Production of sophorolipid biosurfactants by multiple species of the Starmerella (Candida) bombicola yeast clade

Cletus P. Kurtzman1, Neil P.J. Price2, Karen J. Ray2 & Tsung-Min Kuo2

1Foodborne Bacterial Pathogens and Mycology Research Unit, National Center for Agricultural Utilization Research, Agricultural Research Service, US Department of Agriculture, Peoria, IL, USA; and 2Renewable Product Technology Research Unit, National Center for Agricultural Utilization Research, Agricultural Research Service, US Department of Agriculture, Peoria, IL, USA

Abstract

Sophorolipids are carbohydrate-based, amphiphilic biosurfactants that are of increasing interest for use in environmentally benign cleaning agents. Sophorolipid production was tested for 26 strains representing 19 species of the Starmerella yeast clade, including Starmerella bombicola and Candida apicola, which were previously reported to produce sophorolipids. Five of the 19 species tested showed significant production of sophorolipids: S. bombicola, C. apicola, Candida riodocensis, Candida stellata and a new species, Candida sp. NRRL Y-27208. A high-throughput matrix-assisted laser desorption/ionization-time of flight MS assay was developed that showed S. bombicola and C. apicola to produce a lactone form of sophorolipid, whereas C. riodocensis, C. stellata and Candida sp. NRRL Y-27208 produced predominantly free acid sophorolipids. Phylogenetic analysis of sequences for the D1/D2 domains of the nuclear large subunit rRNA gene placed all sophorolipid-producing species in the S. bombicola subclade of the Starmerella clade.

Introduction

The annual worldwide production of surfactants has been reported to be about 10 million tons, with use about equally divided between household detergents and various industrial applications (Van Bogaert et al., 2007). Currently used surfactants are often petroleum derived, but production of these compounds from renewable substrates is now of considerable interest. Sophorolipids, which consist of the sugar sophorose linked to a long-chain hydroxy fatty acid, are among candidate compounds that can be produced from renewable sources.

Interest in sophorolipids is not limited to production of surfactants. The unique chemical structure of sophorolipids can serve as the basis for synthesizing certain hydroxy fatty acids and other compounds (Van Bogaert et al., 2007). Perhaps of greater interest are reports that these glycolipids have antimicrobial activity against certain yeasts (Ito et al., 1980), plant pathogenic fungi (Yoo et al., 2005) and bacteria (Mager et al., 1987; Lang et al., 1989). Furthermore, Shah et al. (2005) showed inhibition of the HIV virus by sophorolipids, and Chen et al. (2006) provided evidence that the compounds have anticancer activity.

Sophorolipids are synthesized by a phylogenetically diverse group of yeasts. The earliest report appears to be that of Gorin et al. (1961), who demonstrated sophorolipid biosynthesis by the anamorphic ascomycetous yeast Candida apicola, which was initially identified as Candida magnoliae. Later, Spencer et al. (1970) showed sophorolipid production by Candida bombicola, and Konoshi et al. (2008) reported Candida batistae to also form sophorolipids. The preceding three Candida sp. are closely related, but sophorolipid biosynthesis was also demonstrated for the less closely related Wickerhamiella domercqiae (Chen et al., 2006) as well as for the basidiomycetous yeast Rhodotorula bogoriensis (Tulloch et al., 1968).

Phylogenetic analysis of sequences for the D1/D2 domains of the nuclear large subunit (LSU) rRNA gene has shown that C. apicola and C. bombicola are members of a clade that is well separated from other ascomycetous yeasts.
Sophorolipids from yeasts

(Kurtzman & Robnett, 1998). Candida bombicola is the first member of the clade for which ascospore formation was discovered and the species was reassigned to the teleomorphic genus Starmerella (Rosa & Lachance, 1998). With the application of sequence analysis to yeast identification, the group of yeasts related to Starmerella bombicola, now termed the Starmerella clade, has increased markedly in the past decade to over 40 species. Many of these species have not been described as yet but are presently recognized from their gene sequences, which have been deposited in GenBank. Candida apicola, C. batistae and S. bombicola are the only members of the Starmerella clade that have been reported to produce sophorolipids. In the present work, we examined additional species of the Starmerella clade for production of sophorolipids using a matrix-assisted laser desorption/ionisation-time of flight (MALDI-TOF) MS-based screen similar to that used previously to identify bacterial biosurfactants, rhamnolipids, surfactins and iturins (Price et al., 2007, 2009; Rooney et al., 2009). From this analysis, we found three additional sophorolipid-producing species of the Starmerella clade, one of which is a novel species, and have determined that two forms of sophorolipids are selectively synthesized by different species within the clade.

**Materials and methods**

**Yeast cultures**

The strains examined in this study were obtained from the ARS Culture Collection (NRRL), National Center for Agricultural Utilization Research, Peoria, IL, and maintained on YM agar (3 g L\(^{-1}\) yeast extract, 3 g L\(^{-1}\) malt extract, 5 g L\(^{-1}\) peptone, 10 g L\(^{-1}\) glucose and 20 g L\(^{-1}\) agar, in distilled water).

**Sophorolipid production medium and growth conditions**

The medium used for production of sophorolipids was termed SL medium and composed of 100 g L\(^{-1}\) glucose, 87.5 g L\(^{-1}\) (100 mL L\(^{-1}\)) oleic acid (Aldrich, technical grade), 1.5 g L\(^{-1}\) yeast extract, 4 g L\(^{-1}\) NH\(_4\)Cl, 1 g L\(^{-1}\) KH\(_2\)PO\(_4\), 0.1 g L\(^{-1}\) NaCl and 0.5 g L\(^{-1}\) MgSO\(_4\)·7H\(_2\)O in distilled water. The initial pH was adjusted to 4.5 with 6 N KOH. Unless specified otherwise, cultures were grown at 25 °C in 50-mL Erlenmeyer flasks with 10 mL of SL medium and shaken at 200 r.p.m. in an Innova 4335 shaker incubator. Incubation times were either 96 or 168 h and the time is given with each reported experiment. The pH of the flask cultures was adjusted to 3.5 twice daily by the addition of 1 N NaOH.

**Separation and quantitation of sophorolipids and oleic acid**

The 10 mL of spent SL medium from each shake flask was acidified with 0.4 mL 6 N HCl and extracted twice with 40 mL of ethyl acetate to remove sophorolipids and unmetabolized oleic acid. The ethyl acetate extract was reduced to dryness in a rotoevaporator, redissolved in 2 mL chloroform, transferred to a glass vial and reduced to dryness under a nitrogen gas stream. Oleic acid was separated from sophorolipids in the mixture by three separate 3 mL hexane extractions. The hexane was evaporated and the concentration of oleic acid was quantified by weight, which was confirmed by gas–liquid chromatography (Price et al., 2009). The residue that remained after hexane extraction was the sophorolipid fraction and the amount was determined by weight following confirmation of the presence of sophorolipids by MALDI-TOF MS as described below. Yields of sophorolipids and consumption of oleic acid are reported as averages and were determined from duplicate cultures, which varied no more than 11%.

**MALDI-TOF mass spectrometric screen for sophorolipids**

MALDI-TOF MS screening was accomplished using a Bruker Omniflex instrument in reflectron mode with positive ion detection. The samples were dissolved in ethyl acetate and the matrix used was 2,5-dihydroxybenzoic acid. The instrument conditions were used as described previously (Price et al., 2009). Determinations were performed in duplicate.

**DNA preparation, gene sequencing and phylogenetic analysis**

The methods used for DNA isolation, purification and sequencing were reported earlier (Kurtzman & Robnett, 1998). DNA characterization was initiated by PCR amplification of the D1/D2 domain of the LSU rRNA gene followed by sequencing reactions using the ABI Big Dye Terminator v3.0 Cycle Sequencing Kit. Sequences of both DNA strands were determined by capillary electrophoresis using an ABI 3130 genetic analyzer (Applied Biosystems, Foster city, CA). Phylogenetic analysis of the gene sequences was determined using the maximum parsimony program included in **PAUP**^*^ 4.063a (Swofford, 1998). Sequences were visually aligned for analysis and *Saccharomyces cerevisiae* was the designated outgroup species.

**Results and discussion**

In the present study, 26 strains representing 19 species of the Starmerella clade were analyzed for production of sophorolipids. Results are reported in Fig. 1, which gives the yield for strains of each species, and the phylogenetic placement...
Fig. 1. Phylogenetic tree of the Starmerella clade determined from maximum parsimony analysis of gene sequences for the D1/D2 domains of nuclear LSU rRNA. Species names are followed by the culture collection strain number and the GenBank accession number for the sequence analyzed. Strains analyzed in the present study are presented in bold type and the yield of sophorolipids is given in brackets. For sophorolipid production, cultures were grown in SL medium in a shaker incubator at 25°C, 200 r.p.m. for 96 h. Strains with Y or YB prefixes designate NRRL numbers. Additional strain information may be obtained using the GenBank accession numbers. Numbers at tree nodes are bootstrap values based on 1000 replicates. TType strain, NTeotype strain. Saccharomyces cerevisiae was designated as outgroup species in the phylogenetic analysis and Candida bombihila served as an unrelated reference species for sophorolipid analysis.
of the strains as determined from the analysis of D1/D2 LSU rRNA gene sequences. Five of the 19 species tested showed significant production of sophorolipids: C. apicola, S. bombicola, Candida riodocensis, Candida stellata and a new species of Candida, NRRL Y-2708, which will be described in a future study. In our earlier work, phylogenetic analysis detected 12 species in the Starmerella clade (Kurtzman & Robnett, 1998) and they separated into two subclades, one represented by C. bombicola and the other by C. magnoliae. With the widespread application of gene sequence analysis in yeast taxonomy, 41 separate lineages (species) are known for the clade and all were included in the phylogenetic analysis shown in Fig. 1 to lend perspective to the placement of species that were tested for the biosynthesis of sophorolipids. However, many of the lineages are undescribed species, which are recognized only from their gene sequences, and cultures are not presently available for analysis.

Even with the addition of many new lineages to the Starmerella clade, the two subclades originally recognized are still evident. Based on the present analysis, sophorolipids are produced only by members of the S. bombicola subclade. Although not included in our analysis, C. batistae was shown by Konoshi et al. (2008) to form sophorolipids, and this species is a member of the S. bombicola subclade (Fig. 1). As seen from Fig. 1, not all members of the subclade produce sophorolipids, and of particular interest for C. apicola, NRRL Y-2481 gave the greatest yield of any strain tested, whereas NRRL Y-6688, a somewhat divergent strain of this species, produced essentially no sophorolipids. In earlier studies of sophorolipid biosynthesis by C. apicola, Tulloch et al. (1968) reported a yield of 40 g L⁻¹ without optimizing the culture medium, much as we found in our assays. Our goal in this study was to test previously unexamined species for sophorolipid production without optimization. We did, however, examine the effect of incubation time, shaker speed and glucose concentration on sophorolipid production by C. bombicola NRRL Y-17069 and Candida sp. NRRL Y-27208, which as described below, produce sophorolipids with a different molecular structure.

Starmerella bombicola NRRL Y-17069 gave maximum sophorolipid yield after 144-h incubation, whereas Candida sp. NRRL Y-27208 produced near-maximum yield after 168 h, the time of the final analysis (Table 1). For both species, increasing yields of sophorolipids were accompanied by decreasing concentrations of oleic acid, which was expected because of the incorporation of oleic acid into the sophorolipid molecule. The requirement for high aeration in production of sophorolipids was reported earlier (Guilmanov et al., 2002) and again shown in this study for both S. bombicola NRRL Y-17069 and Candida sp. NRRL Y-27208 (Table 2). The maximum yield of sophorolipids was obtained at a shaker speed of 350 r.p.m.

Glucose concentration noticeably affected sophorolipid production by both S. bombicola NRRL Y-17069 and Candida sp. NRRL Y-27208 (Table 2). For S. bombicola, 50 g L⁻¹ glucose yielded 48.8 g L⁻¹ sophorolipid, whereas 150 g L⁻¹ glucose yielded 95.4 g L⁻¹ sophorolipid. The increased sophorolipid production was not fully reflected in the reduced concentration of residual oleic acid (Table 2), suggesting that a portion of the lipid moiety was synthesized by the yeast. During production of sophorolipids, the pH of the culture medium declined from 4.5 to as low as 1.8. To sustain production, the pH was readjusted twice daily to 3.5 with 1N NaOH. The precipitous decrease in pH during sophorolipid production and its impact on reducing yield was reported earlier by Gobbert et al. (1984).

Table 1. Production of sophorolipids by Starmerella bombicola NRRL Y-17069 and Candida sp. NRRL Y-27208 over a period of 168 h*

<table>
<thead>
<tr>
<th>Time (h)</th>
<th>S. bombicola</th>
<th>Candida sp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>0.7/84.3</td>
<td>0.6/84.5</td>
</tr>
<tr>
<td>48</td>
<td>11.8/65.4</td>
<td>3.4/79.2</td>
</tr>
<tr>
<td>72</td>
<td>54.4/28.7</td>
<td>11.0/64.3</td>
</tr>
<tr>
<td>96</td>
<td>61.0/33.9</td>
<td>11.4/67.8</td>
</tr>
<tr>
<td>120</td>
<td>62.6/27.3</td>
<td>30.4/35.9</td>
</tr>
<tr>
<td>144</td>
<td>64.0/23.5</td>
<td>39.6/14.8</td>
</tr>
<tr>
<td>168</td>
<td>63.6/16.3</td>
<td>44.5/6.2</td>
</tr>
</tbody>
</table>

*Growth was in SL medium, 25 °C, with a shaker speed of 200 r.p.m. The initial concentration of oleic acid was 87.5 g L⁻¹.

Table 2. Effect of glucose concentration and shaker speed on production of sophorolipids by Starmerella bombicola NRRL Y-17069 and Candida sp. NRRL Y-27208

<table>
<thead>
<tr>
<th>Growth conditions</th>
<th>S. bombicola</th>
<th>Candida sp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose concentration (g L⁻¹)*</td>
<td>50</td>
<td>48.8/9.7</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>69.9/10.3</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>95.4/10.0</td>
</tr>
<tr>
<td>Shaker speed (r.p.m.)</td>
<td>150</td>
<td>35.6/52.5</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>62.2/34.2</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>62.6/31.1</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>67.9/27.1</td>
</tr>
<tr>
<td></td>
<td>350</td>
<td>75.1/19.4</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>76.4/18.3</td>
</tr>
</tbody>
</table>

*SL medium with glucose concentrations as indicated, shaker speed 200 r.p.m., incubation 168 h, 25 °C.
The solvent extracts obtained from all 26 strains examined were initially screened for the presence of sophorolipids by MALDI-TOF MS using techniques developed previously by Price et al. (2009). The spectra were characterized by molecular adduct ions for sophorolipids in the mass range 620–720 Da (Fig. 2). Major ions at \( m/z \) 711 and \( m/z \) 729 are respectively attributed to the \([M+Na]^+\) molecular adduct ions for the lactone and free acid forms of the major diacetylated sophorolipid, \( 6'6''\)-diacetyl-\( \beta\)-\( \delta\)-glucopyranosyl-2'-O-\( \beta\)-\( \delta\)-glucopyranosyl-oxy-octadecenoic acid (Asmer et al., 1988). The observed 18 Da difference between these two ions corresponds to the mass difference between the free carboxylic acid form and the ester-linked \( 4'\)-O-lactone (Fig. 2). Less intense ions at \( m/z \) 669 and \( m/z \) 687 correspond to the monoaacetylated forms of the major sophorolipids, and \( m/z \) 627 and \( m/z \) 645 correspond to the non-acetylated forms (Fig. 2). The 18 Da mass difference between these two sets of ions is again indicative of the free acid and lactone forms of the minor sophorolipids, and the 42 Da difference between di-, mono- and non-acetylated species is characteristic of \( O\)-linked acetyl groups (Price et al., 2009). Similar sophorolipid ions were also observed previously for \( C. bombicola \) by fast atom bombardment MS (Asmer et al., 1988; De Koster et al., 1995).

The five species of the \textit{Starmerella} clade tested that showed the most prominent production of sophorolipids: \( S. bombicola \) NRRL Y-17069, \( C. stellata \) NRRL Y-1446, the new species of \textit{Candida}, NRRL Y-27208, \( C. riodocensis \) NRRL Y-27859 and \( C. apicola \) NRRL Y-2481, were further examined by MALDI-TOF MS. These five strains showed a clear structural diversity for the sophorolipids produced (Fig. 2). \( S. bombicola \) NRRL Y-17069 produced a major di-\( O\)-acetylated lactone form of sophorolipid \([M+Na]^+, m/z \) 711), plus a minor component of this as the free acid form \([M+Na]^+, m/z \) 729). This latter ion is complicated by an adjacent ion at \( m/z \) 727 that is assigned as the potassium adduct \([M+K]^+\) of the major lactone form. By contrast, \( C. stellata \), \textit{Candida} sp. NRRL Y-27208 and \( C. riodocensis \) produced very little of this lactone form (Fig. 2), and the major ion \((m/z \) 729) for these three species is attributed to a di-\( O\)-acetyl free acid form. These strains also produced free acid forms of the monoacetylated and non-acetylated sophorolipids characterized by MALDI-TOF MS ions at \( m/z \) 687 and \( m/z \) 645. \textit{Candida riodocensis} and \textit{Candida} sp. NRRL Y-27208, but not \( C. stellata \), also produced monoacetylated sophorolipid in the lactone form \([M+Na]^+, m/z \) 669). The greatest heterogeneity for sophorolipid production was observed for \( C. apicola \) NRRL Y-2481. Similar to \( S. bombicola \), this strain mainly produced lactone sophorolipids, although with \( C. apicola \), the di-\( O\)-acetyl \((m/z \) 711), mono-\( O\)-acetyl \((m/z \) 669) and non-acetyl \((m/z \) 627) forms were observed. The free acid forms of these three sophorolipids were also observed as minor components from \( C. apicola \), as characterized by ions 18 Da larger at \( m/z \) 729, \( m/z \) 687 and \( m/z \) 645, respectively (Fig. 2). Interestingly, the free acid form was the major component of sophorolipids produced by \( C. batistae \) (Konoshi et al., 2008).

Conclusions

This study demonstrated that in addition to \( S. bombicola \), \( C. apicola \) and \( C. batistae \), three other species of the \textit{Starmerella}...
clade synthesize significant amounts of sophorolipids, i.e., C. riodocensis, C. stellata and Candida sp. NRRL Y-27208. Based on our phylogenetic analysis, sophorolipids were produced only by members of the S. bombicola subclade of the Starmerella clade.

MALDI-TOF MS showed certain of the species to produce sophorolipids predominantly in the lactone form, whereas the other species predominantly gave the free acid form. It should be noted that although MALDI-MS is well suited for the rapid screening and characterization of sophorolipids with diverse molecular mass, it is unable to distinguish between positional isomers, such as differences in the location of acetyl groups, or the fatty acid double bond or acyl-glycosidic linkage. For this reason, a more complete structural characterization of the sophorolipids from Candida sp. NRRL Y-27208 will be published later.

**Acknowledgements**

We thank Eleanor Basehoar-Powers and Trina Hartman for technical assistance and Jamie Schroderus for graphics assistance in preparation of Fig. 1. The mention of firm names or trade products does not imply that they are endorsed or recommended by the US Department of Agriculture over other firms or similar products not mentioned.

**References**


**Table 3. Deduced structures of the sophorolipids identified by MALDI-TOF MS**

<table>
<thead>
<tr>
<th>Forms</th>
<th>R/R**</th>
<th>Calc. mass*</th>
<th>[M+Na]+ (m/z)</th>
<th>R/O formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lactone</td>
<td>OH/OH</td>
<td></td>
<td>627.34</td>
<td>C_{30}H_{52}O_{12}</td>
</tr>
<tr>
<td></td>
<td>OH/OH</td>
<td></td>
<td>645.35</td>
<td>C_{30}H_{54}O_{13}</td>
</tr>
<tr>
<td></td>
<td>OH/OH</td>
<td></td>
<td>669.36</td>
<td>C_{30}H_{54}O_{14}</td>
</tr>
<tr>
<td></td>
<td>OH/OH</td>
<td></td>
<td>687.36</td>
<td>C_{32}H_{56}O_{14}</td>
</tr>
<tr>
<td></td>
<td>OH/OH</td>
<td></td>
<td>711.36</td>
<td>C_{34}H_{58}O_{15}</td>
</tr>
<tr>
<td></td>
<td>OH/OH</td>
<td></td>
<td>729.37</td>
<td>C_{36}H_{58}O_{15}</td>
</tr>
</tbody>
</table>

*Accurate masses were calculated using ISOPRO 3.0.*

Supplied by the U.S. Department of Agriculture, National Center for Agricultural Utilization Research, Peoria, Illinois