RSC Advances

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *RSC Adv.*, 2026, **16**, 12604

Variovoracins: new glycolipopeptide biosurfactants identified by surfactant screening and genome mining

Bradley Haltli, †*[abe] Douglas H. Marchbank, †[ab] Fabrice Berrué, [d] Hebelin Correa, [a] Nicholas G. McCarville,[a] Noelle Duncan,[a] Joshua Kelly,[a] Alyssa L. Grunwald, [a] Martin Lanteigne,[a] Katherine McQuillan[a] and Russell G. Kerr [abc]

Increasing awareness of the environmental issues associated with conventional surfactants has incentivized the chemical industry to develop more sustainable alternatives such as biosurfactants, which are a class of surface-active molecules produced by microorganisms. Biosurfactants are environmentally benign and have the potential to be manufactured sustainably through microbial fermentations fed by renewable feedstocks, unlike conventional surfactants produced from petrochemical feedstocks. Biosurfactants therefore have an important role in the development of sustainable surfactants for consumer and industrial use. To identify new biosurfactants, we screened a microbial library for strains producing emulsifiers. Analysis of active microbial extracts by UHPLC-HRMS led to the identification of the variovoracins, a new class of glycolipopeptide biosurfactants, from *Variovorax paradoxus* RKNM0096. Their chemical structures were elucidated by NMR spectroscopy, and the absolute configuration was determined through multiple chiral derivatization techniques. Meanwhile, tandem mass spectrometry revealed the presence of numerous homologues differing in the position of the C_8 or C_{12} acyl chain. Genome sequencing and bioinformatic analysis identified a candidate biosynthetic gene cluster (BGC) whose role in variovoracin biosynthesis was confirmed by functional characterization of a key rhamnosyltransferase, which was shown to efficiently convert a mono-rhamnosylated pathway intermediate to the final di-rhamnosylated variovoracin. Genome mining identified 123 BGCs homologous to the variovoracin cluster distributed in 12 genera within the *Pseudomonadota*, but concentrated within the *Variovorax* genus. Production of variovoracin analogues was demonstrated in two strains predicted to produce mono-rhamnosylated congeners, further validating the assignment of the variovoracin BGC. Despite their structural similarity to the rhamnolipids, the characteristic curvature (Cc) of the variovoracins has the opposite sign, indicating that their non-ionic and hydrophobic properties confer unique functional characteristics and consequently different surfactant applications.

Received 2nd January 2026
Accepted 16th February 2026

DOI: 10.1039/d6ra00033a

rsc.li/rsc-advances

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Introduction

Surfactants are molecules possessing both hydrophobic and hydrophilic moieties which allow them to be adsorbed at the interface of air–water and water–oil phases or self-assemble to form micelles. As a result, they exhibit a range of detergency, emulsifying, wetting and foaming properties.[1] Global annual consumption of surfactants is estimated at 18.25 million tons in 2024, with a corresponding value of $47.36 billion, which is projected to grow to more than $70 billion by 2032.[2,3] These ubiquitous chemicals are used extensively in consumer products and industrial processes. The majority of synthetic surfactants are alkyl sulphates or sulphonates, which are derived from petrochemical sources[4] and typically toxic and recalcitrant to biodegradation.[5] Due to environmental and sustainability concerns, there is a need to develop more sustainable surfactants as partial or full replacements for petrochemical-derived surfactants. Bio-based surfactants derived, at least in part, from renewable raw materials have emerged as a more sustainable alternative.[6] However, their production still requires chemical synthesis and there remains

[a]*Croda Canada Limited, Regis and Joan Duffy Research Centre, 550 University Avenue, Charlottetown, PE, C1A 4P3, Canada. E-mail: Bradley.Haltli@croda.com*
[b]*Department of Biomedical Sciences, University of Prince Edward Island, 550 University Avenue, Charlottetown, PE, C1A 4P3, Canada*
[c]*Department of Chemistry, University of Prince Edward Island, 550 University Avenue, Charlottetown, PE, C1A 4P3, Canada*
[d]*Aquatic and Crop Resource Development, National Research Council of Canada, Halifax, NS, B3H 3Z1, Canada*
[e]*Department of Chemistry, Technical University of Munich, Lichtenbergstraße 4, 85748 Garching, Germany*
† These two authors contributed equally to this work.

© 2026 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

concerns over deforestation from sourcing raw materials commonly used in bio-based surfactants such as palm oil.[7]

An alternative to synthetic (biobased or petrochemical-derived) surfactants are biosurfactants produced by microorganisms. Microbial biosurfactants are an especially promising replacement for synthetic surfactants due to the scalability and sustainability of fermentation processes. In the producing microorganism, these molecules play important roles in motility,[8,9] biofilm maintenance,[10] virulence,[11–13] and hydrophobic substrate utilization.[14,15] Biosurfactants offer a number of advantages over conventional surfactants, including sustainable production, biodegradability and lower toxicity.[16] In addition to their surface-active properties, biosurfactants can exhibit a wide range of potentially beneficial bioactivities such as antimicrobial, antiviral, anticancer, and insecticidal activity.[17,18]

Biosurfactants can be classified according to size as high- or low-molecular-weight biosurfactants. High-molecular-weight biosurfactants are comprised of large amphiphilic macromolecules such as lipopolysaccharides and lipoproteins, while low-molecular-weight biosurfactants can broadly be divided into lipopeptide and glycolipid subclasses.[6] Both subclasses of low molecular weight biosurfactants possess a hydrocarbon moiety, which mediates hydrophobic interactions and is typically comprised of saturated, unsaturated and/or hydroxylated fatty acids or fatty alcohols. The hydrophilic moiety that mediates polar interactions is often comprised of amino acids, peptides, sugars or functionalities such as carboxylic acids.[19] Lipopeptides consist of peptides linked to fatty acids and are best exemplified by surfactin, which is produced by *Bacillus* species. Lipopeptides with similar surfactant properties are produced by other genera of bacteria such as *Pseudomonas* and *Streptomyces*.[20] Glycolipid surfactants can be categorized into five major groups: rhamnolipids (RLs), trehalose lipids, mannosylerythritol lipids (MELs), cellobiose lipids and sophorolipids.[21] The structure of these biosurfactants plays a major role in determining their functional characteristics, thus the discovery of novel biosurfactants is needed to expand the functional performance of biosurfactants so they can be applied to a wider diversity of applications. In addition, the discovery of novel biosurfactants from non-pathogenic and non-genetically-modified organisms would offer an advantage over existing biosurfactants, such as the RLs, which are produced by pathogenic or genetically modified organisms.[22,23]

Variovorax bacteria are *Betaproteobacteria* belonging to the family Comamonadaceae. Members of this genus are globally distributed and are common inhabitants of soil and water.[24] The genus currently consists of 24 species and *V. paradoxus* is the type species of the genus.[25] *Variovorax* species exhibit a remarkable catabolic capacity to degrade and detoxify a wide variety of organic and inorganic pollutants and are frequently isolated from contaminated soils.[24] *Variovorax* species, including *V. paradoxus*, are also notable for their ability to promote plant growth and produce siderophores.[26–30]

Herein we report the discovery of a new group of glycolipopeptide biosurfactants, which we have named the variovoracins. These compounds were identified by ultra-high-performance liquid chromatography-high-resolution mass spectrometry (UHPLC-HRMS) analysis of a *V. paradoxus* RKNM0096 fermentation that exhibited strong emulsification activity, and their structures were characterized by NMR spectroscopy and tandem mass spectrometry. A putative variovoracin biosynthetic gene cluster (BGC) was identified in the *V. paradoxus* RKNM0096 genome and the involvement of this BGC in variovoracin biosynthesis was established by characterizing a key glycosyltransferase encoded in the BGC. Genome mining identified homologous BGCs in several other bacteria, and the production of variovoracin analogues was demonstrated in two additional *Betaproteobacteria* species. Using the hydrophilic-lipophilic difference (HLD) model, the characteristic curvature (Cc) of the variovoracins was determined to identify potential applications for this new group of biosurfactants.

Results and discussion

Identification of a new class of small molecule biosurfactants

As part of a program aimed at identifying microbial producers of biosurfactants, a clarified culture broth of a bacterium designated as strain RKNM0096 exhibited strong emulsification activity in a kerosene/water emulsification assay, generating an emulsifying index (E_{24}) of 53%. Identification of the bacterium by analysis of its 16S rRNA gene sequence (1538 bp) using the EZ BioCloud identification tool indicated it was most closely related to *V. paradoxus* EPS (CP002417; 99.93% identity).[31] To determine if the observed emulsification activity was due to the production of a small molecule biosurfactant, culture broths of RKNM0096 were analyzed by UHPLC-HRMS, revealing the presence of three major metabolites exhibiting protonated ions of *m*/*z* 1007.6635 $[M + H]^+$ (5.61 min) and 1049.6734 $[M + H]^+$ (overlapping peaks; 5.79 and 5.90 min) (Fig. S1). These *m*/*z* values were consistent with proposed molecular formulae of $C_{51}H_{94}N_2O_{17}$ and $C_{53}H_{96}N_2O_{18}$, respectively. A database search using the Identifier of Natural Products Database (Wiley, 2025 release) and the Dictionary of Natural Products (CHEMnetBASE, v 33.2) using a 5 ppm mass window resulted in no hits for these *m*/*z* values, indicating these metabolites were likely new.

Structural characterization of the variovoracins

The peak corresponding to *m*/*z* 1007.6635 $[M + H]^+$ (**1**) was purified from crude extracts using automated normal-phase flash chromatography followed by semi-preparative reversed phase (RP) HPLC. The structure of **1** was elucidated using a combination of NMR spectroscopy and tandem mass spectrometry. The ^1H and ^{13}C NMR spectra of **1** (Fig. S2–S8) showed the presence of two carbohydrate substituents, as indicated by anomeric proton resonances H-1′ (δ_H 4.79, 1H, d, J = 1.6 Hz) and H-1″ (δ_H 5.01, 1H, d, J = 1.6 Hz), which showed HSQC correlations with carbon resonances C-1′ (δ_C 101.4) and C-1″ (δ_C 103.9), respectively (Table S1). Analysis of two-dimensional NMR spectra confirmed the presence of two separate spin systems involving H-1′ and H-1″. Selective 1D TOCSY spectra acquired by irradiating H-1′ and H-1″ allowed the assignment of proton signals within their respective spin systems. The

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

relatively small ^3J coupling constant of 1.6 Hz between H-1′ and H-2′ (δ_H 3.86, 1H, dd, J = 3.3, 1.6 Hz) showed the axial–equatorial relationship between these protons. A similar coupling constant of 3.3 Hz was observed between H-2′ and H-3′ (δ_H 3.73, 1H, dd, J = 9.6, 3.3 Hz), also indicating their axial–equatorial relationship. Meanwhile, the larger ^3J coupling constants observed for H-4′ (δ_H 3.40, 1H, app. t, J = 9.6 Hz) indicated its axial relationship with both H-3′ and H-5′ (δ_H 3.74–3.68, 1H, m). The methine proton H-5′ exhibited a COSY correlation with methyl proton signal H-6′ (δ_H 1.26, 3H, d, J = 6.2 Hz), revealing the presence of an α-rhamnopyranose moiety. Similarly, the second carbohydrate substituent consisting of H-1″ was also identified as an α-rhamnopyranose. A key HMBC correlation between H-3′ and C-1″ established the α-1,3-glycosidic linkage between these two rhamnopyranose residues.

The anomeric proton H-1′ exhibited an HMBC correlation with C-3C (δ_C 76.5), while methine proton H-3C (δ_H 4.06–4.00, 1H, m) showed COSY correlations with methylene protons H-2aC (δ_H 2.68–2.61, 1H, m), H-2bC (δ_H 2.53–2.48, 1H, m), and H-4C (δ_H 1.59–1.53, 2H, m). The HMBC spectrum also showed coupling between protons H-2bC/H-3C and carbonyl resonance C-1C (δ_C 172.2), establishing the presence of a 3-hydroxyalkanoate moiety. Two additional 3-hydroxyalkanoate moieties were also elucidated from COSY and HMBC correlations exhibited by methine protons H-3A (δ_H 5.27–5.21, 1H, m) and H-3B (δ_H 5.23–5.18, 1H, m). These proton resonances also showed HMBC correlations with C-1B (δ_C 171.7) and C-1C, respectively, indicating ester linkages between the three 3-hydroxyalkanoate moieties. The higher chemical shift values of H-3A and H-3B arising from the electron withdrawing nature of ester carbonyl groups was consistent with this assignment. Overlapping methyl proton (δ_H 0.92–0.88, 9H, m) and methylene envelope (δ_H 1.36–1.27, 1.32–1.24) signals and their integration suggested that the three 3-hydroxyalkanoate moieties were comprised of linear alkyl chains. The length of each alkyl chain was confirmed by tandem mass spectrometry, which established the presence of three 3-hydroxydecanoic acid moieties (Fig. S9, S10 and Table S2).

Methylene protons H-3S (δ_H 3.78–3.72, 2H, m) exhibited a COSY correlation with H-2S (δ_H 4.40, 1H, t, J = 5.7 Hz) and a HMBC correlation with C-1S (δ_C 172.1), showing the presence of a serine amino acid residue, an assignment also consistent with the chemical shifts of carbon resonances C-2S (δ_C 56.7) and C-3S (δ_C 63.1). The spin system of a leucinol residue was also identified through interpretation of COSY and HMBC correlations. The two anticipated amide linkages were established through HMBC correlations between H-2L (δ_H 4.03–3.97, 1H, m) and C-1S and from H-2S to C-1A (δ_C 172.7), confirming the attachment of the dipeptide moiety to the carboxyl terminus of the 3-hydroxydecanoic acid moieties. The planar structure of **1** was therefore elucidated and determined to be the first member of a new group of glycolipopeptides, which we have named the variovoracins (Fig. 1).

The chemical structure of **1**, which possesses both hydrophilic and hydrophobic regions, suggested it was responsible for the observed emulsification activity. As expected, **1** showed strong emulsification activity with an E_{24} value of 53% at a concentration of 1.0 mg mL^{-1} (0.99 mM), consistent with the observed emulsifying properties of the RKNM0096 culture broth. In the diH_2O/kerosene system, the emulsion formed by variovoracin **1** was readily dispersed in water, indicating that the continuous phase of the emulsion was water and that an oil-in-water (o/w) emulsion was formed. The E_{24} values of **1** at pH 6 and 8 were 45% and 34%, respectively, while no emulsification was observed at pH 3 and 10.

Compounds represented by overlapping peaks at 5.79 and 5.90 min in the UHPLC-HRMS chromatogram (Fig. S1) exhibited protonated ions with a m/z value of 1049.6734 $[M + H]^+$. The mass difference of 42 amu compared to **1**, suggested these compounds were mono-acetylated congeners. Purification of the putative acetylated congeners by normal-phase chromatography followed by semi-preparative RP-HPLC led to the purification of mono-acetylated congener **2** (Fig. 2 and Table 1). The ^{1}H and ^{13}C NMR data of **2** closely resembled that of **1** except for the presence of signals belonging to an acetyl group. Through analysis of 2D NMR data, the acetyl moiety was determined to be attached to the 3″ position (Fig. S11–S15 and Table S3). Although RP-HPLC purification of other mono-acetylated congeners was attempted, compound mixtures were obtained and therefore they were not extensively characterized.

Further examination of the *V. paradoxus* culture broth by UHPLC-HRMS analysis also revealed the presence of another variovoracin analogue whose protonated ion of m/z 979.6307 $[M + H]^+$ suggested its molecular formula differed from **1** by C_2H_4. This compound was obtained during RP-HPLC purification of **1** and characterized by NMR spectroscopy (Fig. S16 and S17). The ^{1}H NMR spectrum of this analogue was identical to that of **1** except for the smaller integration of the methylene envelope region, suggesting one of the 3-hydroxydecanoic acid moieties had been substituted with a 3-hydroxyoctanoic acid moiety. Tandem mass spectra confirmed that **3** (Fig. 2 and Table 1) consisted of one 3-hydroxyoctanoic acid moiety and revealed fragment ions with m/z of 509.34 (c – H_4O_2), 339.34 (d – H_4O_2), 311.25 (d – H_4O_2) (Fig. S18), which indicated that this substitution could occur at any one of the three fatty acid positions. Therefore, **3** represents the three homologues **3a–3c**.

Mono-acetylated variovoracins analogous of **3a–3c** were also detected during the RP-HPLC purification of **2**. Variovoracins **4** and **5** (Fig. 2 and Table 1) exhibited protonated ions of m/z 1021.6415 $[M + H]^+$ and m/z 1077.7046 $[M + H]^+$, respectively, indicating their molecular formulae differed from **2** by C_2H_4. Similar to **2**, NMR analysis also showed that **4** and **5** possessed an acetyl group at the 3″ position (Fig. S19–S22). Tandem mass spectrometry showed that **4** and **5** possessed 3-hydroxyoctanoic acid and 3-hydroxydodecanoic acid moieties, respectively, which can also be substituted at any one of the three 3-hydroxydecanoic acid moieties (Fig. S23 and S24). Therefore, **4** and **5** each represent three homologues of **2**, **4a–4c** and **5a–5c**, respectively. Given the presence of other peaks in the UHPLC-HRMS chromatogram exhibiting protonated ions of m/z 1049.6734 $[M + H]^+$, other mono-acetylated variovoracins existed in the *V. paradoxus* crude extract. However, chromatographic overlap between these many analogues, including their putative

 © 2026 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

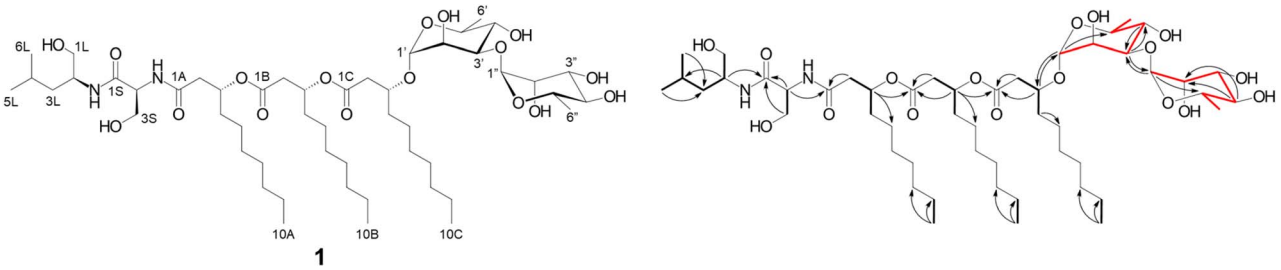


Fig. 1 Structure of variovoracin **1** isolated from *V. paradoxus* (RKNM0096) showing selected COSY (black bold bonds), TOCSY (red bold bonds) and HMBC ($^{1}H \rightarrow {}^{13}C$) correlations.

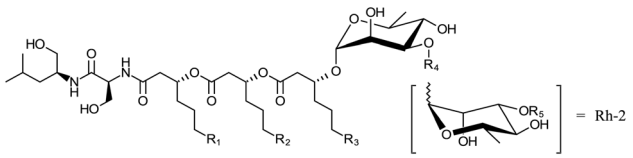


Fig. 2 General structure of variovoracins reported in this study. Groups R_1, R_2, R_3, R_4 and R_5 are defined in Table 1.

Table 1 Variovoracin congeners and homologues characterized in this study

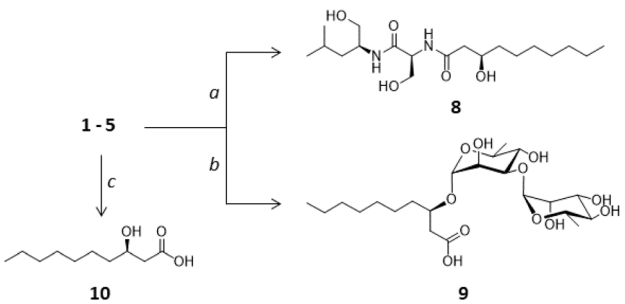
Analog	$[M + H]^+$	R_1	R_2	R_3	R_4	R_5
1	1007.6635	C_4H_9	C_4H_9	C_4H_9	Rh-2	H
2	1049.6734	C_4H_9	C_4H_9	C_4H_9	Rh-2	Ac
3a	979.6307	C_4H_9	C_4H_9	C_2H_5	Rh-2	H
3b		C_4H_9	C_2H_5	C_4H_9	Rh-2	H
3c		C_2H_5	C_4H_9	C_4H_9	Rh-2	H
4a	1021.6415	C_4H_9	C_4H_9	C_2H_5	Rh-2	Ac
4b		C_4H_9	C_2H_5	C_4H_9	Rh-2	Ac
4c		C_2H_5	C_4H_9	C_4H_9	Rh-2	Ac
5a	1077.7046	C_4H_9	C_4H_9	C_6H_{13}	Rh-2	Ac
5b		C_4H_9	C_6H_{13}	C_4H_9	Rh-2	Ac
5c		C_6H_{13}	C_4H_9	C_4H_9	Rh-2	Ac
6	861.6049	C_4H_9	C_4H_9	C_4H_9	H	—
7a	833.5734	C_4H_9	C_4H_9	C_2H_5	H	—
7b		C_4H_9	C_2H_5	C_4H_9	H	—
7c		C_2H_5	C_4H_9	C_4H_9	H	—

3-hydroxyoctanoic acid and 3-hydroxydodecanoic acid homologues, precluded their purification and characterization.

The absolute configuration of the variovoracins was determined using chiral derivatization techniques and UHPLC-HRMS analysis (Scheme 1). A partial acid hydrolysis was performed to generate lipopeptide **8**, which was purified and subjected to the Marfey's method.[32] This analysis demonstrated that leucinol and serine each possessed the L-configuration (Fig. S25 and S26). Meanwhile, a base hydrolysis was also performed to separately obtain glycolipid **9** for chiral derivatization using Tanaka's method (Scheme 1).[33] This approach involved synthesizing diastereotopic thiocarbamoyl-thiazolidine derivatives from the released sugar and comparing UHPLC-HRMS retention time values to commercial L- and D-rhamnose standards. This analysis showed that α-rhamnopyranose moieties possess the L-configuration (Fig. S27). Finally, a mixture of variovoracins were subjected to acid hydrolysis to release 3-hydroxydecanoic acid (**10**), which was subsequently esterified with (*S*)-naproxen chloride.[34,35] Analysis of this derivative by UHPLC-HRMS and comparison of its retention time with standards derived from (*R*)- and (*S*)-3-hydroxydecanoic acid showed that all 3-hydroxydecanoic acid moieties of the variovoracins exhibited the *R*-configuration (Fig. S28).[36]

Identification of the variovoracin biosynthetic gene cluster

To establish the genetic basis of variovoracin biosynthesis, the *V. paradoxus* RKNM0096 genome was sequenced using single-molecule real-time (SMRT) sequencing technology. A total of 140 476 raw subreads with an average length of 11 269 bp were assembled into a single contiguous sequence consisting of 7 191 304 bp. The GC content was 66.70%, which is consistent with other *Variovorax* genomes.[24] Genome annotation identified 6621 putative coding sequences (NZ_CP046508.1). Genome mining for BGCs using AntiSMASH v7.1.0 identified nine BGCs (Table S4).[37] Based on the structure of **1** we hypothesized that its biosynthesis would require a non-ribosomal peptide synthetase (NRPS) to synthesize the dipeptide and one or more glycosyltransferases. AntiSMASH region 9 contained the only dimodular NRPS proximal to two glycosyltransferases in the genome, thus we reasoned that this region most likely contained the BGC for **1** (Table S4; BGC9). The BGC spanned 49 851 bp, encoding 36 putative open reading frames (ORFs), six of which were predicted to play an integral role in variovoracin biosynthesis



Scheme 1 Preparation of variovoracins for Marfey's analysis and other chiral derivatization methods. Reagents and conditions: (a) 0.8 M HCl, H_2O : THF (3 : 1), 60 °C, 18 h. (b) 0.1 M NaOH, H_2O : acetone (9 : 1), rt, 40 h. (c) 1 M HCl, H_2O : CH_3CN (1 : 1), 95 °C, 18 h.

© 2026 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

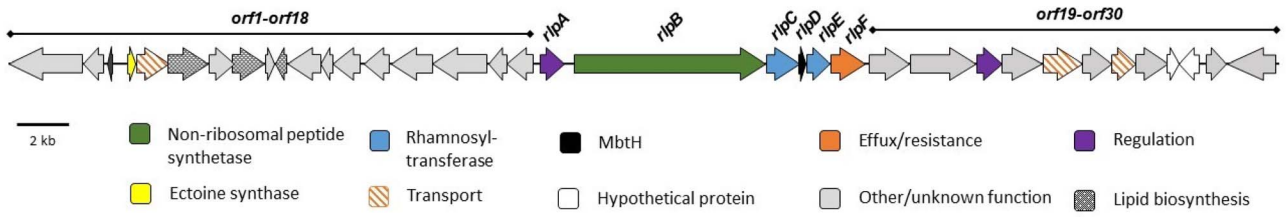


Fig. 3 Variovoracin biosynthetic locus identified in the *V. paradoxus* RKNM0096 genome. The locus shown here corresponds to the complementary sequence between nucleotides 5 338 111 to 5 387 431 in the genome record (NZ_CP046508.1).

(Fig. 3 and Table S5). The six genes, designated *rlpA* to *rlpF* (rhamnosyl-lipo-peptide), are oriented in the same direction, forming a putative operon.

Genes involved in regulation. The first gene of the variovoracin BGC, *rlpA*, encodes a protein that exhibits similarity to transcriptional regulators belonging to the LysR family. Conserved domain analysis indicated that RlpA contained an amino-terminal helix-turn-helix domain and a carboxy-terminal LysR substrate binding domain, which is consistent with the domain architecture of LysR transcriptional regulators. This family of regulators can function as transcriptional activators or repressors,[38] thus it is likely that RlpA plays a role in the regulation of variovoracin biosynthesis.

Genes involved in peptide biosynthesis. Following *rlpA* is a large gene, *rlpB* (7476 bp), which encodes a NRPS. NRPS enzymes are multimodular enzymatic assembly lines that catalyze the condensation of amino acids covalently tethered to the enzyme *via* a thioester linkage to a 4′-phosphopantetheine prosthetic group attached to the active site serine of the peptidyl carrier protein (PCP) domain. NRPS modules minimally consist of three domains, a condensation (C), an adenylation (A) and a PCP domain. The adenylation domain selects amino acid substrates and activates them as aminoacyl-adenylates and then transfers the aminoacyl group to the thiolate of the adjacent PCP domain. The C domain catalyses peptide chain elongation *via* the formation of a peptide bond.[39] Domain analysis of RlpB indicated that the NRPS consists of two modules (M1 and M2) with the domain organization $(\text{C-A-PCP})_{M1}$-$(\text{C-A-PCP-R})_{M2}$.[40] The dimodular structure and domain organization suggests that RlpB generates a dipeptide, which is consistent with the structure of **1**. The first domain of module 1 is a condensation (C) domain. The presence of a C-domain at the beginning of a NRPS initiation module is characteristic of acylated peptides. Amino-terminal C-domains can catalyze amide bond formation between the first amino acid of a peptide and a fatty acid which can be presented to the C-domain as an acyl-ACP or acyl-CoA intermediate, in a process known as lipoinitiation.[41,42] Phylogenetic analysis of the initiation module C-domain (residues 12-437) using the NaPDoS program[43] revealed that the C-domain clustered closely with initiation module C-domains that catalyze the condensation of a fatty acid precursor with an amino acid (bacillibactin NRPS, 38% identity).[44] This observation suggests that variovoracin biosynthesis starts with the condensation of a fatty acid with serine, the first amino acid of the dipeptide. Similar analysis of the second C-domain indicated it was most closely related to the second C-domain of the bacillibactin dimodular NRPS, DhbF (54% identity). Phylogenetic analysis revealed that the M2 C-domain clustered with C-domains catalysing the condensation of two L-amino acids,[43] which is consistent with the structure of **1** and the absence of epimerization domains in the NRPS modules.[39] The substrates of the RlpB A-domains were predicted in antiSMASH by extracting the 10 amino acid substrate specificity codes (A_{M1} – DVWHISLVDK; A_{M2} – DAMLVGAICK) from the deduced amino acid sequence and were predicted to be serine and leucine, respectively.[37] The PCP domains were also analyzed and both contained the core PCP domain motif (DxFFxxLGG[HD]S[LI]) with an invariant serine which represents the 4′-phosphopantetheine attachment site (PCP1-DNFFELGGHSL, PCP2 – DNFFEIGGHSL).[45] The final domain of RlpB is a reductase (R)-domain. R-domains utilize NAD(P)H as a co-factor to reductively release PCP-bound final products as an aldehyde or alcohol.[46] The presence of a leucinol residue at the carboxy-terminus of the variovoracin dipeptide moiety is consistent with release of an acylated dipeptide intermediate by an R-domain. Collectively, the domain structure and organization of RlpB, as well as the predicted substrate specificity of the individual domains are consistent with the structure of **1** and support the assignment of this locus as the variovoracin BGC.

A small gene (*rlpD*) encoding a 70 amino acid protein that shows similarity to MbtH-like proteins was found downstream of *rlpB*. These proteins are often found in association with NRPSs and have been demonstrated to be essential for peptide production and have been shown to facilitate adenylation reactions *via* direct interaction with A-domains.[47,48] Thus we predict that RlpD interacts with one or both A-domains of RlpB to facilitate dipeptide formation.

Genes involved in glycosylation. Glycosylation of the acylated dipeptide generated by RlpB is likely catalyzed by two ORFs (*rlpC* and *rlpE*) downstream of *rlpB.* The deduced amino acid sequence of *rlpC* (439 aa) shows similarity to the GT1 family of glycosyltransferases, which utilize activated sugars as substrates to transfer sugar moieties to a diverse array of acceptor molecules.[49] The deduced amino acid sequence of *rlpE* (322 aa) shows similarity to dTDP-rhamnosyltransferases. In RL biosynthesis two glycosyltransferases act sequentially to transfer two rhamnosyl units to the lipid component of RL.[50] RhlB transfers rhamnose from dTDP-L-rhamnose to the free 3-hydroxyl group of 3-(3′-hydroxydecanoyloxy)decanoic acid (HDD) to generate mono-RL, while di-RL is formed by the

© 2026 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

transfer of an additional rhamnose from dTDP-L-rhamnose to mono-RL by RhlC.[51] The relationship between RlpC/RlpE and the RhlB/RhlC homologs from *P. aeruginosa* PAO1, *B. thailandensis* E264 and *B. pseudomallei* 1710B was investigated by conducting a phylogenetic analysis. In this analysis RlpC clustered with the RhlB orthologs and RlpE clustered with the RhlC orthologs (Fig. S29). Overall sequence similarity of the variovoracin and rhamnolipid enzymes were relatively low with RlpC exhibiting 19–23% identity to the RhlB orthologs and RlpE exhibiting 40–41% identity with the RhlC orthologs. This data suggests that RlpC and RlpE perform similar functions as RhlB and RhlC, respectively. We hypothesize that RlpC catalyses the rhamnosylation of an acylated dipeptide intermediate utilizing dTDP-L-rhamnose as the carbohydrate donor. The limited sequence homology between RlpC and the RhlB orthologs may reflect the significant difference in glycosylation substrates utilized by the enzymes. RlpE is predicted to catalyze the second glycosylation reaction, transferring rhamnose from dTDP-L-rhamnose to the RlpC reaction product.

Genes encoding dTDP-L-rhamnose biosynthesis were not found near the variovoracin gene cluster. Scanning the genome for homologs of *P. aeruginosa* PAO1 rhamnose biosynthetic genes (*rmlBDAC*) identified four genes that exhibited strong sequence similarity to those from *P. aeruginosa* (identity/similarity: RmlB – 79%/89%, RmlD – 60%/71%, RmlA – 78%/89%, RmlC – 66%/80%). In the *V. paradoxus* RKNM0096 genome, the four genes are clustered and are found in the same order as in *P. aeruginosa* (*rmlBDAC*).[52] This locus likely provides the dTDP-L-rhamnose substrates utilized by RlpC and RlpE.

Genes involved in the biosynthesis of the lipid moiety. The *R*-3-hydroxydecanoyl units of the variovoracin 3-(3-(3-hydroxydecanoyloxy)decanoyloxy)-decanoyl (HDDD) lipid moiety may be derived from three likely sources: *de novo* fatty acid biosynthesis[53] or β-oxidation[54] as has been reported for RL biosynthesis, or *via* a dedicated fatty acid biosynthetic pathway located within the BGC, as is the case in the biosynthesis of many lipopeptides such as calcium dependant antibiotic.[55] AntiSMASH analysis did not identify fatty acid synthase genes in the variovoracin BGC, suggesting that the fatty acid precursors are likely derived from *de novo* fatty acid biosynthesis and/or β-oxidation. In RL biosynthesis the HDD moiety is produced by RhlA, which condenses two *R*-3-hydroxydecanoyl-ACP molecules from *de novo* fatty acid biosynthesis or two *R*-3-hydroxydecanoyl-CoA molecules from β-oxidation.[53,54] Scanning of the *V. paradoxus* RKNM0096 genome for RhlA homologs did not identify proteins with significant similarity, suggesting a different mechanism is used in variovoracin biosynthesis. To identify candidate genes that may play a role in the production of HDDD, we analyzed the genes flanking the core *rlpA-F* gene cluster. Three ORFs upstream of *rlpA* (ORFs 6, 8 and 10) exhibit homologies to enzymes that modify fatty acid substrates. The deduced amino acid sequence of *orf6* shows similarity to long-chain fatty acid CoA ligases (FACLs). FACLs activate fatty acids as AMP adenylates and then catalyze the transfer of the fatty acid to the pantetheine group of CoA. Related enzymes in the surfactin,[42] *amphi*-enterobactin[56] and amphibactin[57] BGCs have been demonstrated to provide long-chain fatty acid CoA precursors to N-terminal C-domains in NRPS systems. Orf6 may play a similar role in activating fatty acid intermediates as CoA esters, which may serve as substrates for the N-terminal C-domain of RlpB. The deduced amino acid sequence of *orf8* is similar (~30–50% identity) to putative acyl-CoA dehydrogenases and monooxygenases from a variety of bacteria. Conserved domain analysis indicated that the most significant homology (*E*-value 9.82^{-62}) was to naphthocyclinone hydroxylase (NcnH), which has been demonstrated to hydroxylate polyketide substrates.[58] Thus, Orf8 may hydroxylate decanoyl-CoA to yield 3-hydroxydecanoyl-CoA. Orf10 shows similarity to a variety of genes annotated as 4-hydroxybenzoyl-CoA thioesterases. It also exhibits similarity to acyl-CoA thioesterases belonging the YbgC/YbaW family, which have been shown to catalyze thioester hydrolysis of long chain fatty acyl-CoAs.[59] In the absence of a RhlA homolog a key outstanding question is how the *R*-3-hydroxydecanoyl residues are polymerized to form HDDD prior to being presented to the RlpB N-terminal C-domain. An intriguing hypothesis for this step is the diversion of HDDD intermediates from polyhydroxyalkanoate (PHA) biosynthesis to variovoracin biosynthesis. PhaC is the key enzyme responsible for polymerization of *R*-3-hydroxyalcanoate precursors into short and medium chain-length PHAs.[60] A PhaC homolog is present in the *V. paradoxus* RKNM0096 genome (QSI34435.1) indicating the strain has the ability to produce PHAs that could be used as substrates for variovoracin biosynthesis. Biosynthesis of the HDDD moiety of the variovoracins is not obvious, although it seems likely that *orfs6*, *-8* and *-10* may play a role in its genesis. Further studies are required to establish the involvement of these, and potentially other genes, in the biosynthesis of the lipid component of the variovoracins.

Genes involved in variovoracin acetylation. Acetylated variovoracin analogs are abundant in *V. paradoxus* RKNM0096 fermentation broths. No genes encoding acetyltransferases were identified in the BGC. Thus, it is likely that acetylation is catalyzed by an enzyme encoded elsewhere in the genome.

Genes involved in transport. Directly downstream of *rlpE* is an ORF (*rlpF*) encoding a protein similar to major facilitator superfamily transporters from a variety of bacteria. RlpF exhibits 38% identity and 54% similarity to PA1131 from *P. aeruginosa* PAO1, which is immediately upstream of *rhlC*.[61] RlpF is likely involved in variovoracin efflux.

Proposed biosynthesis

Variovoracin biosynthesis presumably starts with the formation of the HDDD moiety from fatty acid precursors derived from *de novo* fatty acid biosynthesis or β-oxidation. As described above, the origin of this intermediate is not known but its biosynthesis may involve *orfs6*, *-8* and *-10* and potentially intermediates from PHA biosynthesis. After formation of the lipid moiety, it is likely presented as an acyl-coA or acyl-ACP intermediate to the C-domain of RlpB M1 which condenses the lipid moiety with L-serine. RlpB M2 then incorporates L-leucine to form a PCP-bound acylated dipeptide intermediate which is released from the enzyme by the C-terminal R-domain of RlpB, resulting in the formation of a terminal L-leucinol residue. dTDP-L-rhamnose,

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

produced by the *rmlBDAC* operon, is then utilized by the rhamnosyltransferases RlpC and RlpE to sequentially glycosylate the aglycone resulting in the production of the final glycosylated variovoracin **1** (Fig. 4).

Genome mining to identify related biosurfactant producers. To provide support for the identification of the variovoracin BGC and to identify producers of variovoracin analogs we employed a genome mining approach. We reasoned that since *rlpC* exhibited low sequence similarity to the homologous *rhlB* genes involved in rhamnolipid biosynthesis despite utilizing similar substrates, that this gene may encode a glycosyltransferase unique to variovoracin-type biosurfactant BGCs. A BlastP search of the GenBank database using the RlpC sequence, identified 123 hits with sequence identity (ID) > 50%. Manual inspection of the genomic context of the hits revealed that hits with sequence identities >50% were clustered with a mono- or di-modular NRPS, suggesting that they may be involved in glycolipopeptide biosynthesis. The taxonomic distribution of BGCs similar to the variovoracin BGC was limited to the *Pseudomonadota* and was most prevalent in the *Betaproteobacteria* with the following distribution at the genus level: *Variovorax* (71), *Paracidovorax* (15), *Acidovorax* (12), *Collimonas* (5), *Aquabacterium* (1), *Massilia* (1) and *Janthinobacterium* (1). Variovoracin-like BGCs had a limited distribution in the *Alphaproteobacteria* (*Inquilinus* – 9, *Bradyrhizobium* – 5, *Acidocella* – 1) and *Gammaproteobacteria* (*Methylocaldum* – 1, *Tahibacter* – 1). Putative glycolipopeptide BGCs were extracted from the representative genomes of taxa identified in the RlpC blast search and aligned to identify core genes in BGCs homologous to the variovoracin cluster (Fig. S30). This alignment showed that the NRPS, MbtH and RlpC glycosyltransferase homologs were conserved across the 24 aligned clusters. The domain structure of the NRPS in these clusters was partially conserved with 11 clusters (46%) showing a domain structure identical to the variovoracin cluster suggesting that these clusters incorporate a dipeptide into their glycolipopeptide product. An equal number of clusters exhibited a single module structure, suggesting these clusters may generate analogs with a single amino acid.

To test the hypothesis that bacteria possessing a *rlpC* homolog proximal to NRPS and MbtH homologs would produce variovoracin-like glycolipopeptides we screened two strains for glycolipopeptide production, *V. paradoxus* B4 (DSM 21786; genome acc. NC_022247.1) and *Janthinobacterium agaricidamnosum* DSM 9628 (genome acc. NZ_HG322949.1). These strains possess close homologs to RlpC (89.0% and 68.3% identity, respectively). The gene synteny in the *Variovorax* and *Janthinobacterium* homologous BGCs was very similar to that of the variovoracin BGC with a dimodular NRPS followed by *rlpC* and *rlpD* homologs (Fig. S30). The NRPSs from these BGCs exhibited a $(C\text{-}A\text{-}PCP)_{M1}$-$(C\text{-}A\text{-}PCP\text{-}R)_{M2}$ domain structure, with the same predicted A-domain substrate specificity as RlpB. Notably, these BGCs did not contain a *rlpE* homolog. Based on these characteristics we predicted *V. paradoxus* DSM 21786 and *J. agaricidamnosum* DSM 9628 would generate monorhamnosylated variovoracin analogs with a serine-leucinol dipeptide. Screening of fermentation extracts from both strains revealed the presence of three major peaks (t_R 3.07 min, 5.01 min and 5.50 min) in the UHPLC-HRMS chromatogram of *J. agaricidamnosum* DSM 9628 and one major peak (t_R 5.50 min) in the *V. paradoxus* DSM 21786 chromatogram (Fig. S31). The HRMS data corresponding to the peak eluting at 3.07 min was consistent with the mass of the lipopeptide jagaricin (HRESIMS *m*/*z* 1181.6229 $[M + H]^+$, calcd for $C_{56}H_{85}N_{12}O_{16}$, 1181.6201),

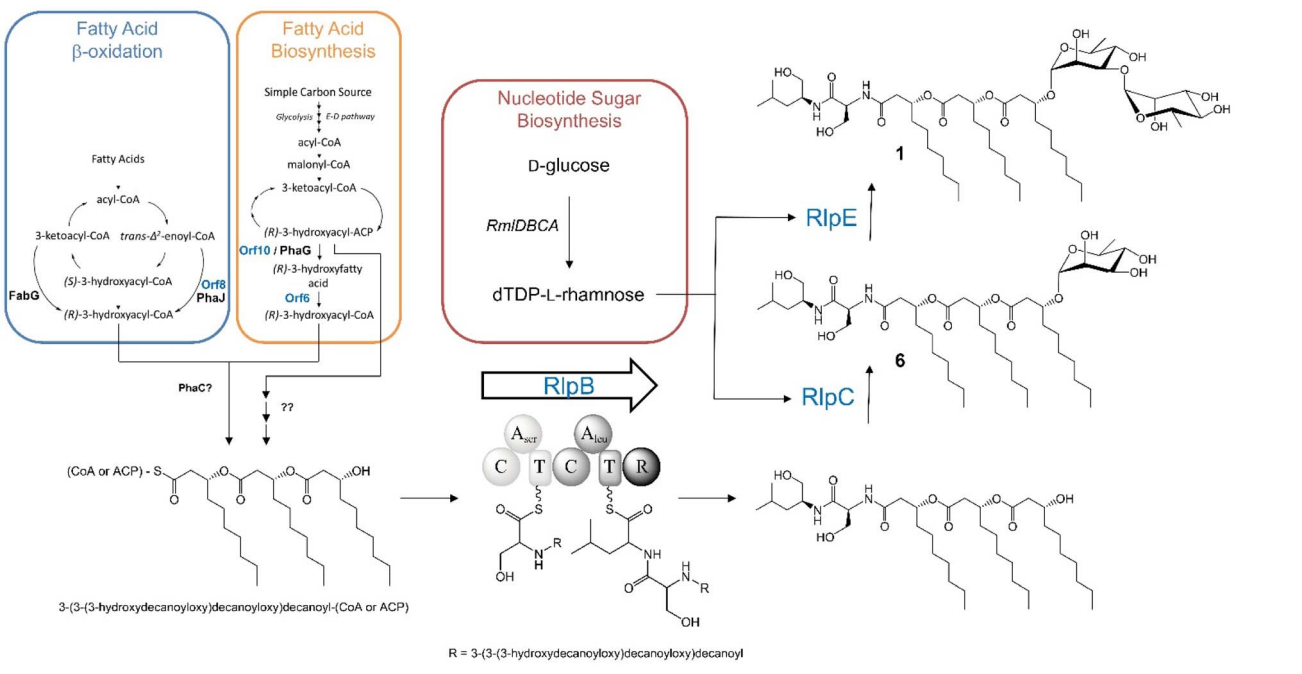


Fig. 4 Proposed variovoracin biosynthetic scheme.

© 2026 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

which was previously reported from this strain.[62] Extraction of the mass spectra for peaks eluting at 5.01 and 5.50 min revealed $[M + H]^+$ ions of *m*/*z* 833.5757 and *m*/*z* 861.6072, respectively. The observed $[M + H]^+$ ions showed a 2.88 and 3.13 ppm mass difference from the predicted *m*/*z* $[M + H]^+$ ions for the monorhamnosyl-analogs of **3a–c** and **1**, respectively (**7a–c** and **6** in Table 1). The peak eluting at 5.50 min in the chromatograms of *J. agaricidamnosum* DSM 9628 and *V. paradoxus* DSM 21786 exhibited the same retention time and monoisotopic mass, indicating the two strains were producing the same metabolite.

To provide further evidence for the production of monorhamnosyl variovoracin analogs the peaks eluting at 5.01 and 5.50 min were purified from *J. agaricidamnosum* DSM 9628 fermentation extracts and the structures elucidate by NMR and tandem mass spectrometry analysis.

The ^{1}H and ^{13}C NMR spectra of the purified metabolites were nearly identical to that of **1**, except that the NMR spectra lacked resonances belonging to the second α-rhamnopyranose moiety (Fig. S32–S39). This data confirmed that the 5.50 min peak corresponded to **6**, the monorhamnosyl-analog of **1**. The 5.01 min peak also lacked α-rhamnopyranose resonances in the NMR spectra, and tandem MS analysis (Fig. S40) revealed that the peak consisted of three compounds (**7a–c**) containing 3-hydroxyoctanoic acid substitutions at each of the three fatty acid positions, similar to **3a–c** (Table 1). Interestingly, no acetylated variovoracin congeners were detected in the *J. agaricidamnosum* DSM 9628 or *V. paradoxus* DSM 21786 culture broth extracts. The absolute configuration of **6** was determined using the previously described chiral derivatization approaches (Scheme S1), showing that the rhamnose, serine, leucinol and 3-hydroxydecanoic acid moieties possessed the same configurations as described for variovoracins **1–5** (Fig. S41b–S43). The structural assignment of **6** and **7a–c**, corroborated our assignment of the *rlpA-F* locus as the variovoracin BGC.

Proving the role of the putative BGC in variovoracin biosynthesis

To unequivocally prove the involvement of the *rlpA-F* locus in variovoracin biosynthesis we attempted to generate insertional mutants in both *V. paradoxus* and *J. agaricidamnosum* but were unable to transform the strains despite an extensive effort. As **6** is the predicted precursor of **1** we reasoned that we could provide strong support for the role of the putative BGC in variovoracin biosynthesis by proving that RlpE could convert **6** to **1**. RlpE was expressed as an N-terminal hexahistidine-tagged protein (6His-RlpE) and purified to near homogeneity by nickel-affinity chromatography. The apparent molecular weight of purified 6His-RlpE by SDS-PAGE was 33.05 kDa (Fig. S44a), which is in good agreement with the expected size (38.2 kDa). The activity of 6His-RlpE was established by incubating the enzyme with dTDP-L-rhamnose and **6**. Reactions were monitored by UHPLC-HRMS and showed complete conversion of **6** to **1** after 4 h (Fig. S44b). This data indicates that RlpE catalyses the second rhamnosylation step in variovoracin biosynthesis in *V. paradoxus* RKNM0096. To provide additional support for the role of RlpE in variovoracin biosynthesis we determined the steady state kinetics to determine the K_m of RlpE for **6**. The K_m was 0.63 mM, indicating a strong affinity of RlpE for **6**. This K_m is similar to that reported for SpnG, which catalyses rhamnosylation of the insecticide spinosyn.[63] As genes for the biosynthesis of natural products in bacteria are typically clustered, these findings provide strong evidence supporting the *rlpA-F* locus as the variovoracin BGC. Future gene inactivation or heterologous expression experiments would be required to unequivocally prove involvement of the *rlpA-F* locus in variovoracin biosynthesis.

Comparison to known biosurfactants

The variovoracins clearly bear resemblance to the RL class of biosurfactants reported from *Pseudomonas* and *Burkholderia*.[22,64,65] Rhamnolipids (RLs) consist of a mono- or di-rhamnose moiety linked *via* an O-glycosidic linkage to the hydroxyl terminus of one, two or rarely three 3-hydroxy fatty acids (Fig. S45). When two L-rhamnose residues are present, they are linked *via* an α-1,2 glycosidic linkage. The hydroxy fatty acids are in the *R*-configuration and can vary in chain length (C_8–C_{16}) and degree of unsaturation and are linked to each other *via* ester bonds between the carboxyl group of one fatty acid and the 3-hydroxyl group of the following fatty acid.[36] The most common RL congener is α-L-rhamnopyranosyl-α-L-rhamnopyranosyl-3-hydroxydecanoyl-3-hydroxydecanoic acid (di-RL). *P. aeruginosa* primarily produces a mixture of mono- and di-RL bearing two 3-hydroxyacyl moieties with C_8–C_{12} chain lengths.[66] In comparison to the RLs, the variovoracins exhibit several important structural distinctions with the most obvious being the presence of a dipeptide moiety attached to the carboxyl end of the terminal 3-hydroxy fatty acid. This modification has important ramifications for the surfactant properties of the variovoracins whose dipeptide moiety renders the variovoracins non-ionic, as opposed to the anionic carboxylate functionality in the case of the RLs. To the best of our knowledge, the only naturally occurring non-ionic RL reported in the literature are methyl esters of mono- and di-RLs.[67] Among the RLs, only one congener produced by *Burkholderia plantarii* has been reported in which the di-rhamnose moiety was linked to three 3-hydroxy fatty acids.[68] Finally, the rhamnose sugars of the variovoracins are bound together through an α-1,3 glycosidic linkage, while all reported di-RLs possess an α-1,2 glycosidic linkage.[66]

The variovoracins also resemble rhizoleucinoside produced by *Bradyrhizobium* sp. BTAi1 and a family of glycyl-glucolipids produced by *Alcanivorax borkumensis* (Fig. S45)[69,70] Rhizoleucinoside is most similar to the variovoracins in that it is comprised of an L-leucinol moiety linked to three sequential *R*-configured 3-hydroxydecanoic acid subunits which are in turn linked to a single α-L-rhamnose. Despite its structural similarity to the rhamnolipids, no efforts to characterize the surfactant properties of rhizoleucinoside were made, although moderate cytotoxicity towards murine and rat microglia were reported.[69] Interestingly, genomic analysis of the *Bradyrhizobium* sp. BTAi1 genome identified a BGC exhibiting strong homology to the variovoracin BGC, suggesting that rhizoleucinoside and variovoracins have a shared biosynthetic origin (Fig. S30). *A.*

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

borkumensis was previously reported to produce anionic biosurfactants consisting of four 3-hydroxydecanoic acids linked to glucose. These glucolipids reduced the surface tension of water to 28 mN m^{-1} with a critical micelle concentration of 0.03 mM.[71] Subsequently, Abraham and colleagues purified a glycine-containing analog from cell wall extracts of the same organism but did not assess its surfactant properties.[70]

Determination of critical micelle concentrations

The critical micelle concentrations (CMCs) of variovoracins **1**, **2**, **3a–3c** and **6** were determined from surface tension measurements as a function of surfactant concentration using the du Noüy-Padday method (Fig. 5 and Table 2).[72] The CMC is the surfactant concentration when the water/air interface of the solution becomes saturated and further addition of surfactant results in the formation of micelles and no further reduction in surface tension. Non-acetylated variovoracins **1** and **3a–3b** both reduced the surface tension of water to slightly below 36 mN m^{-1} with a CMC of 0.20 mM whereas the mono-acetylated and mono-rhamnosylated variovoracins **2** and **6** reduced the surface tension to similar levels between 36-37 mN m^{-1} but with a CMC of 0.85 mM. In comparison, mono- and di-RLs, which were obtained by RP-HPLC purification from a commercially available RL mixture, lowered the surface tension of water to 28.2 and 39.9 mN m^{-1}, respectively, at their CMC of 0.06 mM. The reduction in surface tension by the non-acetylated variovoracins was significantly greater across a narrower concentration range compared to the mono-acetylated and mono-rhamnosylated variovoracins **2** and **6**, whose lower water solubility may have interfered with partitioning into the water/air interface. Meanwhile, the additional rhamnose moiety in di-RL may increase intermolecular repulsion at the water/air interface due to the larger size of its hydrophilic head, resulting in a higher surface tension at its CMC compared to mono-RL. However, a similar trend was not observed in comparing variovoracins **1** and **6** as the issue of water solubility had a more pronounced effect on the ability of **6** to partition into the water/air interface.

Characteristic curvature (Cc) of variovoracin 1

To further characterize the surfactant properties of variovoracins we applied the hydrophilic-lipophilic difference (HLD) model to determine the characteristic curvature (Cc) of **1**, which represents the primary constituent of the RKNM0096 extract. The Cc value is a surfactant characteristic that reflects the type of micelle formed under reference conditions, as described by Acosta *et al.*[73] The HLD (HLD_{mix}) of a formulation predicts its emulsion type and can be calculated from the Cc of each surfactant constituent and other formulation parameters such as temperature and salinity, as described by eqn (1).

$$HLD_{mix} = F(S) - k \times EACN + x_{SDHS} \times Cc_{SDHS} + x_{Var} \times Cc_{Var} \quad (1)$$

where $F(S)$ is the salinity, k is a surfactant constant and equal to 0.17, EACN is the effective alkane carbon number (EACN = 1 for toluene), x_{SDHS} is the molar fraction of SDHS, Cc_{SDHS} is the characteristic curvature of SDHS (−0.92),[73] x_{Var} is the molar fraction of **1**, and Cc_{Var} is the characteristic curvature of **1**.

The Cc value of **1** was determined using a mixed surfactant method with sodium dihexyl sulfosuccinate (SDHS) as the reference surfactant.[73] Electrolyte scans with NaCl were performed for each variovoracin/SDHS molar ratio to find the concentration of NaCl marking a transition from a Type I oil-in-water (o/w) to Type II water-in-oil (w/o) emulsion (Fig. 6). During this transition, a Type III emulsion consisting of a middle bicontinuous phase is obtained, representing the optimal electrolyte concentration (S^*) where HLD = 0. Eqn (2) expresses the optimal salinity (S^*) as a function of the characteristic curvature (Cc_{Var}) of **1** wherein $HLD_{mix} = 0$. Using eqn (2) and plotting S^* as a function of the molar ratio of variovoracin **1** (Fig. 7), the Cc value was determined to be +5.2. On the scale of surfactant Cc, this value is relatively high and consistent with the hydrophobicity of variovoracin **1** and its relative insolubility in water.

$$F(S^*) = k \times EACN - x_{SDHS} \times Cc_{SDHS} - x_{Var} \times Cc_{Var} \quad (2)$$

The Cc value of **1** can be used as a guide to determine potential commercial applications. For instance, an interpretation of the HLD equation (eqn (1)) implies these biosurfactants may have useful formulation applications for attaining a Type II (w/o) emulsion when using an oil possessing a high equivalent carbon number (EACN). The Cc value can also facilitate the identification of synthetic surfactants as potential targets for replacement. Many types of sucrose esters and sulfosuccinate surfactants, such as sucrose distearate (Cc 4.0) and

Table 2 Critical micelle concentrations (CMC) and surface tension (ST) at the CMC of variovoracins and RLs

Compound	CMC (mM)	ST at CMC (mN m^{-1})	Minimum ST (mN m^{-1})
1	0.20	36.8	32.7
2	0.85	36.9	33.5
3a–3c	0.20	36.5	32.5
6	0.85	36.2	30.9
Mono-RL	0.06	28.2	27.1
Di-RL	0.06	39.9	30.1

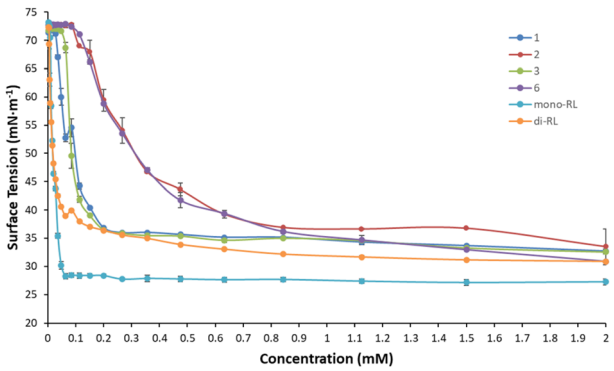


Fig. 5 Surface tension measurements of variovoracins and RLs using the du Noüy-Padday method.

© 2026 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

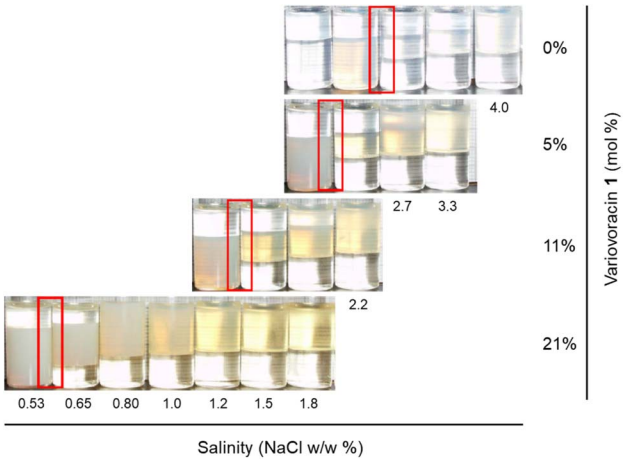


Fig. 6 Determination of optimal salinities (*S**) by the mixed surfactant method using SDHS/variovoracin mixtures with different concentrations of NaCl and molar ratios of variovoracin **1**. Red boxes represent the value of *S** showing the transition from a Type I o/w emulsion (left) to a Type II w/o emulsion (right) with increasing salinity.

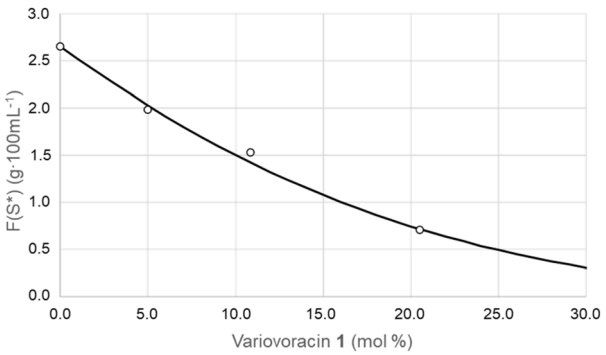


Fig. 7 Plot of optimal salinities represented by *F*(*S**) as a function of **1** molar fraction (mol%). Line of best fit represents Cc values of 5.2 for **1** as derived from eqn (2).

dioctyl sulfosuccinate (Cc 2.6), are potential substitution candidates given the positive sign of their Cc. If blending with a co-surfactant of negative Cc, the HLD equation shows that the higher Cc of **1** could allow one to achieve a HLD of zero at a lower molar concentration compared to these surfactants. In this type of situation, **1** could also be significantly more potent as a co-surfactant compared to, for instance, lauryl glucoside (Cc 1.3) whose smaller Cc would require higher concentrations to move the HLD toward zero. Although sugar esters and alkyl glucosides have bio-based content and are known to be biodegradable,[74,75] their overall environmental impact must also include an assessment of the sustainability of their chemical syntheses and raw materials. Notably, the fatty alcohol portion of alkyl glucosides is mainly derived from the petrochemical industry.[76]

While structurally similar to the variovoracins, the RLs possess a negative Cc, indicating their surfactant properties are significantly different. A Cc value of −1.41 was previously reported from a commercially available mixture of mono- and di-RLs produced by *P. aeruginosa*.[77] Having an additional fatty acid moiety and lacking the anionic carboxylate group renders the variovoracins more hydrophobic than the RLs, which is consistent with the higher Cc value of **1**. According to the HLD model, given the opposite signs of their Cc, these biosurfactants will move the overall HLD of a formulation in separate directions as the variovoracins will tend to form a Type II w/o emulsion while the rhamnolipids will favour a Type I o/w emulsion. As a result, the variovoracins are anticipated to have different formulation applications. Based on their chemical structures, **2** and **6** are anticipated to have a higher Cc than **1**. The hydrophobicity of these analogs may result in disproportionate partitioning to the oil phase rather than the oil/water interface, possibly limiting their functionality as surfactants.

To widen their potential surfactant applications, the variovoracins could be modified to generate analogues exhibiting a range of Cc values. For instance, modification of the NRPS by module or domain swapping or by mutation of the A-domain active site residues could introduce positively or negatively charged amino acids into the peptide component of variovoracins. Alternatively, replacement of the terminal reductase domain of RlpB with a thioesterase domain would generate a carboxylic acid at the peptide C-terminus.[78] Both of these approaches could be used to generate anionic or cationic variovoracins with Cc values significantly lower than that of the naturally occurring non-ionic variovoracin. Such modifications may impart more diverse surfactant properties that may be beneficial depending on the application. The identification of the variovoracin BGC in this study sets the stage for the rational generation of analogs through pathway engineering.

Antimicrobial and cytotoxicity testing

Variovoracins **1**, **2**, and **6** were tested for antimicrobial and cytotoxic activity. No antimicrobial activity against methicillin-resistant *Staphylococcus aureus*, vancomycin-resistant *Enterococcus faecium*, *S. warneri*, *P. aeruginosa*, *Proteus vulgaris*, or *Candida albicans* was observed at a concentration of 128 μg mL^{-1}. Cytotoxicity was assessed against two epithelial cell lines: adult human epiderma keratinocytes and normal human fibroblasts. Variovoracins **1**, **2** and **6** exhibited a similar degree of moderate cytotoxicity towards the fibroblasts and keratinocytes with IC_{50}s between 15–20 μM (Table 3). The observed cytotoxicity is consistent with that expected for biosurfactants that interact with cell membranes[79] and was significantly lower

Table 3 Cytotoxicity (IC_{50}) of three variovoracin congeners against normal adult human keratinocytes and normal human fibroblasts. Zinc pyrithione was used as the cytotoxicity control

	Cell line IC_{50} (μM)	
Compound	Keratinocytes	Fibroblasts
1	15.40 ± 1.69	19.37 ± 2.38
2	18.40 ± 3.81	17.83 ± 1.53
6	18.01 ± 1.86	18.94 ± 0.35
Zinc pyrithione	0.630 ± 0.003	6.93 ± 0.94

© 2026 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

than that observed for zinc pyrithione, an antifungal agent used extensively in consumer products for the topical treatment of dandruff. The cytotoxicity and antimicrobial testing results suggest that the variovoracins would be well tolerated in topical applications and would not be expected to adversely affect the skin microbiome if used as a co-surfactant in topical formulations.

Experimental

General methods

See SI.

Bacterial strains and fermentations

V. paradoxus RKNM0096 was isolated from soil collected from the Battle Bluffs area west of Kamloops, British Columbia, Canada (GPS: 50.731553 N 120.566409 W). Strain RKNM0096 was isolated as a mucoid, yellow pigmented colony on humic acid vitamin agar.[80] *J. agaricidamnosum* DSM 9628 and *V. paradoxus* DSM 21786 were obtained from the DSMZ GmbH culture collection. All strains were maintained on ISP2 agar[81] at room temperature (22 °C ± 2 °C; RT). For fermentations, ISP2 broth seed cultures were inoculated from agar cultures and incubated at RT with shaking at 200 rpm for 24 h. Seed cultures were used to inoculate fermentations (3% v/v) which were incubated at RT (DSM 9628 and DSM 21786) or 30 °C (RKNM0096) with shaking at 200 rpm for 5 days. Small scale 10 mL fermentations were conducted in 25 × 150 mm tubes, while 1 L fermentations were performed in 2.8 L Fernbach flasks. Fermentation medium was ISP2 broth[81] for RKNM0096 and BFM15 (g L^{-1}: sucrose – 20, Bacto peptone – 5, cane molasses – 5, $CaCO_3$ – 5, KI 0.5, $MgSO_4 \cdot 7H_2O$ – 0.5, $FeSO_4 \cdot 7H_2O$ – 0.1) for the DSM strains.

Biosurfactant screening

Emulsification activity was determined using the method of Cooper *et al.*[82] The cell-free broth of *V. paradoxus* RKNM0096 fermentations were added to an equal volume of kerosene and vortexed at maximum speed for 2 min. The emulsification index (E_{24}) was recorded after standing for 24 h at RT and the height of the emulsion (h_{emuls}) and the total height (h_{total}) of the liquid in the tube were measured. The E_{24} was calculated using the equation $E_{24} = h_{emuls}/h_{total} \times 100\%$.

Purification of variovoracins 1–5

V. paradoxus RKNM0096 fermentation broth (7 L) was extracted with Diaion® HP-20 adsorbent resin (700 g, 10% w/w). After centrifugation at 12,500*g* for 10 min, the supernatant was removed and the resin was washed three times with 7 L of deionized water. The resin was extracted twice with 4 L of CH_3OH, which was then evaporated *in vacuo* to provide an extract that was partitioned between ethyl acetate and deionized water. The ethyl acetate layer was removed and evaporated *in vacuo* to provide a dark yellow oil (405.2 mg), which was fractionated by automated normal-phase flash chromatography using the following chromatographic conditions: column, Teledyne Isco RediSep® Gold 40 g silica column (60 Å, 20–40 μm); mobile phase flow rate, 35 mL min^{-1}; linear gradient, 100% hexanes at 3.0 min to hexanes:acetone (70 : 30) at 4.0 min and then 100% acetone at 29.0 min, which was held for 5 min. Fractions containing variovoracins **1–5** were further separated by semi-preparative RP-HPLC using methodology described in the SI.

Purification of variovoracins 6–7

J. agaricidamnosum DSM 9628 fermentation broth (1 L) was extracted twice with 0.5 L of ethyl acetate. The organic layer was evaporated *in vacuo* yielding a crude extract (110.4 mg) that was fractionated by automated normal-phase flash chromatography using the following chromatographic conditions: column, Teledyne Isco RediSep® Gold 15.5 g C_{18} column (200 Å, 20–45 μm); mobile phase flow rate, 30 mL min^{-1}; linear gradient, CH_3OH : H_2O (10 : 90) at 1.0 min to 100% CH_3OH at 11.0 min, which was held for 8.0 min. Fractions containing variovoracins **6–7** were further separated by semi-preparative RP-HPLC using methodology described in the SI.

Variovoracins structure elucidation

The structures of the variovoracin congeners were elucidated using combination of 1D and 2D ^1H and ^{13}C NMR and tandem mass spectrometry. Spectroscopic methodologies and data are provided in the SI.

Stereochemical analysis of variovoracins

See SI for derivatization and UHPLC-HRMS methodologies as well as UHPLC-HRMS chromatograms of chiral derivatives.

Purification of rhamnolipid (RL) standards

The mono- and di-RL standards were obtained by semi-preparative RP-HPLC separation [column, Phenomenex Gemini-NX 110 Å C_{18} column (250 × 10 mm, 5.0 μm); mobile phase flow rate, 3.0 mL min^{-1}; CH_3CN : H_2O (58 : 42, 0.1% formic acid); R_t (di-RL): 19.1 min, R_t (mono-RL): 38.1 min] of a commercially available 90% rhamnolipid mixture (AGAE Technologies, P/N: R90-10 G). The samples were evaporated *in vacuo* to provide 6.9 and 5.2 mg of mono- and di-RL, respectively.

Determination of the critical micelle concentration (CMC) and characteristic curvature (Cc)

CMC values were determined by surface tension measurements on a Kibron Delta-8 multichannel microtensiometer based on the du Noüy-Padday method as described in the SI.[72] See SI for a derivation of equation expressing optimal salinity (S^*) expressed as a function of the characteristic curvature (Cc_{Var}) of the variovoracins.

Cytotoxicity and antimicrobial assays

See SI.

© 2026 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Genome sequencing and bioinformatic analysis

Genomic DNA was isolated from RKNM0096 using the Ultra-Clean® Microbial DNA Isolation Kit according to the manufacturer's recommendations (Mo Bio, Carlsbad, CA, USA). Genome sequencing was performed by the McGill University and Genome Quebec Innovation Centre (Montreal, QC, CA) using 2 SMRT Cells in a PacBio RSII sequencer (Pacific Biosciences, Menlo Park, CA, USA). A total of 140 476 raw subreads with an average length of 11 269 bp were generated and genome assembly was achieved using a HGAP workflow.[83] A detailed description of the genome assembly and annotation methodology is provided in the SI.

Characterization of RlpE enzymatic activity

The predicted dTDP-rhamnosyltransferase encoding rlpE gene was cloned into pET28a and expressed as and N-terminal histidine-tagged protein and purified by Ni-affinity chromatography. The sequence of the expressed protein is available under GenBank accession number (PX521120). Cloning, expression, protein purification and enzymatic characterization is described in the SI.

Conclusions

Herein we present the isolation and characterization of a new family of glycolipopeptide biosurfactants, which we have named the variovoracins. Although the variovoracins share structural features with the anionic RLs, the incorporation of a dipeptide moiety drastically alters the functional properties of the variovoracins by rendering them non-ionic. Application of the HLD model to characterize the surfactant properties of the variovoracins revealed they had a high positive Cc (+5.2), which clearly differentiates them from the RLs which have a negative Cc. This data indicates that the variovoracins would perform well as sustainable alternatives to high HLD surfactants such as sucrose esters and sulfosuccinates and could function as a complementary surfactant in mixed surfactant formulations.

Analysis of the *V. paradoxus* RKNM0096 genome led to the identification of a putative variovoracin BGC. Involvement of this locus in variovoracin production was provided by demonstrating that RlpE, a predicted rhamnosyltransferase, efficiently converted the predicted mono-rhamnosylated variovoracin substrate **6** to the dirhamnosylated **1**. Genome mining was utilized to provide further evidence to support the identification of the variovoracin BGC by identifying homologous gene clusters in publicly available genome sequences and then demonstrating production of the predicted variovoracin congener in two strains. Genomic analysis also established the distribution of related variovoracin-like BGCs in the *Pseudomonadota*. Collectively, the genomic analyses set the stage for extending the chemical diversity of the variovoracin family by targeted genetic modification of the *V. paradoxus* BGC or by isolation of new congeners from bacteria harbouring related BGCs. These efforts may lead to the discovery of novel variovoracin analogs with different surfactant properties or other biological activities.

With nearly 20 million tonnes of surfactants used annually, the replacement of synthetic surfactants with more sustainable and environmentally friendly biosurfactants is essential to reduce the negative impacts of surfactant use on the environment. The identification of new biosurfactants with different surface-active properties is an important strategy to reduce societal reliance on synthetic surfactants and the discovery of the variovoracins provides a new tool in the biosurfactant toolbox. To be truly viable commercial alternatives to synthetic surfactants, biosurfactants need to be produced at very high titers using highly efficient and cost-effective bioprocesses to be cost-competitive with synthetic surfactants. To develop the variovoracins into a viable commercial biosurfactant, future research will utilize the bioprocess and genomic groundwork laid by this study to achieve industrially-relevant yields.

Conflicts of interest

BH, DM, NM, ND, JK, AG, ML, KM, and HC are employees of Croda Canada Ltd, which is a subsidiary of Croda International Plc. BH, DM, FB and RK are inventors of patents EP3440179B1, US11608352B2, and US12404294B2 which are assigned to Croda International Plc.

Data availability

Mass spectrometry and NMR data used to establish the structure of variovoracin analogs is provided in the supplementary information (SI). The genome sequence of *Variovorax paradoxus* RKNM0096 is available in GenBank under accession number NZ_CP046508.1. Other sequence data used in this study were obtained from publicly available data held in GenBank under accession numbers that are provided in the supplementary information. Supplementary information: general methods and methods and data for bacterial isolation, stereochemical determinations, CMC determination, HLD determination cytotoxicity and antimicrobial testing, genome sequencing, bioinformatic analyses and RlpE characterization. See DOI: https://doi.org/10.1039/d6ra00033a.

Acknowledgements

We gratefully acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada, Canada Foundation for Innovation, Innovation PEI, the Jeanne and Jean-Louis Lévesque Foundation and Croda Canada Ltd. We thank Gwenola Le Mouee, Jose Ignacio Martinez Sanchez and Sander van Loon of Van Loom Chemical Innovations for performing Cc measurements. We also thank Dr Christopher Kirby and Maike Fischer of Agriculture and Agri-Food Canada for NMR services, Dr Bettina Sommers for assistance cloning *rlpE*, Dr Jonathan Booth of Croda for lending his expertise in surfactant chemistry and Prof. Dr Thomas Brück, Werner Siemens Chair of Synthetic Biology at the School of Natural Sciences of the Technical University of Munich, Germany for helpful discussions.

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Notes and references

1 G. O. Reznik, P. Vishwanath, M. A. Pynn, J. M. Sitnik, J. J. Todd, J. Wu, Y. Jiang, B. G. Keenan, A. B. Castle, R. F. Haskell, T. F. Smith, P. Somasundaran and K. A. Jarrell, *Appl. Microbiol. Biotechnol.*, 2010, **86**, 1387.

2 Surfactants Market by Type (Anionic, Non-ionic, Cationic, Amphoteric, and Others) and Application (Household Detergents, Personal Care, Industrial & Institutional Cleaners, Food Processing, Oilfield Chemicals, Agricultural Chemicals, Textiles, Emulsion Polymerization (Plastics), Paints & Coatings, Construction, and Others): Global Opportunity Analysis and Industry Forecast, https://www.alliedmarketresearch.com/surfactant-market, accessed 2025-09-24, 2025.

3 Surfactants Market Size, Share & Industry Analysis, By Type (Anionic, Nonionic, Cationic, and Amphoteric), By Application (Home Care, Personal Care, Textile, Food & Beverages, Industrial & Institutional Cleaning, Plastics, and Others), and Regional Forecast, 2024-2032, https://www.fortunebusinessinsights.com/surfactants-market-102385, accessed 2025-09-24.

4 D. K. Santos, R. D. Rufino, J. M. Luna, V. A. Santos and L. A. Sarubbo, *Int. J. Mol. Sci.*, 2016, **17**, 401.

5 C. C. V. S. Nagtode, H. K. Ahmad Yasin, S. N. Mali, S. M. Tambe, P. Roy, K. Singh, A. Goel, P. D. Amin, B. R. Thorat, J. N. Cruz and A. P. Pratap, *ACS Omega*, 2023, **8**, 11674.

6 D. G. Hayes and G. A. Smith, *Biobased Surfactants: Overview and Industrial State of the Art*, OACS Press, Oxford, UK, 2nd edn, 2019.

7 N. S. Badrul Azhar, C. Leong Puan, N. Kamarudin, N. Aziz, S. Nurhidayu and J. Fische, *Global Ecol. Conserv.*, 2015, **3**, 553.

8 T. Kohler, L. K. Curty, F. Barja, C. van Delden and J. C. Pechere, *J. Bacteriol.*, 2000, **182**, 5990.

9 J. Tremblay, A. P. Richardson, F. Lepine and E. Deziel, *Environ. Microbiol.*, 2007, **9**, 2622.

10 M. E. Davey, N. C. Caiazza and G. A. O'Toole, *J. Bacteriol.*, 2003, **185**, 1027.

11 R. Kownatzki, B. Tummler and G. Doring, *Lancet*, 1987, **1**, 1026.

12 C. D. McClure and N. L. Schiller, *Curr. Microbiol.*, 1996, **33**, 109.

13 M. Van Gennip, L. D. Christensen, M. Alhede, R. Phipps, P. O. Jensen, L. Christophersen, S. J. Pamp, C. Moser, P. J. Mikkelsen, A. Y. Koh, T. Tolker-Nielsen, G. B. Pier, N. Hoiby, M. Givskov and T. Bjarnsholt, *APMIS*, 2009, **117**, 537.

14 R. A. Al-Tahhan, T. R. Sandrin, A. A. Bodour and R. M. Maier, *Appl. Environ. Microbiol.*, 2000, **66**, 3262.

15 W. H. Noordman and D. B. Janssen, *Appl. Environ. Microbiol.*, 2002, **68**, 4502.

16 E. Rosenberg and E. Z. Ron, *Appl. Microbiol. Biotechnol.*, 1999, **52**, 154.

17 R. Sen, *Adv. Exp. Med. Biol.*, 2010, **672**, 316.

18 E. J. Gudiña, J. A. Teixeira and L. R. Rodrigues, *Mar. Drugs*, 2016, **14**, 38.

19 S. A. Jackson, E. Borchert, F. O'Gara and A. D. Dobson, *Curr. Opin. Biotechnol.*, 2015, **33**, 176.

20 M. Ines and G. Dhouha, *Peptides*, 2015, **71**, 100.

21 T. J. P. Smyth, A. Perfumo, S. McClean, R. Marchant and I. M. Banat, in *Handbook of Hydrocarbon and Lipid Microbiology*, ed. K. N. Timmis, Springer-Verlag, Berlin Heidelberg, Germany, 2010, ch. 27, p. 3687.

22 V. U. Irorere, L. Tripathi, R. Marchant, S. McClean and I. M. Banat, *Appl. Microbiol. Biotechnol.*, 2017, **101**, 3941.

23 C. Bettenhausen, *Chem. Eng. News*, 2020, 98.

24 B. Satola, J. H. Wubbeler and A. Steinbuchel, *Appl. Microbiol. Biotechnol.*, 2013, **97**, 541.

25 Y. H. Jiang, C. C. Yin, L. L. Yang, Y. H. Xin, Q. Liu and J. Ye, *Int. J. Syst. Evol. Microbiol.*, 2025, **75**, 006895.

26 O. M. Finkel, I. Salas-Gonzalez, G. Castrillo, J. M. Conway, T. F. Law, P. Teixeira, E. D. Wilson, C. R. Fitzpatrick, C. D. Jones and J. L. Dangl, *Nature*, 2020, **587**, 103.

27 R. Garcia Teijeiro, A. A. Belimov and I. C. Dodd, *New Biotechnol.*, 2020, **56**, 103.

28 C. Kurth, S. Schieferdecker, K. Athanasopoulou, I. Seccareccia and M. Nett, *J. Nat. Prod.*, 2016, **79**, 865.

29 C. W. Johnston, M. A. Skinnider, M. A. Wyatt, X. Li, M. R. Ranieri, L. Yang, D. L. Zechel, B. Ma and N. A. Magarvey, *Nat. Commun.*, 2015, **6**, 8421.

30 A. W. Robertson, N. G. McCarville, L. W. MacIntyre, H. Correa, B. Haltli, D. H. Marchbank and R. G. Kerr, *J. Nat. Prod.*, 2018, **81**, 858.

31 S. H. Yoon, S. M. Ha, S. Kwon, J. Lim, Y. Kim, H. Seo and J. Chun, *Int. J. Syst. Evol. Microbiol.*, 2017, **67**, 1613.

32 S. Hessa, K. R. Gustafson, D. J. Milanowski, A. Edgardo, M. A. Lipton and L. K. Pannell, *J. Chromatogr. A*, 2004, **1035**, 211.

33 T. Tanaka, T. Nakashima, T. Ueda, K. Tomii and I. Kouno, *Chem. Pharm. Bull.*, 2007, **55**, 899.

34 H. Spahn, *Arch. Pharm.*, 1988, **321**, 847.

35 H. Spahn, W. Henke, P. Langguth, S. Jürgen and E. Mutschler, *Arch. Pharm.*, 1990, **323**, 465.

36 A. M. Abdel-Mawgoud, F. Lépine and E. Déziel, *Appl. Microbiol. Biotechnol.*, 2010, **86**, 1323.

37 K. Blin, S. Shaw, H. E. Augustijn, Z. L. Reitz, F. Biermann, M. Alanjary, A. Fetter, B. R. Terlouw, W. W. Metcalf, E. J. N. Helfrich, G. P. van Wezel, M. H. Medema and T. Weber, *Nucleic Acids Res.*, 2023, **51**, W46.

38 S. E. Maddocks and P. C. F. Oyston, *Microbiology*, 2008, **154**, 3609.

39 M. A. Fischbach and C. T. Walsh, *Chem. Rev.*, 2006, **106**, 3468.

40 B. O. Bachmann and J. Ravel, *Methods Enzymol.*, 2009, **458**, 181.

41 F. Kopp, U. Linne, M. Oberthür and M. A. Marahiel, *J. Am. Chem. Soc.*, 2008, **130**, 2656.

42 F. I. Kraas, V. Helmetag, M. Wittmann, M. Strieker and M. A. Marahiel, *Chem. Biol.*, 2010, **17**, 872.

43 N. Ziemert, S. Podell, K. Penn, J. H. Badger, E. Allen and P. R. Jensen, *PLoS One*, 2012, **7**, e34064.

© 2026 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 06 March 2026. Downloaded on 3/15/2026 2:10:48 PM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

44 J. J. May, T. M. Wendrich and M. A. Marahiel, *J. Biol. Chem.*, 2001, **276**, 7209–7217.
45 D. Konz and M. A. Marahiel, *Chem. Biol.*, 1999, **6**, R39–R48.
46 L. Du and L. Lou, *Nat. Prod. Rep.*, 2010, **27**, 255–278.
47 R. H. Baltz, *J. Ind. Microbiol. Biotechnol.*, 2014, **41**, 357–369.
48 D. A. Herbst, B. Boll, G. Zocher, T. Stehle and L. Heide, *J. Biol. Chem.*, 2013, **288**, 1991–2003.
49 C. Breton, L. Snajdrová, C. Jeanneau, J. Koca and A. Imberty, *Glycobiology*, 2006, **16**, 29R.
50 E. Déziel, F. Lépine, S. Milot and R. Villemur, *Microbiology*, 2003, **149**, 2005.
51 A. M. Abdel-Mawgoud, R. Hausmann, F. Lépine, M. M. Müller and E. Déziel, *Biosurfactants*, 2011, 13.
52 R. Rahim, L. L. Burrows, M. A. Monteiro, M. B. Perry and J. S. Lam, *Microbiology*, 2000, **146**, 2803.
53 W. Zhu and C. O. Rock, *J. Bacteriol.*, 2008, **190**, 3147–3154.
54 A. M. Abdel-Mawgoud, F. Lepine and E. Deziel, *Chem. Biol.*, 2014, **21**, 156.
55 Z. Hojati, C. Milne, B. Harvey, L. Gordon, M. Borg, F. Flett, B. Wilkinson, P. J. Sidebottom, B. A. Rudd, M. A. Hayes, C. P. Smith and J. Micklefield, *Chem. Biol.*, 2002, **9**, 1175.
56 H. K. Zane, H. Naka, F. Rosconi, M. Sandy, M. G. Haygood and A. Butler, *J. Am. Chem. Soc.*, 2014, **136**, 5615.
57 M. P. Kem, H. K. Zane, S. D. Springer, J. M. Gauglitz and A. Butlera, *Metallomics*, 2014, **6**, 1150.
58 P. Brünker, O. S. Sterner, J. E. Bailey and W. Minas, *Antonie van Leeuwenhoek*, 2001, **79**, 235.
59 A. Angelini, L. Cendron, S. Goncalves, G. Zanotti and L. Terradot, *Proteins*, 2008, **72**, 1212.
60 S. Z. Neoh, F. Chek, H. T. Tan, J. A. Linares-Pastén, A. Nadakumar, T. Hakoshima and K. Sudesh, *Curr. Res. Biotechnol.*, 2022, **4**, 87.
61 D. Dubeau, E. Déziel, D. E. Woods and F. Lépine, *BMC Microbiol.*, 2009, **9**, 263.
62 K. Graupner, K. Scherlach, T. Bretschneider, G. Lackner, M. Roth, H. Gross and C. Hertweck, *Angew. Chem., Int. Ed.*, 2012, **51**, 13173.
63 Y. L. Chen, Y. H. Chen, Y. C. Lin, K. C. Tsai and H. T. Chiu, *J. Biol. Chem.*, 2009, **284**, 7352.
64 J. R. Edwards and J. A. Hayashi, *Arch. Biochem. Biophys.*, 1965, **111**, 415.
65 D. Dubeau, E. Deziel, D. E. Woods and F. Lepine, *BMC Microbiol.*, 2009, **9**, 263.
66 A. M. Abdel-Mawgoud, F. Lepine and E. Deziel, *Appl. Microbiol. Biotechnol.*, 2010, **86**, 1323.
67 T. Hirayama and I. Kato, *FEBS Lett.*, 1982, **139**, 81–85.
68 J. Andra, J. Rademann, J. Howe, M. H. Koch, H. Heine, U. Zahringer and K. Brandenburg, *Biol. Chem.*, 2006, **387**, 301.
69 J. Chen, J. Sun, R. W. Deering, N. DaSilva, N. P. Seeram, H. Wang and D. C. Rowley, *Org. Lett.*, 2016, **18**, 1490.
70 W. R. Abraham, H. Meyer and M. Yakimov, *Biochim. Biophys. Acta*, 1998, **1393**, 57.
71 D. Schulz, A. Passeri, M. Schmidt, S. Lang, F. Wagner, V. Wray and W. Gunkel, *Z. Naturforsch., C:J. Biosci.*, 1991, **46**, 197.
72 P. Suomalainen, C. Johans, T. Soderlund and P. K. J. Kinnunen, *J. Med. Chem.*, 2004, **47**, 1783.
73 E. J. Acosta, J. S. Yuan and A. S. Bhakta, *J. Surfactants Deterg.*, 2008, **11**, 145.
74 I. J. A. Baker, R. I. Willing, D. N. Furlong, F. Grieser and C. J. Drummond, *J. Surfactants Deterg.*, 2000, **3**, 13.
75 T. Madsen, G. Petersen, C. Seierø and J. Tørsløv, *J. Am. Oil Chem. Soc.*, 1996, **73**, 929.
76 S. M. Mudge and P. C. DeLeo, *Environ. Sci.:Processes Impacts*, 2014, **16**, 74.
77 T. T. Nguyen and D. A. Sabatini, *J. Surfactants Deterg.*, 2009, **12**, 109.
78 S. Hwang, N. Lee, S. Cho, B. Palsson and B.-K. Cho, *Front. Mol. Biosci.*, 2020, **7**, 87.
79 N. Mishra, K. Rana, S. D. Seelam, R. Kumar, V. Pandey, B. P. Salimath and D. Agsar, *Front. Bioeng. Biotechnol.*, 2021, **9**, 761266.
80 M. Hayakawa and H. Nonomura, *J. Ferment. Technol.*, 1987, **65**, 501.
81 E. B. Shirling and D. Gottlieb, *Int. J. Syst. Evol. Microbiol.*, 1966, **16**, 313.
82 D. G. Cooper and B. Goldenberg, *Appl. Environ. Microbiol.*, 1987, **53**, 224.
83 C. S. Chin, D. H. Alexander, P. Marks, A. A. Klammer, J. Drake, C. Heiner, A. Clum, A. Copeland, J. Huddleston, E. E. Eichler, S. W. Turner and J. Korlach, *Nat. Methods*, 2013, **10**, 563.

© 2026 The Author(s). Published by the Royal Society of Chemistry