Structural characterization of novel sophorolipid biosurfactants from a newly identified species of Candida yeast

Neil P. J. Price a,⁎, Karen J. Ray a, Karl E. Vermillion b, Christopher A. Dunlap c, Cletus P. Kurtzman d

aUSDA-ARS-NCUAR, Renewable Product Technology, 1815 North University Street, Peoria, IL 61604, USA
bUSDA-ARS-NCAUR, Functional Foods, 1815 North University Street, Peoria, IL 61604, USA
cUSDA-ARS-NCAUR, Crop Bioprotection, 1815 North University Street, Peoria, IL 61604, USA
dUSDA-ARS-NCUAR, Bacterial Foodborne Pathogens & Mycology Research Units, 1815 North University Street, Peoria, IL 61604, USA

⁎ Corresponding author. Tel.: +1 309 681 6246; fax: +1 309 681 6040.
E-mail address: neil.price@ars.usda.gov (N.P.J. Price).

A R T I C L E   I N F O

Article history:
Received 12 June 2011
Received in revised form 14 July 2011
Accepted 18 July 2011
Available online 24 July 2011

Keywords:
Sophorolipids
Biosurfactants
Mass spectrometry
MALDI-TOFMS
NMR
Starmerella

A B S T R A C T

Sophorolipids are a group of O-acylsophorose-based biosurfactants produced by several yeasts of the Starmerella clade. The known sophorolipids are typically partially acetylated 2-O-β-D-glucopyranosyl-α-glucopyranose (sophorose) 0-β-glycosidically linked to 17-α-hydroxy-A9-octadecenoic acid, where the acyl carboxyl group often forms a 4'-lactone to the terminal glucosyl residue. In a recent MALDI-TOFMS-based screen for sophorolipid-producing yeasts we identified a new species, Candida sp. NRRL Y-27208, that produces significant amounts of novel sophorolipids. This paper describes the structural characterization of these new compounds, using carbohydrate and lipid analysis, mass spectrometry, and NMR spectroscopy. Unlike those reported previously, the NRRL Y-27208 sophorolipids contain an α-hydroxy-linked acyl group (typically 18-hydroxy-A9-octadecenoate), and occur predominantly in a non-lactone, anionic form. In addition, 17 dimeric and trimeric sophoroses were identified by MALDI-TOFMS from this strain. The surfactant-like properties of these sophorolipids have value as potential replacements for petroleum-based detergents and emulsifiers.

Published by Elsevier Ltd.

1. Introduction

Biosurfactants are microbially produced, renewable, surface-active compounds that are increasingly being used for oil and mineral recovery, and in detergents, cosmetics, and anti-microbial formulations.1–4 The annual worldwide production of surfactants is currently about 10 million tons, and most of which are derived from petroleum feedstocks, but production from renewable sources is now of great interest.4 Included in the glycolipid-type biosurfactants are sophorolipids from various Starmerella yeast species,5–8 which consist of the disaccharide sophorose linked to a long-chain hydroxy fatty acid. Among yeasts of the Starmerella clade that have been examined, the greatest yield of sophorolipids has been reported from Candida apicola and Starmerella bombicola.9 These sophorolipids are a partially acetylated 2-O-β-D-glucopyranosyl-α-glucopyranose unit attached β-glycosidically to 17-α-hydroxyoctadecanoic or 17-α-hydroxy-A9-octadecenoic acid10,11 and can be acylated on the 6'- and/or 6'-positions. The hydroxy fatty acid is generally 16 or 18 carbon atoms, and may contain one or more unsaturated bonds.12,13 The fatty acid carboxyl group is either free (acidic or open form) or internally esterified at the 4'-position (lactone form). The 1,4'-lactone form of the sophorolipids are nonionic surfactants with a critical micelle concentration (CMC) of 40–100 μg ml−1 and are reported to lower the surface tension of water by 30–40 mN m−1, making them especially useful as emulsifiers for oil/water mixtures.9

The S. bombicola sophorolipids are nearly identical to those of C. apicola.14 Chen et al. observed more than six sophorolipids from Wickerhamiella domiciae, and identified one as 17-α-(oxy)-octadecanoic acid 1,4'-lactone 6,6'-diacetate,15 which is also the major sophorolipid from C. apicola and Candida bombicola. Tulloch et al. also discovered similar sophorolipids from C. bogoriensis16 that were shown to differ in the hydroxy fatty acid moiety, which in this instance is 13-hydroxydocosanoic acid.17 We have recently analyzed 26 strains representing 19 species of the yeast genus Starmerella for sophorolipid production using a MALDI-TOFMS based screen initially developed to identify rhamnolipids.18,19 Five of the 19 species tested showed significant production of sophorolipids: C. apicola, S. bombicola, Candida riodocensis, Candida stellata, and a newly identified species of Candida, NRRL Y-27208. As assessed from the MALDI-TOFMS data, these five strains showed a clear structural diversity for the sophorolipids produced.17 The most heterogeneity was observed for C. apicola, which produced di-O-acetyl, mono-O-acetyl, and non-acetyl sophorolipids in the free-acid and lactone forms. S. bombicola produced a major
di-O-acetyl lactone form of sophorolipid, plus a minor component of this as the free acid form. By contrast, C. stellata, Candida sp. NRRL Y-27208, and C. riodocensis produced very little of this lactone form, and the major sophorolipid for these three species is a di-O-acetyl free-acid form, plus smaller amounts of mono-O-acetyl and non-acetyl sophorolipids.

In the present paper we present evidence for the structure of novel sophorolipids from the newly identified yeast, Candida sp. NRRL Y-27208, using a combination of mass spectrometry, nuclear magnetic resonance spectroscopy, GC–MS, HPLC, and classical carbohydrate analysis techniques. From these analyses the complete structure and surfactant properties of these novel sophorolipids are reported, together with the identification of 17 newly identified polymeric sophorolipids.

2. Results and discussion

2.1. MALDI-TOFMS analysis

The sophorolipids were isolated from S. bombicola NRRL Y-17069 and Candida sp. NRRL Y-27208 grown in liquid culture on glucose plus oleic acid, as described previously. The MALDI-TOFMS-based screen identified m/z 711 as the major [M+Na]+ ion arising from the S. bombicola sophorolipids, whereas those from Candida sp. NRRL Y-27208 gave ions at m/z 729, 687, 669, and 645. Noticeably, the S. bombicola major sophorolipid was absent from the NRRL Y-27208 extract, although the 18 Da mass difference between m/z 711 and m/z 729 indicated that these ions are due to the lactone and free carboxylic acid forms of the sophorolipids, respectively. Hence, these major sophorolipid species were assigned as the lactone and free acid forms of a di-O-acetyl sophorolipid, 6,6'-di-O-acetyl-β-D-glucopyranosyl-2′-O-β-D-glucopyranosyloxy-octadecenoic acid. Smaller ions from the NRRL Y-27208 sophorolipids, at m/z 687 and 645 were attributed to mono-O-acetyl and non-O-acetyl analogs of this major free acid sophorolipid, with a minor ion at m/z 669 indicating the mono-O-acetyl sophorolipid in a lactone form.

Growth of these Candida yeasts on different fatty acid substrates resulted in the formation of various sophorolipid structures, as determined by MALDI-TOFMS (Fig. 1). Hence, S. bombicola grown on oleic acid produced a 6,6'-di-O-acetyl-β-D-Glc-p-2′-O-β-D-Glc-p-oxy-oleate (oxy-octadecenoate) lactone-type sophorolipid that is characterized by m/z 711 ions, plus a smaller ion (m/z 729) due to the free-acid form (Fig. 1 A). These ions were also seen when S. bombicola was cultured on stearic acid, but in this case more major ions were observed that are two mass units larger at m/z 713 and 731 (Fig. 1 B). These are attributed to the formation of 6,6'-di-O-acetyl-β-D-Glc-p-2′-O-β-D-Glc-p-oxy-stearate-type sophorolipids in the corresponding lactone (m/z 713) or free-acid (m/z 731) forms.
rate (oxy-octadecanoate)-type sophorolipids, plus smaller amounts of the previously observed 6,6’-di-O-acetyl-β-D-Glc-p-2'-O-β-D-Glcp-oxy-oleate (oxy-octadecanoate)-type. In contrast, Candida sp. NRRL Y-27208 grown on oleic acid produced only 6,6’-di-O-acetyl-β-D-Glc-p-2'-O-β-D-Glcp-oxy-oleate-type sophorolipid in the free-acid form, with no evidence of lactone formation (Fig. 1C). This was also the major sophorolipid produced when this strain was grown on stearate, plus a minor amount of 6,6’-di-O-acetyl-β-D-Glc-p-2'-O-β-D-Glcp-oxy-stearate, which is also in the free-acid form (m/z 731, Fig. 1D).

2.2. Lipid and sugar compositional analysis by GC–MS

To confirm the MALDI-TOFMS assignments and provide additional structural data the sophorolipid extracts were hydrolyzed and analyzed by GC–MS for composition analysis. Analysis of fatty acid methyl esters by GC–MS revealed that the S. bombicola and Candida sp. strains produce sophorolipids that contain different acyl groups. When S. bombicola was grown on oleic acid (octadecenoic) the major sophorolipid fatty acid observed is 17-hydroxyoleic fatty acid (17-hydroxyoleate, tR 15.1 min Fig. 2, panel A). The methyl 17-hydroxyoleate gave a GC–EIMS molecular ion at m/z 312 and (M–18) at m/z 294, plus fragment ions characteristic of (n–1) hydroxy fatty acids. This is also consistent with the oleate-type sophorolipids ions m/z 711 and 729 observed by MALDI-TOFMS; Fig. 1, panel A). When grown on stearic acid (octadecanoic acid) S. bombicola produced sophorolipids that contain predominantly 17-hydroxy-octadecanoate (17-hydroxystearic, tR = 15.5 min Fig. 2, panel B). The 17-hydroxystearic methyl ester gave no observable molecular ion, but was identified by a strong (M–44) fragment ion (m/z 270) arising from characteristic cleavage alpha to the 17-hydroxy group. A smaller GC peak (Rf = 15.1 min) was attributed to 17-hydroxy-octadecenoic (17-hydroxyoleic acid), and arises from the 6,6’-di-O-acetyl-β-D-Glc-p-2'-O-β-D-Glcp-oxy-oleate-type sophorolipids, as also observed by MALDI-TOFMS as m/z 711 or m/z 729 ions (Fig. 2, panel B).

Sophorolipids from the newly identified Candida sp. NRRL Y-27208 did not contain 17-hydroxy fatty acid, but rather a new GC–MS peak was observed at tR 16.2 min, plus a smaller peak at 16.5 min (Fig. 2, panel C). The larger peak (peak 3) was characterized as 18-hydroxyoleate by EIMS ions at m/z 312 (M+) and m/z 294 (M–18). In addition, (M–CH2OH) was observed at m/z 280 and (M–CH2OH–H2O) at m/z 262 (Fig. 2, panel D). The smaller GC peak (peak 4) gave rise to M+ at m/z 340, (M–18) at m/z 322, (M–CH2OH) at m/z 308, and (M–CH2OH–H2O) at m/z 290 (data not shown), suggesting that a minor amount of 20-hydroxy-C20:1 fatty acid is also present. Taken together with the MALDI-TOFMS and NMR data, the fatty acid analysis supports the conclusion that S. bombicola produces sophorolipids that contain 17-hydroxyoleic or 17-hydroxystearic acyl groups, the majority of which are in a 1,4-lactone formation, whereas Candida sp. NRRL Y-27208 sophorolipids contain α-hydroxyacetyl group, predominantly 18-hydroxyoleic, in the open-chain, free-acid form.

Acid hydrolyzed or saponified samples were also derivatized to form volatile aldononitrile peracetates (PAANs) suitable for GC–MS analysis of monosaccharides. For the sugar derivatives, a single GC peak was observed that co-eluted with a standard of α-glucose PAAN. The MS spectrum arising from this peak showed fragment ions that are characteristic of hexoses and were identical to the standard α-glucose PAAN. The sugar composition of the sophorolipid head-groups was consistent whether the producing yeast was grown on glucose, xylose, mannose, or galactose as the carbohydrate source (data not shown). Hence, sophorolipids with alternative sugar head groups were not observed.

2.3. Sophorolipid surfactant properties

Surface tension properties of the sophorolipids were determined using the pendant drop method (Fig. 3). The critical micelle concentrations (CMCs) for the sophorolipids differed by strain, and hence by the nature of the acyl chain and whether it is in the open-chain or lactone form. S. bombicola sophorolipids 6,6’-di-O-acetyl-β-D-Glc-p-2'-O-β-D-Glcp-17-hydroxystearate 1,4-lactone

---

**Figure 2.** GC–MS analysis of hydroxyl fatty acids from Candida sophorolipids grown on different substrates. (A) S. bombicola grown on oleate; (B) S. bombicola grown on stearate; (C) Candida sp. grown on oleate. The acid-hydrolyzed fatty acids were analyzed as their methyl esters. (1) 17-Hydroxyoleate, tR 15.2 min; (2) 17-hydroxyoleate, tR 15.5 min; (3) 18-hydroxyoleate, tR 16.3 min; (4) 20-hydroxy-C20:1, tR 16.6 min. Panel D shows a detail of the EIMS spectrum of GC peak 3, and it is discussed in the main text: tR = retention time.

**Figure 3.** Surface tension properties as determined using the pendant drop method. The equilibrium surface tension was measured 150 s after drop formation. The critical micelle concentrations (CMCs) for the sophorolipids are S. bombicola 17-hydroxystearate-type sophorolipid in the 1,4-lactone form (SL1OL) = 6.9 mg L–1; S. bombicola 17-hydroxyoleate-type sophorolipid in the 1,4-lactone form (SL1ST) = 5.6 mg L–1; and new Candida sp.-type 18-hydroxystearate-type sophorolipid in open-chain, anionic form (SL2BOL) = 46.4 mg L–1.
and 6′,6″-di-O-acetyl-β-α-Glc–2′-O-β-α-Glc–17-hydroxyoleate 1,4′-lactone have measured CMCs of 5.6 mg/L and 6.9 mg/L, respectively. The novel anionic, open-chain sophorolipid 6′,6″-di-O-acetyl-β-α-Glc–2′-O-β-α-Glc–18-hydroxyoleic from Candida sp. NRRL Y-27208 has a CMC of 46.4 mg/L (Fig. 3). These properties are consistent with those found previously for sophorolipids that had been chemically modified to open the lactone ring.21

2.4. Naturally occurring dimeric and trimeric sophorolipids

Examination of the higher mass range in the MALDI-TOFMS spectra led to the identification of several previously undescribed polymeric sophorolipids (Figs. 4–6). In the mass range of 950–1200 Da, five dimeric sophorolipids (A–E) were detected. These were present in varying amounts in the five Starmerella yeast strains tested (Fig. 5). Sodium adduct [M+Na]+ ions observed at m/z 969.47 (A), 1011.48 (B), 1053.48 (C), 1095.49 (D), and 1137.50 (E) are attributable to dimeric sophorolipids, in which the free carboxylic acid of one sophorose monomer is ester linked to a second monomer (see structures in Fig. 4). The 42 Da mass differences between these ions indicated that they are due to non-O-acetyl, mono-O-acetyl, di-O-acetyl, tri-O-acetyl, and tetra-O-acetyl dimers. S. bombicola produced conspicuously less of these dimeric sophorolipids, relative to the other four strains analyzed, and only dimers B, D, and E were detected (Fig. 5). The new Candida species, NRRL Y-27208, produced greater quantities, predominantly mono-O-acetyl dimer B and tri-O-acetyl dimer D, as well as several larger polymeric compounds (described below). The MALDI-TOFMS dimer profile for C. stellata and C. apicola were similar to that of Candida sp. NRRL Y-27208, but these strains did not produce the larger sophorolipid polymers. C. riodocensis also produced five dimers, although in this case the major compounds detected were the tetra-O-acetyl dimer E and the tri-O-acetyl dimer D (Fig. 5). Additional MALDI-TOFMS ions were detected for dimeric sophorolipid (assigned as F–H) in the mass range of m/z 1300–1450 (Fig. 6, panel 1). These masses correlate with sophorolipid dimers with two acyl chains, and the 42 Da mass difference between them indicating di-O-acetyl (dimer F), tri-O-acetyl (dimer G), and tetra-O-acetyl (dimer H) species. Unlike the smaller, monoacylated dimers, these diacylated dimeric sophorolipids were only produced by the new species Candida sp. NRRL Y-27208, and comparable polymeric compounds were not seen from the S. bombicola strain or with the other yeast species included in our initial screening.17 In addition, the new strain Candida sp. NRRL Y-27208 produced nine trimeric sophorolipids, five (I–M) in the mass range of 1600–1850 Da, and four (N–Q) in the range of 1950–2150 Da (Fig. 6, panels 2 and 3). Sophorolipid trimers I–J correspond to the masses of three sophorose units linked by two acyl groups, and are either di- (I), tri- (J), tetra- (K), penta- (L), or hexa-O-acetyl (M). The larger ions, m/z 1980.16, 2022.14, 2064.12, 2106.20 (Fig. 6, panel 3) correspond to trimeric sophorolipids linked by three acyl chains, and were designated as the tri- (N), tetra- (O), penta- (P), and hexa-O-acetyl (Q) species. As with the diacylated dimers (sophorolipids F–H), all of these trimeric sophorolipids were only detected as produced by the new Candida sp. NRRL Y-27208.

2.5. 1D and 2D NMR assignments

The NMR spectra of the sophorolipids are dominated by essentially three spin systems: (1) the 2-O-linked β-α-glucosyl residue (numbered C-1′ to C-6′); (2) the acylhydroxy-linked β-α-glucosyl residue (numbered C-1′ to C-6′); and, (3) the hydroxylacyl group (numbered C-1 to C-18). The numbering system follows that of de Koster et al.22 NMR assignments have been reported previously for sophorolipids from C. bombicola,22 Candida (Torulopsis) apicola,21 Candida (Torulopsis) bombicola ATCC 22214,23 and for lipase-modified sophorolipids from Candida (Torulopsis) bombicola,24 and our current findings confirm these general structural assignments. Hence, NMR signals due to the glucosyl anomeric nuclei are evident for the S. bombicola sophorolipid H-1′/C-1′ at 4.46 ppm/102.2 ppm and H-1′/C-1′ at 4.63 ppm/104.0 ppm, and for the Candida sp. NRRL Y-27208 sophorolipid H-1′/C-1′ at 4.46 ppm/101.1 ppm and H-1′/C-1′ at 4.57 ppm/104.2 ppm (Table 1). The fatty acid 9,10-double bond is also evident for both sophorolipids at 5.35 ppm/129.3 ppm and 5.36 ppm/129.5 ppm (Table 1). In this paper we will, therefore, focus on the structural difference between the sophorolipids from S. bombicola NRRL Y-17069 and those of the new strain Candida sp. NRRL Y-27208 (Figs. 7 and 8).

For the S. bombicola sophorolipids the anomic H-1′ signal at 4.63 ppm shows a HMBC long range coupling across the

---

Figure 4. The structure of dimeric and trimeric sophorolipids. Seventeen polymeric sophorolipids (A–Q) were detected by MALDI-TOFMS, as shown in Figures 5 and 6. The structures shown are A–E, mono-acyl-disophorose; F–H, di-acyl-disophorose; I–M, di-acyl-trisophorose; N–Q, tri-acyl-trisophorose. –OR = O-acetyl groups. The calculated [M+Na]+ masses (in Daltons) are A, 969.47; B, 1011.48; C, 1053.48; D, 1095.49; E, 1137.50; F, 1333.73; G, 1375.73; H, 1417.74; I, 1657.85; J, 1699.86; K, 1741.86; L, 1783.86; M, 1825.87; N, 1980.10; O, 2022.10; P, 2064.11; Q, 2106.11. Actual masses are shown in Figures 5 and 6.
"anomeric oxygen to the C-2′ carbon (C-2′, 82.3 ppm) on the other glucosyl ring (Fig. 7, panel B). This, and the chemical shift differences between the H-2′/C-2′ nuclei (3.42 ppm/82.3 ppm) and the H-2′/C-2′ nuclei (3.33 ppm/75.3 ppm) are supportive evidence for the O-(1→2)-β-glycoside linkage between the glucosyl rings. The other glucosyl anomeric signal, H-1, also shows a long range coupling, in this case to the methine C-17 acyl carbon at 78.7 ppm (Fig. 7, panel B). An HSQC experiment revealed that this C-17 carbon is coupled to the adjacent H-17 proton at 3.76 ppm (Fig. 7, panel A). This proton signal is noticeably shifted by the attached 17-hydroxy group, which is O-glycosidically linked to the Glc-1 ring. Hence, it is observed as an extra CHOH methine signal in the HSQC and DEPT spectra (Fig. 7, panels A and D, respectively). Noticeably, the corresponding H-17/C-17 for the new Candida sp. sophorolipid gave rise to NMR signals at 1.61 ppm/29.3 ppm, which are characteristic of a non-hydroxylated methylene group.

The terminal acyl methyl group of the S. bombicola sophorolipids is assigned by H-18/C-18 NMR signals at 1.22 ppm/20.2 ppm (Table 1). The H-18 proton showed a small hertz coupling to the previously assigned C-17 'extra' CHOH, resulting in a doublet at 1.22 ppm (Fig. 7, panel C). This coupling was also evident in the TOCSY and COSY correlation spectra (data not shown). Noticeably, as discussed below, this evidence for a terminal methyl doublet was absent for the Candida sp. NRRL Y-27208 sophorolipids (Fig. 8, panel C).

A DEPT spectrum of the S. bombicola sophorolipids revealed two methylene carbons at 62.4 ppm and 63.4 ppm which are assigned, respectively, as the glucosyl C-6′ and C-6″ signals (Fig. 7, panel D). The HSQC showed that one of these (C-6′) was associated with

Figure 5. MALDI TOF/MS spectra of dimeric sophorolipids produced by five species of the Starmerella yeast clade: (1) S. bombicola NRRL Y-17069, (2) C. stellata NRRL Y-1446, (3) Candida sp. NRRL Y-27208, (4) C. riocencis NRRL Y-27859, and (5) C. apicola NRRL Y-2481. The observed [M+Na]+ ions (A–E) are due to sophorolipid dimers in which two sophorose units are linked via a hydroxyleate group. The mass difference of 42 Da indicates non-O-acetyl (A), mono-O-acetyl (B), di-O-acetyl (C), tri-O-acetyl (D), and tetra-O-acetyl (E) structures. The structures are shown in Figure 4.
non-equivalent proton signals at 4.11–4.15 ppm and at 3.68–3.85 ppm, while the other (C-6) was coupled to overlapping proton signals (Fig. 7, panel A). These have previously been assigned, respectively, to the lactone and non-lactone (anionic) forms for the *S. bombicola* sophorolipids. This was also evident from the H-4 signals at 4.92 ppm (lactone form) and 3.32 ppm (anionic form) (Table 1). The lactone H-4 signal at 4.92 ppm was evident as a triplet, and also showed a long-range coupling to carbon C-5 at 72.0 ppm (Fig. 7, panel B). The corresponding non-lactone C-5 signal was observed at 76.5 ppm (Table 1).

An equivalent DEPT spectrum for the *Candida* sp. NRRL Y-27208 sophorolipid also revealed two methylene carbons in the 63–64 ppm region (Fig. 8, panels A and D), which are assigned as C-6 and C-6’ (Table 1). These are coupled to overlapping proton signals at 4.21 ppm and at 4.49 ppm (Fig. 8, panel A). However, for the NRRL Y-27208 sophorolipid an additional methylene group is also evident in the DEPT spectrum at 69.5 ppm (Fig. 8, panel D). This acyl C-18 carbon is coupled to non-equivalent H-18 protons at 3.58 ppm and 3.82 ppm (Fig. 8, panel A), both of which also show long-range HMBC couplings to the anomeric C-1’ (101.6 ppm) for this molecule (Fig. 8, panel B).

Taken together with the MALDI-TOFMS evidence and the GC–MS lipid analysis, these NMR data support the conclusion that the major sophorolipids from the new species *Candida* sp. NRRL Y-27208 are comprised of a sophorose unit O-glycosidically linked to the 18-hydroxy group of 18-hydroxyoleic acid. By contrast, the sophorose headgroup of the *S. bombicola* sophorolipids is linked to the α-1 hydroxyl group of 17-hydroxyoleic acid. Moreover, when grown on stearic acid *S. bombicola* produces 17-hydroxystearate-containing sophorolipids. In addition, whereas the *S. bombicola* sophorolipid is predominantly in the 1,4’-lactone form, with lesser amounts of the carboxylate in the free anionic form, the new species *Candida* sp. sophorolipid is entirely in the open free-carboxylate form.

The α-linked sophorolipids described are similar to those recently identified from *Candida batistae*, although phylogenetic
3. Experimental

3.1. Materials

*Candida* species were obtained from the ARS Culture Collection (NRRL) at NCAUR, Peoria, IL. Oleic acid (technical grade at 90% purity) was from Aldrich. All other chemicals were reagent grade and used without further purification.

3.2. Yeast fermentation and extraction

*Candida* species were grown in 3 mL of YM medium which contained 5 g/L peptone, 10 g/L glucose, 3 g/L yeast extract, and 3 g/L malt extract in a test tube at 25 °C, 200 rpm for 24 h. This culture was used to inoculate (1%) the production medium (10 mL), 100 g/L glucose, 100 mL/L oleic acid, 1.5 g/L yeast extract, 4 g/L NH₄Cl, 1 g/L KH₂PO₄, 0.1 g/L NaCl, 0.5 g/L MgSO₄ in a 50-mL Erlenmeyer flask. This was cultivated at 25 °C, 200 rpm, for typically 96 h. The pH was adjusted to ~3.5 at 24, 32, 48, and 56 h into the reaction using 1 M NaOH. The samples that included both cells and broth were acidified to pH 2 and extracted twice with 40 mL of EtOAc. Solvents were evaporated from the combined extracts with a rotary evaporator. The oleic acid was separated from the remaining sophorolipids with hexane washes, with the oleic acid going into the hexane. The purified sophorolipids were analyzed using MALDI-TOFMS.

3.3. Compositional analysis by gas chromatography/mass spectrometry (GC–MS)

Samples were hydrolyzed in trifluoroacetic acid (2 M, 110 °C, 30 min) on a reaction block. After cooling, the solvent was removed by evaporation and aldononitrile acetates were prepared as described previously. GC–MS analysis was performed on an Agilent (Santa Clara, CA) 6890N gas chromatograph interfaced with an Agilent 5973N mass-selective detector configured in electron impact (EI) mode, and with a Hewlett-Packard (Santa Clara, CA) 7863 series autoinjector. Chromatography was achieved on a Hewlett-Packard DB-5 ms column (30 m by 0.2 mm) using helium as the carrier gas. The oven temperature was ramped over a linear gradient from 150 to 300 °C at 10 °C per min. Mass spectra were recorded in the positive-ion mode over the range of m/z 60–550. Injector and detector/interface temperatures were 275 and 300 °C, respectively. Data analysis was done off-line using an HP Chemstation.

3.4. Surface tension measurements

The surface tension of sophorolipid solutions was determined using the pendant drop method. Samples were prepared with deionized distilled water and analyzed with an FTA 4000 surface tension instrument (First Ten Angstroms, Inc. Portsmouth, VA). Measurements were made with a 22-gauge blunt needle with a 6-μL drop. The values reported represent an equilibrium surface tension determined 150 s after drop formation. The reported values are the average of a minimum of three separate measurements. The critical micelle concentration determined with software provided by the instrument's manufacturer (FTA32 version 2.0 build 275).

3.5. MALDI-TOFMS analysis

MALDI-TOF mass spectra were recorded on a Bruker Daltonics Omniflex instrument (Bruker Daltonics, Billerica, MA) operating in reflectron mode. Samples were typically dissolved in acetonitrile, and the matrix used was 2,5-dihydroxybenzoic acid. The samples plus matrix were dried onto the 49-place target at room temperature prior to introduction into the spectrometer. Ion source 1 was set to 19.0 kV, and source 2 to 14.0 kV, with lens and reflector voltages of 9.20 and 20.00 kV, respectively. A 200 ns pulsed ion extraction was used with matrix suppression up to 200 Da. The instrument was calibrated externally on a dp series of malto-oligosaccharides. Excitation was at 337.1 nm, typically at 60% of 150 μl maximum output, and 80 shots were accumulated. The reflectron mass resolution (FWHM) for m/z = 2465 (ACTH 18–39) was >20,000.

3.6. Nuclear magnetic resonance (NMR) spectroscopy

All NMR experiments were performed on a Bruker Avance spectrometer (Bruker BioSpin Corp., Billerica, MA) operating at 500.11 MHz using a standard 5-mm z-gradient BBI probe at 600.11 MHz.

Table 1

<table>
<thead>
<tr>
<th>Residue</th>
<th>¹H Chemical shifts</th>
<th>¹³C Chemical shifts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>d-Glu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1H</td>
<td>4.46</td>
<td>4.46</td>
</tr>
<tr>
<td>2H</td>
<td>3.32</td>
<td>3.32</td>
</tr>
<tr>
<td>3H</td>
<td>3.56</td>
<td>3.57</td>
</tr>
<tr>
<td>4H</td>
<td>3.30</td>
<td>3.32</td>
</tr>
<tr>
<td>5H</td>
<td>3.47</td>
<td>3.47</td>
</tr>
<tr>
<td>6H</td>
<td>4.37, 4.20</td>
<td>4.38, 4.20</td>
</tr>
<tr>
<td>7H</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>8H</td>
<td>2.07</td>
<td>2.06</td>
</tr>
</tbody>
</table>

₃H⁻ and ¹³C NMR data and assignments

₃H⁻ linkage, with no evidence for lactone forms. In addition, dimeric and trimeric sophorolipids, but these were not reported for *C. batistae*. Lastly, sophorolipids with hexane washes, with the oleic acid going into the hexane. The purified sophorolipids were analyzed using MALDI-TOFMS.
Figure 7. NMR data obtained on sophorolipids from *Starmerella bombicola* NRRL Y-17069. The data support the 1,4-lactone configuration of the major sophorolipid, and the presence of the 17-hydroxyoleic O-acyl group.

Figure 8. NMR data obtained on sophorolipids from new species *Candida* sp. NRRL Y-27208. The data support the non-lactone, anionic form of the sophorolipids, and the presence of the unusual ω-hydroxy-linked 18-hydroxyoleic group.
27 °C. Chemical shifts are reported as ppm from tetramethylsilane calculated from the lock solvent. The deuterated solvents used were obtained from Cambridge Isotope Labs (Andover, MA). The pulse sequences used were those supplied by Bruker, and processing was done with the Bruker TOPSPIN software package (v. 1.3).

Acknowledgments

We acknowledge the technical assistance of Trina Hartman, and helpful discussion and pre-review editing of the manuscript by Karen Hughes and Dr. Joseph Rich (NCAUR-ARS-USDA, Peoria).

References