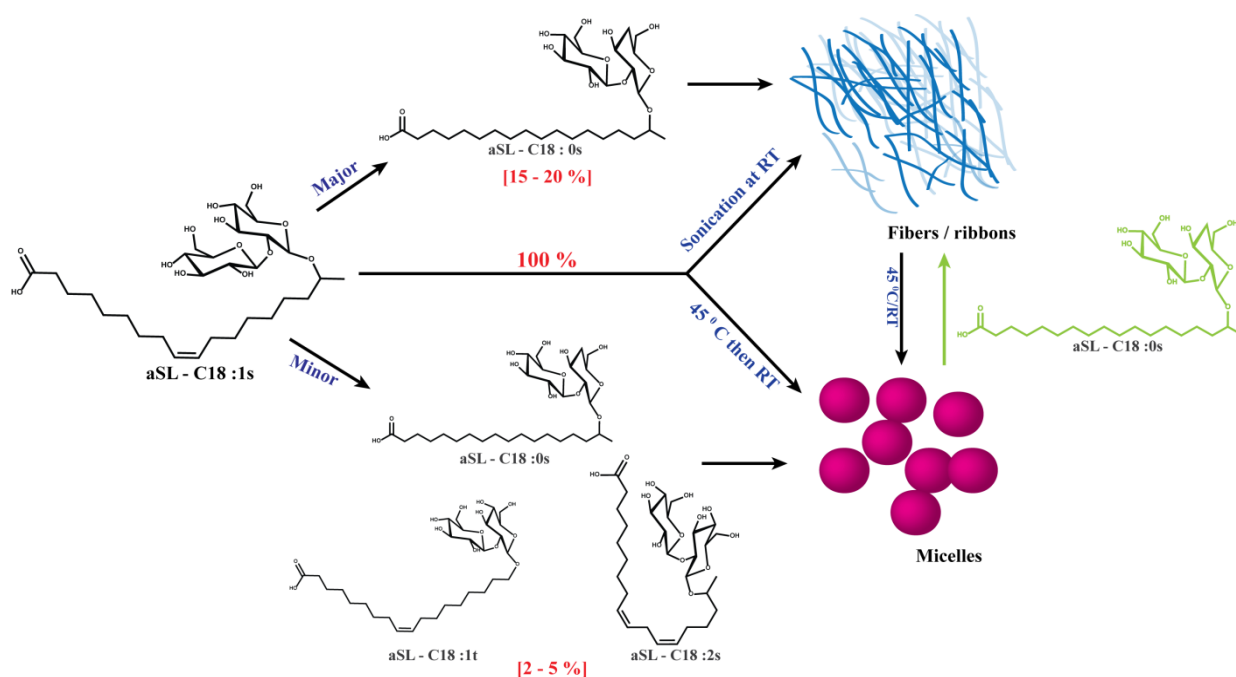


Figure 4. Overview of the general self-assembly pathway of a SL-C18:1s sophorolipids as a function of the batch composition. Percentages are given in wt%

Conclusions

In conclusion, we investigated the role of congeners in the formation of self-assembled structures of a bio-derived surfactant namely acidic sub terminal C18 sophorolipids (aSL-C18:1s) by cryo-TEM, NMR, DLS, HPLC-ELSD and SAXS techniques. We were able to recover various samples from different research teams worldwide and a set of highly pure and enriched aSL-C18:1s. The results indicate that, all the systems studied exhibit self-assembled structures in solution. For all samples having a aSL-C18:1s majority ($> 85\%$ based on HPLC-ELSD peak ratio) and variable fractions of other, classical, congeners (aSL-C18:1t, aSL-C18:0s, aSL-C18:2s, etc.) a stable micellar solution is always formed at room temperature, despite the preparation method. Micelles can also be obtained at RT starting from the 100% pure aSL-C18:1s sample by solubilizing ($45\text{ }^{\circ}\text{C}$, 1 min) in water. Micelles can be modelled with core-shell ellipsoids of revolutions and their average radius and thickness are, respectively, $0.83 \pm 0.05\text{ nm}$ and $1.12 \pm 0.11\text{ nm}$. Considering the low relative error (less than 10%), one can safely say that the micellar structure is not influenced by the congener mixture. On the other hand, we find that if the sophorolipid batch contains more than 15-20 wt% of the stearic acid sophorolipid aSL-C18:0s derivative, crystalline fibers and ribbon structures are majoritarian, whereas micelles can still coexist. The study of a set of 100% pure aSL-C18:1s

references show that this compound is not water-soluble and it forms crystalline fibers at RT upon simple sonication. However, as mentioned, a stable micellar solution can be obtained if the compound is solubilized under gentle heating. We also find that the stability of the micellar solution at RT strongly varies with the congener content. At 100% aSL-C18:1s and in the presence of 15-20 wt% aSL-C18:0s, we evaluated a stability of about 30 min, before fiber formation occurs. On the contrary, as soon as minor amounts of congeners (less than 0.5 %) other than aSL-C18:0s are present, the micellar solution displays an “infinite” stability. Our results assume great significance if we look into the molecular self-assemblies as seen in microbial biosurfactants. By this work we both clarify a controversial point in the literature concerning the self-assembly of aSL-C18:1s and we give the limit of tolerance beyond which a typical acidic sophorolipid batch sample should be considered as a mixture and not as pure sample. In addition, we also provide a practical guide to those who want to work with one of the most common biosurfactants without lengthy and expensive purification processes: the conditions for micellar and fiber formations are now established.

Associated content

This Supporting Information is available free of charge via the Internet at <http://pubs.acs.org>.”

Acknowledgments

This work was supported by Raman-Charpak Fellowship Scheme to P. D and the authors thank CEFIPRA, DST - India and Campus France for funding the same. P. D thanks UGC - New Delhi (Govt. of India) for Senior Research Fellowship. The SAXS experiments were performed on beam line ID02 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. P. D and B. L. V. P are grateful to Dr. A. A. Prabhune and Dr. C. V. Ramana (CSIR-NCL, Pune) for the SLs production and purification. Dr. G. Laurent and Dr. F. Ribot (LCMCP, UPMC, Paris, France) are kindly acknowledged for their assistance in NMR experiments.

Reference

-
- [1] R. Montgomery, *Biores. Technol.* **2004**, *91*, 1–29
 - [2] Toxic Substances Control Act (TSCA), passed in 1976 and recently amended by the Toxic Substances Control Act (TSCA) Modernization Act (<https://www.congress.gov/bill/114th-congress/house-bill/2576/text>)
 - [3] <http://echa.europa.eu/regulations/reach>; <http://www.unido.org/index.php?id=5279>;

- [4] J. J. Bozell, G. R. Petersen, *Green Chem.* **2010**, *12*, 539–554
- [5] S. Laurichesse, L. Avérous, *Progr. Polym. Sci.* **2014**, *39*, 1266–1290
- [6] S. Lang, D. Wullbrandt, *Appl. Microbiol. Biotechnol.* **1999**, *51*, 22–32
- [7] S. Itoh, H. Honda, F. Tomita, T. Suzuki, *J. Antibiot.* **1971**, *24*, 855–859
- [8] C. Sylđatk, S. Lang, V. Matulovic, F. Z. Wagner, *Naturforsch.* **1985**, *40c*, 61–67
- [9] S. Spoeckner, V. Wray, M. Nimtz, S. Lang, *Appl. Microbiol. Biotechnol.* **1999**, *51*, 33–39
- [10] A. M. Abdel-Mawgoud, F. Lépine, E. Déziel, *Appl. Microbiol. Biotechnol.* **2010**, *86*, 1323–1336
- [11] I. N. A. Van Bogaert, J. Zhang, W. Soetaert, *Process. Biochem.* **2011**, *46*, 821 – 833
- [12] H. J. Asmer, S. Lang, F. Wagner, V. Wray, *J. Am. Oil Chem. Soc.* **1988**, *65*, 1460 - 1466
- [13] J. F. K. Lintiac, C.J. Jones, C. K. Larive, *Anal. Chem.* **2009**, *81*, 10116–10123
- [14] R. Marchant, I. M. Banat, *Trends Biotechnol.* **2012**, *30*, 558–565
- [15] M. L. Chen, J. Penfold, R. K. Thomas, T. J. P. Smyth, A. Perfumo, R. Marchant, I. M. Banat, P. Stevenson, A. Parry, I. Tucker, I. Grillo, *Langmuir* **2010**, *26*, 18281–18292
- [16] N. Baccile, A.-S. Cuvier, C. Valotteau, I. N. A. Van Bogaert, *Eur. J. Lipid Sci. Technol.* **2013**, *115*, 1404–1412
- [17] A. Smith, *J. Colloid Interf. Sci.* **1978**, *66*, 575–576
- [18] J. E. Gebhardt, D. W. Fuerstenau, *J. Colloid Interf. Sci.* **1984**, *101*, 278–279
- [19] K. J. Mysels, *Langmuir* **1986**, *2*, 423–428
- [20] P. Kékicheff, C. Grabielle-Madellmont, M. Ollivont, *J. Colloid Interf. Sci.* **1989**, 112–132
- [21] D. W. G. Develter, L. M. L. Lauryssen, *Europ. J. Lipid Sci. Technol.* **2010**, *112*, 628–63
- [22] I. N. Van Bogaer, K. Saerens, C. De Muynck, D. Develter, W. Soetaert, E. J. Vandamme, *Appl. Microbiol. Biotechnol.* **2007**, *76*, 23–34
- [23] S. Zhou, C. Xu, J. Wang, W. Gao, R. Akhverdiyeva, V. Shah, R. Gross, *Langmuir* **2004**, *20*, 7926 - 7932;
- [24] P. Dhasaiyan, A. Banerjee, N. Visaveliya, B. L. V. Prasad, *Chem. Asian J.* **2013**, *8*, 369 - 372
- [25] N. Baccile, F. Babonneau, J. Jestin, G. P. Arnaudet, I. N. A. Van Bogaert, *ACS Nano* **2012**, *6*, 4763 - 4776
- [26] N. Baccile, J. K. Pederson, G. P. Arnaudet, I. N. A. Van Bogaert, *Soft Matter* **2013**, *9*, 4911 - 4922
- [27] S. Manet, A. S. Cuvier, C. Valotteau, G. C. Fadda, J. Perez, E. Karakas, S. Abel, N. Baccile, *J. Phys. Chem. B* **2015**, *119*, 13113 - 13133
- [28] A. S. Cuvier, F. Babonneau, J. Berton, C. V. Stevens, G. C. Fadda, G. P. Arnaudet, P. Le Griel, S. Prevost, J. Perez, N. Baccile, *Chem. Eur. J.* **2015**, *21*, 19265 - 19277
- [29] J. Penfold, M. Chen, M., R. K. Thomas, C. Dong, T. J. P. Smyth, A. Perfumo, R. Marchant, I. M. Banat, P. Stevenson, A. Parry, I. Tucker, I. Grillo, *Langmuir* **2011**, *27*, 8867–8877
- [30] A. S. Cuvier, J. Berton, C. V. Stevens, G. C. Fadda, F. Babonneau, I. N. A. Van Bogaert, W. Soetaert, G. P. Arnaudet, N. Baccile, *Soft Matter* **2014**, *10*, 3950 - 3959
- [31] I. N. A. Van Bogaert, J. Sabirova, D. Develter, W. Soetaert, E. J. Vandamme, *FEMS Yeast Res.* **2009**, *9*, 610 - 617
- [32] N. Baccile, A. S. Cuvier, S. Prevost, C. V. Stevens, E. Delbeke, J. Berton, W. Soetaert, I. N. A. Van Bogaert, S. Roelants, *Langmuir* **2016**, *32*, 6343

-
- [33] W. Gao, R. Hagver, V. Shah, W. Xie, R. A. Gross, M. F. Ilker, C. Bell, K. A. Burke, E. B. Coughlin, *Macromolecules* **2007**, *40*, 145 - 147
- [34] P. G. Vekilov, S. Chung, K. Olafson, Shape change in crystallization of biological macromolecules, *MRS Bulletin* **2016**, *41*, 375-380
- [35] A. Raudino, L. Cantù, M. Corti, E. Del Favero, *Curr. Op. Colloid Interf. Sci.* **2000**, *5*, 13-18
- [36] L. Poppe, H. van Halbeek, D. Acquotti, S. Sonnino, *Biophys. J.* **1994**, *66*, 1642-1652