View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Ulster University's Research Portal

27 Isolation and Analysis of Lipopeptides and High Molecular Weight Biosurfactants

*T. J. P. Smyth · A. Perfumo · S. McClean · R. Marchant · I. M. Banat**
School of Biomedical Sciences, University of Ulster, Coleraine, County Londonderry, Northern Ireland, UK
*IM.Banat@ulster.ac.uk

1 Introduction 3689

2 Experimental Approach 3691
2.1 Extraction: The Aim is to Obtain a Crude Extract Free from Aqueous Culture Medium 3691
2.1.1 *Bacillus* Lipopeptides 3691
2.1.2 *Pseudomonas* Lipopeptides 3692
2.1.3 Lipoproteins and Polymerics Extraction 3692
2.2 Detection/Purification 3693
2.2.1 TLC of Lipopeptides 3693
2.2.2 HPLC-UV/HPLC-MS for Purification of Different Types of Peptides/Lipopeptides 3693
2.2.3 One Dimensional Sodium Dodecyl Sulphate Polyacrylamine Gel Electrophoresis (1D SDS-PAGE) 3694
2.3 Analysis 3695
2.3.1 Bradford Assay for Protein Determination 3695
2.3.2 Fatty Acid Analysis by GC-MS 3696
2.3.3 Carbohydrate Determination for High Molecular Weight Biosurfactants 3697
2.3.4 Colorimetric Assay (Phenol-Sulphuric Acid) for Carbohydrate Determination 3697
2.3.5 GC-MS Determination of Carbohydrate Portion 3697
2.3.6 Protein Digestion with Proteases 3698
2.3.7 MALDI-TOF Mass Spectrometry 3699
2.3.8 Edman Degradation for Amino Acid Sequence 3699
2.3.9 Capillary Chromatography with Ion-Trap/QTOF MS/MS for *De Novo* Sequencing 3700

K. N. Timmis (ed.), *Handbook of Hydrocarbon and Lipid Microbiology*, DOI 10.1007/978-3-540-77587-4_290,
© Springer-Verlag Berlin Heidelberg, 2010

3 Solutions and Materials 3701

4 Time Considerations 3702

5 Trouble Shooting 3703

6 Research Needs 3703

Abstract: High molecular weight biosurfactants are produced from a number of different bacteria and comprise lipoproteins, proteins, polysaccharides, lipopolysaccharides or complexes containing several of these structural types, many of which have yet to be fully characterized. Lipopeptide biosurfactants are cyclic compounds and are most commonly isolated from *Bacillus* and *Pseudomonas* strains. A great deal of research has been carried out on these cyclic lipopeptides and their structures have been fully characterized. The ability to isolate, purify and characterize these structures is extremely important, providing detailed information with regard to different cultivation condition and biological activities. Similar methods can be used for both lipopeptides and biopolymers especially when attempting to determine their amino acid sequences. The experimental techniques used to isolate, purify and analyses these biosurfactant are widely varied from simple colorimetric assays giving an approximate indication of the type of compounds present to the more complex mass spectrometric techniques that provide information on molecular mass and structural features. Mass spectrometry provides essential information in the identification of these structures using sophisticated MS/MS experiments and software technologies. The use of Edman degradation can help confirm the results observed from *de novo* sequencing experiments, thus providing further confirmation of amino acid sequence of both lipopeptides and protein containing biopolymers. This paper details information on the most commonly reported techniques used to analyze these types of biosurfactants.

1 Introduction

Biosurfactants are amphipilic compounds with both hydrophilic and hydrophobic moieties (mostly hydrocarbons) able to display a variety of surface activities that, among other roles, help solubilize hydrophobic substrates (Desai and Banat, 1997). During the last decade they have been under investigation as potential replacements for synthetic surfactants and are expected to have many potential industrial and environmental applications (Banat et al., 2000; Kosaric, 1993). They generally can be grouped either as low or high molecular weight biosurfactants; the former consisting of glycolipids and lipopetides and the latter of high molecular weight polymeric biosurfactants. Glycolipids were discussed in Chapter 28, Vol. 5, Part 2 while in this paper the focus is on the analysis of lipopeptides and high molecular weight biosurfactants mainly using proteomic based techniques.

Lipopeptides are cyclic structures generally produced by *Bacillus* and *Pseudomonas* species. They mainly consist of hydrophilic peptides, usually between 7 and 10 amino acids long, linked to a hydrophobic fatty acid structure. *Bacillus* cyclic lipopeptides consist of three major groups known as the surfactin, iturin and fengycin families. Surfactin (*Fig. 1a*) is the most commonly studied lipopeptide and its structure consists of a 7 amino acid cyclic sequence connected to a C13–C16 fatty acid (Kakinuma et al., 1969). Iturin also consists of 7 amino acids, though different to surfactin, linked to C14–C17, while fengycin is composed of ten amino acids with fatty acid chain length of C14–C18 (Lang, 2002). *Pseudomonas* lipopeptides consist of a fatty acid attached to a peptide, which is cyclised to form a lactone ring between two amino acids in the peptide chain. *Pseudomonas* cyclic lipopeptides fall into four distinct classification, namely the widely studied viscosin and amphisin groups along with tolaasin and syringomycin families (Raaijmakers et al., 2006). The viscosin group's general structure consists of a 9 amino acid peptide linked to a 3-hydroxy fatty acid, usually decanoic acid

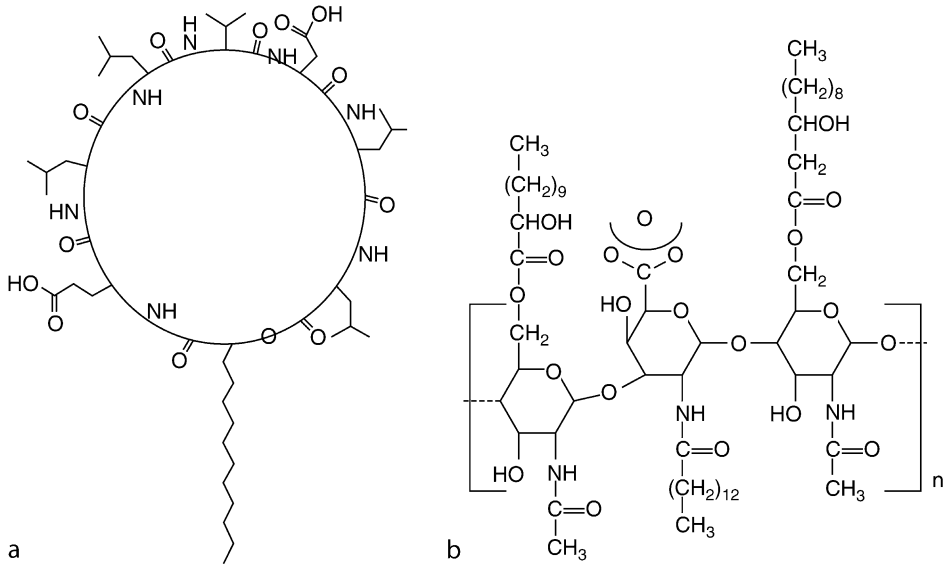

Figure 1

The structure of cyclic lipopeptide surfactin (a) from *Bacillus* sp. and a high molecular weight biopolymer emulsan (b) from *Acinetobacter* RAG-1.

while amphisin has an 11 amino acid peptide also linked to 3-hydroxy fatty acid (Sorenson et al., 2001). A number of other *Pseudomonas* lipopeptides, not belonging to any of these families have also been characterized, such as putisolvin I and II (Kuiper et al., 2004).

High molecular weight biosurfactants are generally grouped together as polymeric biosurfactants. They are produced by a number of different bacteria and are composed of lipoproteins, proteins, polysaccharides, lipopolysaccharides or complexes containing several of these structural types (Ron and Rosenberg, 2001). The most commonly studied biopolymers are Emuslan and Alasan isolated from *Acinetobacter* species. Emulsan (*Fig. 1b*) is a lipopolysaccharide isolated from *Acinetobacter calcoacetius* RAG-1 ATCC 31012 and has a molecular weight of around 1,000 kDa (Rosenberg et al., 1979) while Alasan is a complex of an anionic polysaccharide and a protein with a molecular weight of around 1,000 kDa isolated from *Acinetobacter radioresistens* (Navon-Venezia et al., 1995). A number of other polymeric biosurfactants have been discovered but remain partially or totally uncharacterized.

Chemical and structural analysis of lipopeptides and high molecular weight biosurfactants is carried out using a broad range of techniques varying from simple colorimetric assays to sophisticated mass spectrometry (MS) and sequencing techniques. Due to their size and structural complexity biopolymers are much more difficult to characterize than the lower molecular weight lipopeptide biosurfactants. After isolation and purification procedures, molecular mass determination of the compounds of interest may be facilitated by mass spectrometry. This should be followed by analysis of the fatty acid portion and determination of the peptide sequence using automated Edman degradation sequencing and a number of mass spectrometric techniques. This combined approach would provide the necessary information required for complete structural identification.

The analysis of biopolymeric biosurfactants presents a number of challenges due to their high molecular mass. Generally, identification of the relative percentage of the carbohydrate, lipid and protein portions is carried out using simple colorimetric assays, such as Bradford assay for protein determination. Such information is particularly useful when the biopolymer has not been previously characterized, as it provides information on the techniques required for structural determination. For example, if the biopolymer contains only lipids and proteins, it is not necessary to carry out further analysis for polysaccharides. Determination of the approximate molecular mass can be carried out using sodium dodecyl sulphate polyacrylamine gel electrophoresis (SDS-PAGE) and matrix assisted laser desorption ionization time of flight mass spectrometry (MALDI-ToF MS). However, for more detailed structural analysis to determine the exact composition requires the molecule to be broken down into more manageable portions. Removal of the carbohydrate and fatty acid portions means that these compounds can be analyzed by gas chromatography mass spectrometry (GC-MS) after derivatisation. Structural analysis of the protein portion of biopolymers requires its digestion into smaller peptides by enzymatic degradation. In this way the individual sequences of the smaller cleaved peptides may be carried out using Edman degradation and mass spectrometric techniques. The information obtained can then be used to piece together the full protein structure.

The methods described here represent a variety of common techniques that can be used for analysis of lipopeptides and high molecular weight biosurfactants. Due to the less complex nature of lipopeptides, techniques such as SDS-PAGE and digestions are not required for their analysis. In contrast these techniques are essential for polymeric biosurfactants allowing the smaller components obtained after digestion to be analyzed in a similar manner to lipopeptides. Some of the techniques can be carried out without the need for complex equipment, while others such as MS, use expensive equipment and require a specific expertise. This chapter is therefore intended to provide insights and suggestions on how to approach the analysis of these biosurfactants. Further information on each procedure for analysis and information regarding interpretation of results can be followed up in the relevant literature cited.

2 Experimental Approach

2.1 Extraction: The Aim is to Obtain a Crude Extract Free from Aqueous Culture Medium

2.1.1 *Bacillus* Lipopeptides

This extraction technique is a combination of acid precipitation and solvent extraction (Vater et al., 2002)

1. To remove cells centrifuge at 13,000 × g for 15 min at 4°C
2. Acidify by addition of concentrated HCl to pH 2.0 and allow precipitate to form at 4°C overnight
3. Centrifuge at 13,000 × g for 15 min at 4°C to obtain pellet
4. Remove supernatant and extract the pellet with methanol for 2 h while stirring continuously
5. Filter methanol to remove remaining material and evaporate to dryness using rotary evaporation

2.1.2 *Pseudomonas* Lipopeptides

The extraction method reported above could be used for extraction of *Pseudomonas* lipopeptides, however, the most commonly reported method for this group of lipopeptides is solvent extraction with ethyl acetate (Kuiper et al., 2004).

1. Extract three times with an equal volume of ethyl acetate[1] shaking vigorously each time and allow the two layers to separate in a separating funnel
2. Transfer bottom aqueous layer and the top ethyl acetate layer to separate flasks. Re-extract the aqueous portion twice more with ethyl acetate layer
3. Add a small amount of magnesium sulphate to the ethyl acetate portion, to remove the traces of water present, filter and rotary evaporate to dry extract

2.1.3 Lipoproteins and Polymerics Extraction

Extraction of the majority of high molecular weight biosurfactants is carried out using ammonium sulphate precipitation, followed by dialysis to remove any small molecules that may be present (Rosenberg et al., 1979). The methods reported for isolation of the high molecular biosurfactants are quite varied and generally specific to the actual biosurfactant present. Ammonium sulphate precipitation can be used for the majority of extractions, however, for more specific details refer to related literature of the microorganism under investigation. Other techniques for high molecular weight biosurfactant isolation include TCA/acetone precipitation, acid ethanol and chloroform/methanol.

The ammonium sulphate precipitation method below is based on using 100 ml of culture broth and should be adjusted according to the starting volume.

1. Take 100 ml of culture broth and remove cells by centrifuging at 10,000 × g for 15 min
2. Cool the supernatant at 4°C and add slowly while stirring 23.34 g of ammonium sulphate to obtain a 40% saturation solution.[2] Leave at 4°C overnight
3. Centrifuge at 10,000 × g for 15 min to obtain pellet
4. Re-suspend in a 40% solution of ammonium sulphate and centrifuge again at 10,000 × g for 15 min to obtain pellet
5. Dissolve pellet in 20 ml water and extract with equal volume of hexane, three times in a separating funnel to remove residual non-polar lipids
6. Using dialysis tubing with a molecular weight cut off point of 5 kDa clamp one end and rinse the tubing with distilled water to check for leaks
7. Fill tubing with product from step 5 and clamp top. Place in beaker with dialysis buffer of distilled water, place on a stirring plate in a cold room overnight
8. Change dialysis buffer once after 6 h
9. Remove sample after completion of dialysis and lyophilize to obtain biosurfactant product

[1] A small amount of formic acid (1%) can also be added for extraction.
[2] Higher saturation levels maybe required for precipitation of some proteins. Refer to literature of the specific biopolymer under investigation.

2.2 Detection/Purification

A number of impurities are often co-extracted together with the target biosurfactant and may need to be separated.

2.2.1 TLC of Lipopeptides

Thin layer chromatography (TLC) is a simple method, which can be used to detect the presence of lipopeptides while preparative TLC can be used to purify small quantities (Symmank et al., 2002).

1. Dissolve a small quantity of crude extract in chloroform and apply 10 µl onto a TLC plate (silica gel 60) and apply at point of origin near the bottom of the plate
2. Once dried, develop plate in solvent system of chloroform:methanol:water (65:25:4)[3]
3. When developed remove plate and allow to air-dry in a fume cupboard
4. Spray the plate evenly with water or a solution of 5% sulphuric acid and place in an oven at 110°C for 20 min to visualize spots
5. At preparative scale remove lipopeptide spots by scraping the silica from plate into a flask and extract with chloroform:methanol (2:1) overnight
6. Filter and rotary evaporate to dryness

2.2.2 HPLC-UV/HPLC-MS for Purification of Different Types of Peptides/Lipopeptides

HPLC is an excellent method for the separation of individual peptide/lipopeptide biosurfactants (Aguilar, 2004). The most commonly employed technique is reversed phase chromatography, which results in the separation of each peptide/lipopeptide structure based on polarity. The separated products are detected by UV absorbance detection and each individual peak can be collected using a fraction collector for further analysis of their structure. Coupling of HPLC with a mass spectrometer provides preliminary information on the molecular mass of each component. Purification with either HPLC-UV or HPLC-MS using different types of column chemistry is also possible. For the purpose of this report the basic steps in HPLC-MS will be discussed below.

1. Interface HPLC to electrospray ionisation mass spectrometry (ESI-MS) system
2. Use semi-preparative HPLC column such as Luna C18 column (250 mm × 10 mm id) (Phenomenex, Cheshire, UK) connected[4]
3. Prepare the sample by dissolving in water/TFA (99.95:0.05, v/v). Centrifuge at 13,000 rpm for 5 min to remove particulate matter

[3] Other TLC solvents could be used for analysis. This system is for general lipopeptide use, however more specific systems can be found in the literature for particular strains.

[4] Using a C18 column provides a good starting point for analysis and generally should result in a good separation, a large variety of other columns could also be used.

4. Prepare mobile phase A (water/TFA (99.95:0.05, v/v)) and B (acetonitrile/water/TFA (80:19.95:0.05)) and perform gradient elution starting with 100% A, 0% B changing to 0% A, 100% B over 80 min[5]
5. Set flow rate at 1.0 ml min^{-1} and injection volume of 500 µl min^{-1}
6. Connect fraction collector and set up the effluent flow from the column so that 90% is fraction collected and the remaining 10% sent to the MS detector
7. Collect fractions at a rate of 1 ml per minute. Fraction collection may be automated or peak collection may also be used
8. MS scan range should be set at m/z 150–2,000
9. MS/MS fragmentation can also be set up on the eluent entering the mass spectrometry. The information obtained can then be used to search a number of databases
10. The individual fraction can then be subjected to Q-TOF MS/MS *de novo* sequencing or Edman degradation to obtain sequence information on each individual peptide/lipopeptide

2.2.3 One Dimensional Sodium Dodecyl Sulphate Polyacrylamine Gel Electrophoresis (1D SDS-PAGE)

SDS-PAGE is commonly used for proteomic experiments and is extremely useful for separation of protein mixtures[6] and for estimating the molecular mass of high molecular weight biosurfactants/proteins. The proteins are applied to one end of a gel in a loading buffer that contains a reducing buffer and SDS. The reducing buffer is used to cleave the disulphide bonds rendering the proteins in a linear form. SDS attaches to the protein relative to its molecular mass, resulting in a net negative charge on the protein. When a voltage is applied to the gel, proteins separate according to molecular mass. The individual bands can then be removed and the proteins extracted (Toren et al., 2001).

1. Prepare 10% running gel[7] using the following; 12.3 ml of double distilled water; 7.5 ml of gel buffer which is 1.5 M Tris-HCl at pH 6.6 (6 g of Tris in 100 ml of water adjusted with HCl); 9.9 ml of acrylamide/Bis-acrylamide (30%/0.8% w/v); 0.3 ml of 10% SDS (w/v); 0.15 ml of 10% ammonium persulfate (APS w/v) and finally 0.02 ml of TEMED (*N*',*N*',*N*',*N*'-tetramethylethylethylenediamine). For the last two parts (APS and TEMED) of the running gel add just before pouring the gel
2. For preparation of stacking gel use the following; 3.05 ml of double distilled water; 1.25 ml of stacking buffer (0.5 M Tris-HCl at pH 6.8); 0.665 ml of acrylamide/Bis-acrylamide (30%/0.8% w/v); 50 µl of 10% SDS; 50 µl of 10% APS and 5 µl of TEMED.[8] For the last two parts (ABS and TEMED) of the stacking gel add just before pouring the gel

[5] Generally a number of trail runs should be carried out to optimise the column conditions and the run time required. Time given is just a rough guide but can be adjusted up or down depending on how the sample.
[6] Only when relatively simple mixture is present. When complex mixtures are being studied 2D gels are more appropriate.
[7] The % gel depends on the size of proteins. 10% is suitable for the range 20KDa to 300kDa approximately.
[8] The majority of these buffers can be purchased as ready made solutions.

3. Running buffer should be prepared as follows; 200 mM glycine, 25 mM Tris-HCl at pH 6.8 and 0.1% (w/v) SDS
4. Prepare Laemmli sample buffer (Laemmli, 1970) by using 0.02 g (w/v) of bromophenol blue; 0.313 ml of stacking buffer (0.5 M Tris-HCl at pH 6.8); 0.5 ml 10% (w/v) SDS; 0.25 ml (v/v) glycerol and make up entire solution to 2.5 ml with double distilled water
5. Set up gel electrophoresis as per standard procedure and pour the running gel prepared in step 1 with APS and TEMED being added just before pouring. Place 1 ml of water saturated with butanol carefully on top of monomer solution and leave for 45 min to allow the gel to polymerize
6. Pour the overlying water off and drain the excess with strips of filter paper
7. Add the stacking gel buffer with APS and TEMED being added just before pouring. Insert sample comb at an angle to allow air bubble to escape and leave to set for 15 min
8. Remove the comb and place gel in the buffer chamber and add running buffer to cover the gel
9. Prepare the protein sample to be investigated by adding 20 µl of proteins to 20 µl of sample buffer (step 4) and 1 µl of beta-mercapto-ethanol[9] (used to reduce the disulphide bonds). Incubate at 65°C for 10 min
10. Load 5 µl molecular weight marker and 10 µl of each sample into separate wells
11. Put on lid and connect power supply with 200 volts and run the gel. Stop the electrophoresis when the marker dye is less than 1 cm from bottom of plate (around 1 h)
12. Remove gel and immerse in Coomassie brilliant blue stain[10] (made by adding 1.25 g of Coomassie R-250, 250 ml methanol, 50 ml of acetic acid and 200 ml of water) and leave overnight
13. Pour off the Coomassie stain and destain using 50% methanol in water containing 10% acetic acid. Cover gel with destain solution and leave for 10 min then remove
14. Record image of gel and determine approximate molecular mass of each protein or high molecular weight biosurfactant present by comparison with the molecular weight ladder

2.3 Analysis

2.3.1 Bradford Assay for Protein Determination

The Bradford assay is a relatively quick method for the determination of protein concentration in the high molecular weight biosurfactants. The assay is based on the binding of Coomassie Brilliant Blue with protein causing a color change that can be quantified using a spectrophotometer (Bradford, 1976).

1. To prepare the Bradford reagent, dissolve 100 mg of Coomassie Brilliant Blue G-250 in methanol and add 100 ml of 85% phosphoric acid. Dilute to 1 L with distilled water.
2. Prepare a series of standards (1 ml each) using Bovine Serum Albumin (BSA) with the following concentrations 0 (blank), 250, 500, 1,000, 1,500 and 2,000 µg ml^{-1}.[11]

[9] Dithiothreitol (DTT) can also be used.
[10] Other stains can also be used.
[11] The method detailed is for a standard assay, scaling down can be carried out for a microassay. For microassay (1–10µg of protein) and use standard concentrations of 1, 5, 7.5, 10, 15 and 20µg of protein.

3. Prepare serial dilutions of unknown samples to be assayed.
4. To assay add 1 ml of Bradford reagent and 20 µl of each of the standards into separate 1 ml plastic cuvettes.[12] Prepare the unknown samples in the same manner, adding 1 ml of Bradford reagent and 20 µl into separate 1 ml plastic cuvettes.
5. Mix well and measure the absorbance at 595 nm wavelength, using the blank to zero.
6. Measure the standards and use the results prepare a standard curve. Measure the unknown sample and use the standard curve to calculate the unknown protein concentrations.

2.3.2 Fatty Acid Analysis by GC-MS

Lipopeptides and high molecular weight biosurfactants vary greatly in the lipid portion, therefore the analysis of this component provides detailed structural information for identification of fatty acid structures. The process involves hydrolytic cleavage of the link between the carbohydrate or peptide/protein part of the biosurfactant and the lipid portions. Subsequent derivatization of the resulting fatty acid chains to fatty acid methyl esters (FAME) and further conversion to trimethylsilyl (TMS) derivatives facilitates analysis by GC or GC-MS (Yakimov et al., 1995).

1. Take 5 mg of purified glycolipid and dry completely.
2. To this add 0.95 ml of methanol and 0.05 ml of sulphuric acid in a sealed tube and heat at 90°C for 15 h.
3. Evaporate the solvents and extract the residue with 1 ml of hexane.
4. Wash with 1 ml of water and evaporate the hexane to obtain extract containing the FAME.[13]
5. To prepare TMS ether derivatives, add 0.5 ml of TMSI (N-trimethylsilylimidazole)[14] to 5–10 mg of dried sample from step 4 in a vial and heat at 60°C for 10–15 min.
6. GC-MS analyses can be carried out using a variety of columns such as DB-23 capillary column or supelco omegawax (Sigma-Aldrich, Dorset, UK) (30 M × 0.25 mm × 0.25 µm id).
7. Analytical conditions that can be used are; injector temperature 250°C, start oven program at 50°C, hold for 1 min then ramp at 40°C per min to 250°C and hold for 25 min.[15]
8. Mass spectrometry settings; electron impact at 70 eV with scan range 50–650 u and an injection volume of 1 µl.
9. Fatty acid methyl esters retention times increase with chain length and degree of separation.

[12] For microassay preparation use 1ml of Bradford reagent and 100µl of standard or sample.
[13] The FAME can be analysed by GC or GC-MS at this stage, however further derivatisation to TMS derivatives provides better sensitivity for analysis.
[14] TMS derivatives can be formed with a variety of different silylation reagents.
[15] As well as a variety of columns that can be used for analyses, the other conditions such as run temperatures and times can also be varied.

2.3.3 Carbohydrate Determination for High Molecular Weight Biosurfactants

The analysis of the carbohydrate portion of glycoproteins, lipopolysaccharides and other lipid containing polysaccharide-protein complexes can be carried out using a variety of techniques. However, total characterization of the exact structural conformation and linkage sites is extremely difficult. A two-stage approach maybe adopted where the relative percentage of the carbohydrate portion of the purified biosurfactant is first determined by colorimetric assay. In the second step the composition of the individual monosaccharides present is determined using GC-MS.

2.3.4 Colorimetric Assay (Phenol-Sulphuric Acid) for Carbohydrate Determination

The carbohydrate portion needs to be removed and hydrolyzed into individual monosaccharides before the assay can be carried out. A number of different methods can be used to determine the carbohydrate content.

1. Hydrolyze the purified biosurfactant product by adding 20 μg to 200 μl of 0.5 M methanolic HCl in a sealed vial. Heat at 65°C for 16 h.
2. Dry down using a stream of nitrogen.
3. Prepare phenol reagent by dissolving 5 ml in 95 ml of water.
4. Prepare a blank by mixing 100 μl of water to 200 μl of phenol reagent. For samples, dissolve 5 mg in 1 ml of water, take 100 μl and add 200 μl of phenol. Prepare a series of standards by dissolving glucose and 1 mg ml^{-1}. Take μl, 40 μl, 60 μl, 80 μl and 100 μg make up to 100 μl with water and add 200 μl of phenol reagent to each standard (Chaplin, 1986).
5. Add 1 ml of concentrated sulphuric acid and leave solution to set for 10 min.
6. Shake vigorously and leave for 30 min.
7. Determine the absorbance at 490 nm and compare against the range of standards prepared in step 4.

2.3.5 GC-MS Determination of Carbohydrate Portion

After hydrolysis of the carbohydrate portion into individual monosaccharide they can be converted to TMS derivatives for analysis by GC-MS (Manzi and Varki, 1993).

1. Hydrolyise the purified biosurfactant product by adding 20 μg to 200 μl of 0.5 M methanolic HCl sealed vial. Heat at 65°C for 16 h.
2. Dry down using a stream of nitrogen and wash with hexane to remove non-polar material.
3. Dissolve 10 mg of the residue with 0.1 ml of TMSI (N-trimethylsilylimidazole) and incubate at room temperature for 30 min.
4. Evaporate the mixture using nitrogen and dissolve in 1 ml of hexane.
5. GC-MS analyses can be carried out using a variety of columns such as DB-5 capillary column or supelco omegawax (Sigma-Aldrich, Dorset, UK) (30 M × 0.25 mm × 0.25 μm id).

6. Analytical conditions are; injector temperature 200°C, start oven program at 50°C, hold for 3 min then ramp at 20°C per min to 170°C and then to 250°C at 6°C per min and hold for 5.7 min.[16]
7. Mass spectrometry conditions; electron impact at 70 eV with scan range of m/z 50–650 and an injection volume of 1 μl for both samples and standards.
8. Identify the components by mass spectra and retention times.

2.3.6 Protein Digestion with Proteases

Using proteomic techniques it is not always possible to obtain sufficient information to identify the amino acid sequence of the intact protein structure. This is due to several reasons, such as an inability to obtain accurate measurements on very large proteins and the fact that Edman degradation and mass spectrometry techniques are more suited to smaller peptide analysis. Therefore digestion of the protein structure into smaller peptides containing between 6–20 amino acids is necessary. Using individual peptides, Edman degradation and MS/MS techniques can be used to provide the information required to identify their structure and help to piece together the full protein structure. The digestion of the protein is carried out using proteases, with trypsin being the most commonly employed. Trypsin cleaves proteins at lysine and arginine residues in the C-terminal direction except where a proline residue follows. It may also be necessary to use other proteases at a latter stage to help determine the order of each peptide sequence in the protein structure. For example, by using a protease such as glu-C, which cleaves at glutamate and asparatate residues, would provide a different set of cleaved peptides. Investigation of these peptides would provide information allowing overlapping portions of the sequence to be identified. The digestion described here is based on in-solution tryptic digestion (Strader et al., 2006), however in-gel digestion is also commonly used.

1. Prepare ammonium bicarbonate buffer by adding 400 mg to 100 ml of ultrapure water.
2. Prepare sample buffer of mixture of acetonitrile/water[17] in ammonium bicarbonate buffer containing 50 mM Tris/HCl (pH7.6) and 10 mM $CaCl_2$.
3. To make reduction solution, add 16 mg of dithiotreitol (DTT) in 500 μl ultrapure water to give concentration of approximately 200 mM. The alkylation solution is prepared by adding 9 mg of iodacetamide to 500 μl of ultra pure water give a concentration of 100 mM.
4. Prepare 10% formic acid solution.
5. Dissolve the purified protein band obtained from gel electrophoresis technique described in Section 2.2.3 in 40 μl of sample buffer made in step 2.
6. For reduction add 4 μl of 200 mM DTT prepare in step 3 and incubate at 37°C for 1 h.
7. For alkylation add 4 μl of iodacetamide solution prepared in step 3 and incubate for 15 min in the dark at room temperature.

[16] As well as a variety of columns that can be used, other conditions such as run temperatures and times can also be varied.

[17] Sample buffer should contain at maximum 80% acetonitrile, the value should depend on the protein solubility and should be checked beforehand.

8. To digest the protein add trypsin[18] in a ratio of trypsin/protein 1/5 and incubate for 2–4 h at 37°C.
9. To stop the digestion, add formic acid to a final concentration of 5% (v/v)
10. The sample can then be submitted for separation using HPLC as described in Section 2.2.2 to obtain individual peptides for MS/MS analysis. Alternatively the mixture can be purified to remove the buffers and used for MALDI-TOF analysis and capillary chromatography with ion trap/QTOF MS/MS
11. The sample can be lyophilized and stored at − 80°C.

2.3.7 MALDI-TOF Mass Spectrometry

MALDI is a soft ionisation mass spectrometry technique that allows the identification of intact compounds. Samples to be analyzed are mixed with a matrix and dried on a platform, onto which a laser is fired with various degrees of energy forming gaseous ions, which can then be separated in a time of flight analyzer (TOF) and detected. MALDI can be used to determine the full molecular mass of high molecular weight biosurfactants and lipopeptide biosurfactants that have been purified. It is also possible to analyze the individual peptide components obtained from the protease digestions and use the information obtained to search a database in a process known as peptide mass fingerprinting or peptide mass mapping. If the high molecular mass biosurfactant has been previously characterized the ions observed for protease digestion would be specific for this biosurfactant and therefore sufficient for identification

1. Prepare the matrix for peptide analysis, α-cyano-4-hydroxy cinnamic acid (αCHCA), which should be used for compounds up to 10,000 Da. Prepare a 10 mg ml^{-1} solution of αCHCA in acetonitrile/Ultrapure water/trifluoroacetic acid (50:50:0.1) and vortex until dissolved.
2. Prepare matrix for high molecular weight biosurfactants (>10,000) as for step one but using sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid).
3. Dissolve sample in methanol, mix 1 µl with 1 µl of matrix and apply to sample plate. Allow plate to dry at room temperature.
4. Insert plate into instrument, move plate to sample position and fire the laser.
5. If the signal is poor increase laser strength or move the laser to a different position on the well.[19]
6. To improve signal remove possible salts that maybe present by washing the sample on-plate with 2 µl of cold ultrapure water containing 0.1% TFA.

2.3.8 Edman Degradation for Amino Acid Sequence

Edman degradation is the classic technique for sequencing peptides using chemical methods and may be applied to peptides and proteins where the N-terminus has not been modified by,

[18] It may be necessary to use other proteases in order to obtain further information about order of the sequences obtained.
[19] Signal is severely influenced by the presence of salts and other impurities.

for example, pyroglutamic acid or acetylation (Zachara and Gooley, 2000). The method provides an assignment for each residue in the peptide, unlike amino acid analysis which only provides an indication of the ratio of AAs in the peptide. Edman experiments take place in an oxygen-free environment and involve the modification of the N-terminal residue with phenylisothiocyanate to provide a cleaved phenylthiohydantoin (PTH) amino acid. All of the chemical processes take place on automated sequencers and are followed by a chromatographic step where the retention time of the cleaved PTH amino acid is compared with the retention times of a series of PTH modified amino acid standards to ascertain its identity. The method is relatively slow taking approximately 45 min for each residue, however the amino acid assignment is called with a high degree of confidence and less subject to interpretation, as can be the case with MS/MS methods. Of particular relevance is the differentiation between isobaric leucine and isolueine which is not possible with low energy collision MS/MS, but readily achievable using Edman degradation. The quality of sequence information obtained by the Edman method is subject to the amount of starting material and its purity. For successful sequencing to take place, peptides and proteins must be purified to near homogeneity by chromatographic methods to prevent mixed sequencing during Edman experiments. Lipopeptides need to be in the open ring form for this type of analysis, which is carried out using mild alkaline hydrolysis.

1. Purify the protein or peptide of interest using chromatographic methods, described in Section 2.2.2 and 2.2.3, removing all the non-volatile salts and buffers that may cause unwanted side reactions.
2. Take a small sample of dried purified cyclic lipopeptides, dissolve in 1 M NaOH and allow to stand at room temperature for 12 h.
3. Neutralize the mixture with HCl and extract the linear form of the biosurfactant by extracting with chloroform.
4. Remove the salts that were used for opening of the cyclic peptide ring using HPLC as per step 2.2.3 and 2.2.3.
5. Apply the sample to the instrument using manufacturers instructions and carry out sequencing. Continue until no further sequence information is obtained.

2.3.9 Capillary Chromatography with Ion-Trap/QTOF MS/MS for *De Novo* Sequencing

The use of QTOF MS/MS and other high-resolution mass spectrometers such as orbitraps and FT-ICR instruments allow *de novo* sequencing of peptides where the mass accuracy capability of the instrument facilitates amino acid assignment of the peptide without recourse to a database Graham et al., 2007). The method is useful where the protein or peptide of interest is deemed novel due to unsuccessful sequencing attempts by database searching with MS/MS datasets. Protease digestion of large proteins must first occur as the technique is most successful for peptides less than 2,000 Da. Unlike Edman degradation, MS/MS experiments do not require the peptide to be purified to homogeneity as the mass analyzer may be used to select only the ions of interest. The analysis can be carried out using direct a infusion method or by capillary chromatography separation of individual peptides before they enter the mass spectrometer. Following fragmentation of the peptide, various proprietary software tools are available to process the raw data and analyze the sequence assigned. Frequently this is used in

conjunction with BLAST searching to determine if homology exists with existing precursor structures. Often low mass resolution instruments, in particular ion-trap technology, are used to obtain MS/MS data, which may be analyzed with appropriate algorithms to provide sequence assignment based on database searching. Well very successful, the resolution of the instrument does not allow for *de novo* sequencing. Advantages of the method include fast analysis times and the small sample volume required. Disadvantages of the method are that algorithms are used to assign a peptide sequence to product ion spectra and a certain level of expertise must be gained to critically analyze the data produced. Often de novo sequencing is used to identify sequence tags within peptide structure rather than the complete sequence. These tags may then be used to design primers for cDNA cloning studies that may further corroborate the mass spectrometric sequencing data. As with all MS equipment, the experimental technique will vary greatly depending on the manufacturer of the instrument and software, therefore, only specific details on sample preparation and basic operation procedures are described below and for specific details the software manuals should be consulted.

1. Following trypsin digestion for protein containing biosurfactants and after chromatographic purification of the peptides or lipopeptides
2. Prepare samples for analysis by dissolving in acetonitrile/water mixture
3. Samples may be analyzed by capillary chromatography mass spectrometry or by introducing the sample directly into the instrument using a spray emitter; according to the manufacturers instructions
4. Capillary liquid chromatography conditions; starting at 100% mobile phase A (HPLC water containing 0.1% formic acid) changing to 100% mobile phase B (acetonitrile/water/formic acid 80:20:0.1) over 1 h
5. Connect appropriate capillary column
6. Set flow rate at 0.3 μl min^{-1} with an injection volume of 5 μl

3 Solutions and Materials

3.1 Extraction of *Bacillus* lipopeptides – Centrifuge, centrifuge tubes, HCl, methanol and filter funnel and paper, rotary evaporator

3.2 Extraction of *Pseudomonas* lipopeptides – Separating funnel, ethyl acetate, magnesium sulphate, filter paper and funnel, rotary evaporator

3.3 Extraction of lipoproteins and polymeric biosurfactants – Centrifuge, ammonium sulphate, distilled water, hexane, separating funnel, Dialysis tubing with molecular weight cut-off point of 5,000 daltons, 2 clamps for tubing, large beaker, stirring plate

3.4 TLC of lipopeptides – TLC tank with lid, TLC plates (silica gel 60), chloroform, methanol, distilled water, fine point capillaries, sulphuric acid, oven, filter paper, rotary evaporator

3.5 HPLC-UV/HPLC-MS purification– HPLC system connected to either UV or MS detector or both, a semi-preparative HPLC column such as Luna C18 or other suitable column, HPLC water, acetonitrile and TFA, Fraction collector and tubes for collecting fractions

3.6 SDS-PAGE – General gel equipment, for running gel; double distilled water, 1.5 M Tris-HCl at pH 6.6, acrylamide/Bis-acrylamide (30%/0.8%), SDS, ammonium persulfate, TEMED (*N*',*N*',*N*',*N*'-tetramethylethylethylenediamine). For stacking gel; same as for running gel except 0.5 M Tris-HCl at pH 6.8. For the running buffer; glycine,

Tris-HCL, SDS. Sample buffer (Laemmli); bromophenol blue, 0.5 M Tris-HCL at pH 6.8, glycerol, SDS, double distilled water. Other reagents include water saturated with butanol, filter paper, molecular weight marker, beta-mercapto-ethanol, coomassie brilliant blue stain, destain containing methanol water and acetic acid

3.7 Bradford assay – Coomassie brilliant blue, methanol, phosphoric acid, distilled water, BSA, Sample tubes, spectrophotometer

3.8 Fatty acid analysis by GC-MS – GC-MS equipment with GC column such as DB-23 capillary column or supelco omegawax, methanol, sulphuric acid, sealed tubes, oven, hexane, distilled water, TMS ether derivative such as TMSI, rotary evaporator

3.9 Carbohydrate determination by colorimetric assay – Methanol, HCl, nitrogen supply, phenol, sulphuric acid, sample curvettes, spectrophotometer

3.10 Carbohydrate analysis by GC-MS – As per Section 3.8

3.11 Protein digestion with proteases – Ammonium bicarbonate, ultrapure water, acetonitrile, Tris-HCl at pH 7.6, calcium chloride, DTT (dithiothreitol), iodacetamide, formic acid, trypsin

3.12 MALDI – Mass spectrometry with MALDI soft ionisation, vortex, matrix (α-cyano-4-hydroxy cinnamic acid (αCHCA) for lipopeptides/peptides, sinapinic acid for large molecular weight biosurfactants, acetonitrile and ultrapure water

3.13 Edman degradation – Edman degradation sequencer and associated solvents and chemicals, acetonitrile, ultrapure water

3.14 Capillary chromatography with ion-trap/QToF MS/MS for *de novo* sequencing – Appropriate capillary liquid chromatography and connected to either a ion-trap or QToF mass spectrometry, Ultrapure HPLC water, HPLC acetonitrile, formic acid

4 Time Considerations

Time required to perform some of the procedures illustrated (e.g., extraction and purification) can be affected by several factors such as sample volume or presence of residual production substrates (oils in particular). The times stated represent a rough guideline for the experimental procedures. However, the time required for analyzing results, in particularly with mass spectrometry techniques is considerably longer.

4.1 Extraction of lipopetides – 2 days

4.2 Extraction of lipoproteins and polymeric biosurfactants – 2 days

4.3 TLC analysis – 3 h

4.4 HPLC For purification – 1–3 h per sample depending on the run time and level of separation required.

4.5 1D SDS-PAGE – approximately 2 days

4.6 Bradford assay – 2 h

4.7 Fatty acid analyses by GC – 1 day for sample preparation and 1 h for sample on GC-MS.

4.8 Carbohydrate determination by colorimetric assay – 2 days

4.9 Carbohydrate determination by GC-MS – 1 day for sample preparation and 1 h for sample on GC-MS

4.10 Protein digestion – 1 day

4.11 MALDI-TOF – 5 min per sample

4.12 Edman degradation – 1–2 days

4.13 Capillary chromatography with ion trap/QTOF MS/MS – 2–3 h per sample, data analysis can take several days per sample. *De novo* sequencing by direct infusion – half an hour per sample

5 Trouble Shooting

5.1 Lipopeptide Extraction – Main problem is inefficient extraction so extra care should be taken when extracting and repeat until no color is further extracted
5.2 Lipoproteins and polymeric extraction – Dialysis may be problematic and care needs to be taken to ensure that both ends of tubing are clamped tightly so none of the biosurfactant product leaks out
5.3 TLC – Spray reagent may show the presence of other spots (caused for instance by carbon source or fatty acids) in addition to the biosurfactant. Use a second spray such as ninhydrin to confirm the presence of amino acids
5.4 HPLC-UV/HPLC-MS – Some lipopeptides/peptides may co-elute together. The use of HPLC-MS is much more useful for determining whether co-elution is occurring. This problem can be overcome by modifying the elution conditions or changing to a different column chemistry
5.5 SDS-PAGE – A large number of problems maybe encountered when using this technique and for guidance refer to published information. One major problem is that protein containing biosurfactants that are similar in size will not be separated using 1D gels and therefore using two dimensional (2D) gel maybe necessary.
5.6 Bradford assay – As with all colorimetric assay the presence of some impurities may influence the results
5.7 Fatty acid analysis – Problems occur when either the fatty acid is not released from the biosurfactant or sufficient derivatisation may not have occurred
5.8 Carbohydrate determination by colorimetric and GC analysis major problem are the same as 5.6 and 5.7
5.9 MALDI-ToF analysis – impurities will severely affect the signal so removal of salts and buffers is essential
5.10 Edman degradation – If no sequence information is obtained during the first three or four cycles this may indicate that the N terminus is blocked and may require treatment with an enzyme such as PyroGlu amino peptidase to remove the modification. Sequencing may then commence at residue two of the truncated peptide

6 Research Needs

The major research needs surround the production of high yields and cost effective downstream processing. Downstream processing is probably the most expensive process when studying lipopeptide and high molecular weight biosurfactants. Lipopeptides are extremely powerful biosurfactants and the ability to produce increased yields with low cost purifying methods would make them commercially attractive. With regard to the biopolymers, research into the further characterization of the unknown structures or those that are only partially characterized is necessary.

References

Aguilar M-L (2004) Methods in molecular biology. In HPLC of Peptides and Proteins; Methods and Protocols. Aguilar ML, Ed. Totowa, NJ: Humana Press.

Banat IM, Makkar RS, Cameotra SS (2000) Potential commercial applications of microbial surfactants. Appl Microbiol Biotechnol 53: 495–508.

Bradford MM (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72: 248–254

Chaplin MF (1986) Monosaccharides. In Carbohydrate Analysis: a Practical Approach. MF Chaplin and JF Kennedy (eds.). Oxford: IRL Press, pp. 1–36.

Desai JD, Banat IM (1997) Microbial production of surfactants and their commercial potential. Microbiol Mol Biol Rev 61: 47–64.

Graham RLJ, Graham C, McMullan G (2007) Microbial proteomics: a mass spectrometry primer for biologists Microbe Cell Fact 15: 6–26.

Kakinuma A, Hori M, Isono M, Tamura G, Arima K (1969) Determination of amino acid sequence in surfactin, a crystalline peptidolipid surfactant produce by *Bacillus subtilis.* Agric Biol Chem 33: 971–997.

Kosaric N (1993) Surfactant science, vol. 48. New York: Marcel Dekker, p. 483.

Kuiper I, Lagendijk EL, Pickford R, Derrick JP, Lamers GEM, Thomas-Oates JE, Lugtenberg BJJ, Bloemberg GV (2004) Characterisation of two *Pseudomonas putida* lipopeptides biosurfactants, putisolvin I and II, which inhibit biofilm formation and break down existing biofilms. Mol Microbiol 54: 97–113.

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680–685.

Lang S (2002) Biological amphiphiles (microbial biosurfactants). Curr Opin In Colloid Interface Sci 7: 12–20.

Manzi AE, Varki A (1993) Composition analysis of glycoproteins. In Glycobiology: a Practical Approach. M Fukuda and A Kobata (eds.). Oxford: Univeristy Press, pp. 27–77.

Navon-Venezia S, Zosim Z, Gottlieb A, Legmann R, Carmeli S, Ron EZ, Rosenberg E (1995) Alasan, a new bioemulsifier from *Acinetobacter radioresistens.* Appl Environ Microbiol 61: 3240–3244.

Raaijmakers JM, De Bruijn I, De Kock MJD (2006) Cyclic lipopeptide production by plant-associated *Pseudomonas* sp: Diversity, activity, biosynthesis and regulation. Mol Plant Microbe Interact 19: 699–710.

Ron EZ, Rosenberg E (2001) Natural roles of biosurfactants. Environ Microbiol 3: 229–236.

Rosenberg E, Zuckerberg A, Rubinovitz C, Gutnick DL (1979) Emulsifier of *Arthrobacter* RAG-1:isolation and emulsifying properties. Appl Environ Microbiol 37: 402–408.

Sorenson D, Nielsen TH, Christophersen C, Sorensen J, Gaijhede M (2001) Cyclic lipoundecapeptide amphisin from *Pseudomonas* sp DSS73. Acta Crystallogr 57: 1123–1124.

Strader MB, Tabb DL, Hervey WJ, Pan C, Hurst GB (2006) Efficient and specific trypsin digestion of microgram to nanogram quantities of proteins in organic-aqueous solvent systems. Anal Chem 78: 125–134.

Symmank H, Franke P, Saenger W, Berhard F (2002) Modification of biologically active peptides: Production of a novel lipohexapeptide after engineering of *Bacillus subtilis* surfactin synthetase. Protein Eng 15: 913–921.

Toren A, Navon-Venezia S, Ron EZ, Rosenberg E (2001) Emulsifying activities of purified alasan proteins from *Acinetobacter radioresisten* HA53. Appl Environ Microbiol 67: 1102–1106.

Vater J, Kablitz B, Wilde C, Franke P, Mehta N, Cameotra, SS (2002) Matrix-assisted laser desorption ionization-time of flight mass spectrometry of lipopeptide biosurfactants in whole cells and culture filtrates of *Bacillus subtilis* C-1 isolated from petroleum sludge. Appl Environ Microbiol 68: 6210–6219.

Yakimov MM, Timmis KN, Wray V, Fredrickson HL (1995) Characterization of a new lipopeptide surfactant produced by thermotolerant and halotolerant subsurface *Bacillus licheniformis* BAS50. Appl Environ Microbiol 61: 1706–1713.

Zachara NE, Gooley AA (2000) Identification of glycosylation sites in mucin peptides by Edman degradation Methods Mol Biol 125: 121–128.