Chapter 9

Screening for Expressed Nonribosomal Peptide Synthetases and Polyketide Synthases Using LC-MS/MS-Based Proteomics

Yunqiu Chen, Ryan A. McClure, and Neil L. Kelleher

Abstract

Liquid chromatography–mass spectrometry (LC-MS)-based proteomics is a powerful technique for the profiling of protein expression in cells in a high-throughput fashion. Herein we report a protocol using LC-MS/MS-based proteomics for the screening of enzymes involved in natural product biosynthesis, such as nonribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs) from bacterial strains. Taking advantage of the large size of modular NRPSs and PKSs (often >200 kDa), size-based separation (SDS-PAGE) is employed prior to LC-MS/MS analysis. Based upon the protein identifications obtained through software search, we can accurately pinpoint the expressed NRPS and/or PKS gene clusters from a given strain and growth condition. The proteomics screening result can be used to guide the discovery of potentially new nonribosomal peptide and polyketide natural products.

Key words Proteomics, Liquid chromatography, Mass spectrometry, Nonribosomal peptide, Polyketide, Natural product

1 Introduction

Many natural products with antibiotic, anticancer, and antifungal activities are synthesized by nonribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs) [1]. Multimodular NRPS and PKS enzymes are often very large proteins (>200 kDa) and contain many functional domains for the assembly of simple building blocks into natural products [2]. Traditionally, natural product discovery has relied on a bioactivity-guided screening approach, which includes repeated fractionation to isolate the compounds of interest [3]. This approach often suffers from a high frequency of rediscovering known natural products [4]. With the advent of the genomic revolution and whole genome sequencing, researchers have realized that microorganisms possess a far greater genetic potential for natural product biosynthesis than what had been observed [5]. Through rational prediction followed by tar-

Bradley S. Evans (ed.), *Nonribosomal Peptide and Polyketide Biosynthesis: Methods and Protocols*, Methods in Molecular Biology, vol. 1401, DOI 10.1007/978-1-4939-3375-4_9, © Springer Science+Business Media New York 2016

geted detection of the natural products based on the genetic information, an approach known as "genome mining" has led to the discovery of a number of new natural products [6–8]. On the other hand, genome mining approaches also have faced roadblocks in discovery as some biosynthetic pathways are not actively expressed or expressed at extremely low levels in laboratory conditions, also known as "cryptic" pathways [9, 10].

To complement the genome mining and bioassay-based natural product discovery approaches, a proteomics approach was developed by the Kelleher group, termed Proteomic Investigation of Secondary Metabolism (PrISM) [11]. By initially identifying the actively expressed biosynthetic proteins, PrISM avoids pursuing "cryptic" pathways. Natural products are predicted according to the genetic information and discovered in a targeted way. The PrISM approach has led to the discovery of several new natural products including koranimine, flavopeptins, gobichelin A and B, as well as the identification of biosynthetic pathways for many known natural products [11–15]. Most recently, the "depth" at which PrISM could be applied (i.e., detecting expression of gene clusters at low levels) was evaluated using a single Actinomycete strain with its genome sequence; this was called Genome-enabled PrISM [16].

In this chapter, we present the detailed protocol for the PrISM screening of Actinobacteria strains for the expressed NRPS and PKS proteins. In this protocol, Actinobacteria strains are grown under a variety of culture conditions and harvested at several time points. Cells are lysed by mechanical disruption using bead beating, and the proteomes are fractionated by size using one dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The high molecular weight proteins (>150 kDa), often representing NRPSs or PKSs, are subjected to in-gel trypsin digestion followed by liquid chromatography coupled with mass spectrometric (LC-MS) analysis. The LC-MS data are automatically analyzed using proteomics software and the protein identifications reveal which NRPS and/or PKS pathways are actively expressed under the growth conditions—guiding the discovery of the corresponding natural products.

2 Materials

2.1 Culture Media

1. *Malt Yeast Glucose*: 10 g/L malt extract, 4 g/L yeast extract, 4 g/L glucose, pH to 7.
2. *ATCC 172*: 10 g/L glucose, 20 g/L soluble starch, 5 g/L yeast extract, 5 g/L N-Z amine type A, 1 g/L $CaCO_3$.
3. *ISP4*: 10 g/L soluble starch, 1 g/L K_2HPO_4, 1 g/L $MgSO_4$, 1 g/L NaCl, 2 g/L $(NH_4)_2SO_4$, 2 g/L $CaCO_3$, 0.001 g/L

$FeSO_4 \bullet 7H_2O$, 0.001 g/L $ZnSO_4 \bullet 7H_2O$, 0.001 g/L $MnCl_2 \bullet 4H_2O$, pH to 7.0.

4. *Arginine Glycerol Salts*: 0.85 g/L l-arginine, 12.5 g/L glycerol, 1 g/L K_2HPO_4, 1 g/L NaCl, 0.5 g/L $MgSO_4 \bullet 7H_2O$, 0.01 g/L $FeSO_4 \bullet 7H_2O$, 0.001 g/L $CuSO_4 \bullet 5H_2O$, 0.001 g/L $ZnSO_4 \bullet 7H_2O$, 0.001 g/L $MnSO_4 \bullet H_2O$, pH to 7.0.
5. *Soy Flour Mannitol*: 20 g/L d-mannitol, 20 g/L soy flour, pH to 7.0 (This medium needs to be autoclaved twice).
6. *4×R2A*: 2 g/L peptone, 2 g/L starch, 2 g/L glucose, 2 g/L yeast extract, 2 g/L casein hydrolysate, 1.2 g/L K_2HPO4, 1.2 g/L sodium pyruvate, 0.1 g/L $MgSO_4$.
7. *MSB medium*: To make 1 L MSB medium, mix separately sterilized A (40 mL), B (10 mL), C (5 mL), and ddH_2O. Supplement with 10 mM (final concentration) of sodium succinate and 0.05 % casamino acids.
   (a) 1 M $Na_2HPO_4 + KH_2PO_4$, pH 7.3. Mix ~38.25 mL 1 M Na_2HPO_4 stock solution and ~11.5 mL KH_2PO_4 stock solutions and adjust pH to 7.3.
   (b) Dissolve 20 g nitrilotriacetic acid and 14 g KOH in 700 mL H_2O. Then add the following chemicals individually in order: 28 g $MgSO_4$, 6.67 g $CaCl_2 \bullet 2H_2O$ (then add KOH to raise pH till dissolved), 0.0185 g $(NH_4)_6Mo_7O_{24} \bullet 4H_2O$, 0.198 g $FeSO_4 \bullet 7H_2O$, 100 mL Hunter's metals 44*. Adjust pH to 6.8 with KOH. Adjust volume to 1 L.

       *Hunter's metals 44: 2.5 g EDTA (free acid) dissolved in 800 mL dH_2O, add 10.95 g $ZnSO_4 \bullet 7H_2O$, 5 g $FeSO_4 \bullet 7H_2O$, 1.54 g $MnSO_4 \bullet H_2O$, 0.392 g $CuSO_4 \bullet 5H_2O$, 0.250 g $Co(NO_3)_2 \bullet 6H_2O$, 0.177 g $Na_2B_4O_7 \bullet 10H_2O$, add five drops of concentrated sulfuric acid, adjust volume to 1 L and store at 4 °C.
   (c) 20 % (wt/vol) $(NH_4)_2SO_4$.

2.2 Reagents

All solutions are prepared using ultrapure water (prepared by purifying deionized water to attain a resistivity of 18 MΩ • cm at 25 °C) and analytical grade reagents.

1. SDS lysis buffer (4×): 125 mM Tris (pH 6.8), 4 % (wt/vol) SDS, 40 % (vol/vol) glycerol, 10 % (vol/vol) β-mercaptoethanol, 0.01 % (wt/vol) bromophenol blue.
2. SDS-PAGE running buffer: 25 mM Tris, pH 8.3, 192 mM glycine, 0.1 % SDS.
3. Ammonium bicarbonate (NH_4HCO_3) solution: 100 mM in water, pH 7.8 and 50 mM in water, pH 7.8.
4. Tris-(2-carboxyethyl) phosphine (TCEP): 10 mM in 50 mM NH_4HCO_3, pH 7.
5. Iodoacetamide (IAA): 50 mM solution in water.

6. Trypsin: dissolve one vial of lyophilized trypsin in 40 μL trypsin resuspension buffer provided by manufacturer, and heat at 30 °C for 15 min, then dilute 40-fold using 50 mM NH_4HCO_3 to a final concentration of 12.5 μg/mL. This solution should be freshly prepared.
7. 1:2 5 % formic acid–acetonitrile: 1 part 5 % v/v formic acid solution in water to 2 parts LC-MS grade acetonitrile.
8. LC-MS buffer A: LC-MS grade water with 0.1 % formic acid.
9. LC-MS buffer B: LC-MS grade acetonitrile with 0.1 % formic acid.

2.3 Equipment

1. LTQ-Orbitrap mass spectrometer (Thermo-Fisher Scientific, MA, USA) with a nanoelectrospray source.
2. Nano LC system with autosampler.
3. Nano LC trap column (2 cm length × 100 μm ID, 5 μm C18 particle).
4. Nano LC analytical column (10 cm length × 75 μm ID, 5 μm C18 particle).
5. Vortexer.
6. Vortexer adapter for 1.5 mL tubes, horizontal orientation (MOBIO, CA, USA).
7. Sonic dismembrator with 1/8 in. microtip.
8. Electrophoresis system and power supply.
9. SpeedVac concentrator.
10. Ultrasonic cleanser.
11. Shaker incubator.
12. 18 × 150 mm glass tubes with caps (autoclave).
13. Centrifuge for 1.5 mL tube.
14. Protein LoBind Tubes (Eppendorf, Germany).
15. 15 mL conical centrifuge tubes.
16. 14 mm culture dish.
17. Razor blades.
18. Carbide beads (0.25 mm, MOBIO, CA, USA).

3 Methods

3.1 Bacteria Strain Growth

This protocol uses Actinobacteria strains as an example, but the PrISM methodology can be applied to other species of bacteria as well. We suggest to use at least three types of culture media and to collect samples at 4 time points (*see* **Note 1**).

1. Inoculate 5 mL of culture media in an 18 × 150 mm sterile glass tube with a few single colonies of the strain of interest to generate a starting culture.
2. The tubes are placed at 45° angle in an incubator shaker and are allowed to grow at 30 °C with rotation at 250 rpm.
3. After ~2–3 days, when the strain has reached log growth phase, transfer an aliquot (500 μL) of the starting culture to the screening media (5 mL each) which are also contained in 18 × 150 mm glass tubes.
4. For each screening medium, four tubes are prepared to be harvested at 4 time points (*see* **Note 2**). For a typical Actinobacteria strain, we recommend harvesting at 24 h, 48 h, 72 h, and 96 h after transferring to screening media.

3.2 Cell Lysis

1. At each time point for harvesting, the entire culture is transferred to a 15 mL conical centrifuge tube and subjected to centrifugation at >4000 × *g* for 10 min (*see* **Note 3**). The culture supernatants are removed and the cell pellets can be stored at −80 °C until ready for lysis.
2. Resuspend the cell pellet in an equal volume of 4× SDS lysis buffer, and transfer it to a 1.5 mL centrifuge tube. Place the sample on a hotplate set at 95 °C for 10 min (*see* **Note 4**).
3. Add a small amount (the same volume as the cell pellet) of carbide beads (0.25 mm) to the sample, and place the tubes on a vortexer with a horizontal orientation. Vortex for 30 min at maximum speed (*see* **Note 5**).
4. Sonicate the sample for 1 min using a sonic dismembrator with a 1/8 in. microtip (*see* **Note 6**).
5. Heat the samples at 95 °C for 10 min.
6. Centrifuge the samples at 20,000 × *g* for 10 min. Extract the supernatant to a new 1.5 mL centrifuge tube for SDS-PAGE analysis. This is the proteomic sample for a given strain under a certain growth condition. The samples can be stored at −80 °C until ready for SDS-PAGE separation.

3.3 Separation of the Proteomic Samples Using SDS-PAGE

1. Assemble the 10 % T precast gels (15 well) onto the electrophoresis system following the manufacturer's instructions. Fill the inner and outside chambers of the cassette with the SDS running buffer.
2. Load 5 μL of the prestained protein standard into the first lane. Load 15 μL of the proteomic samples from Subheading 3.3 into each other lane. To avoid cross-contamination, try to load samples from the same strain on the same gel. Perform the electrophoresis at a constant voltage of 180 V, until the bromophenol blue dye reaches the bottom of the gel.

3. After SDS-PAGE, pry the gel plates open with a spatula. Place the gel in a clean 14 mm culture dish. Wash the gel with 200 mL of water three times for 5 min each. Remove the water and add ~50 mL of Bio-Safe™ Coomassie stain to the dish and let shake for 1 h. After staining, remove the stain and rinse with water for 3 × 10 min. Take a picture of the Coomassie stained gel and check the efficiency of the cell lysis (*see* **Note 7**) (Fig. 1).

3.4 In-Gel Tryptic Digestion

1. Remove most of the water from the dish—leaving a small amount to prevent the gel from drying. For each lane, a razor blade is used to excise the region above 250 kDa as one slice and the 150–250 kDa region as another slice (*see* **Note 8**). Cut each gel slice into 1 × 1 mm gel pieces and transfer using the razor blade into a 1.5 ml LoBind Eppendorf tube (*see* **Note 9**). Always use a new razor blade and a new tube for each lane from a gel sample.
2. Add 300 μL H_2O to each tube and vortex for 15 min.
3. Add 300 μL acetonitrile and wash for another 15 min.
4. Remove the supernatant carefully without picking up the gel pieces. Wash the gel pieces with 300 μL of 100 mM NH_4HCO_3 for 15 min. Discard the supernatant. Wash the gel pieces with 300 μL of 100 mM NH_4HCO_3–acetonitrile (50:50 vol/vol) for 15 min (*see* **Note 10**).

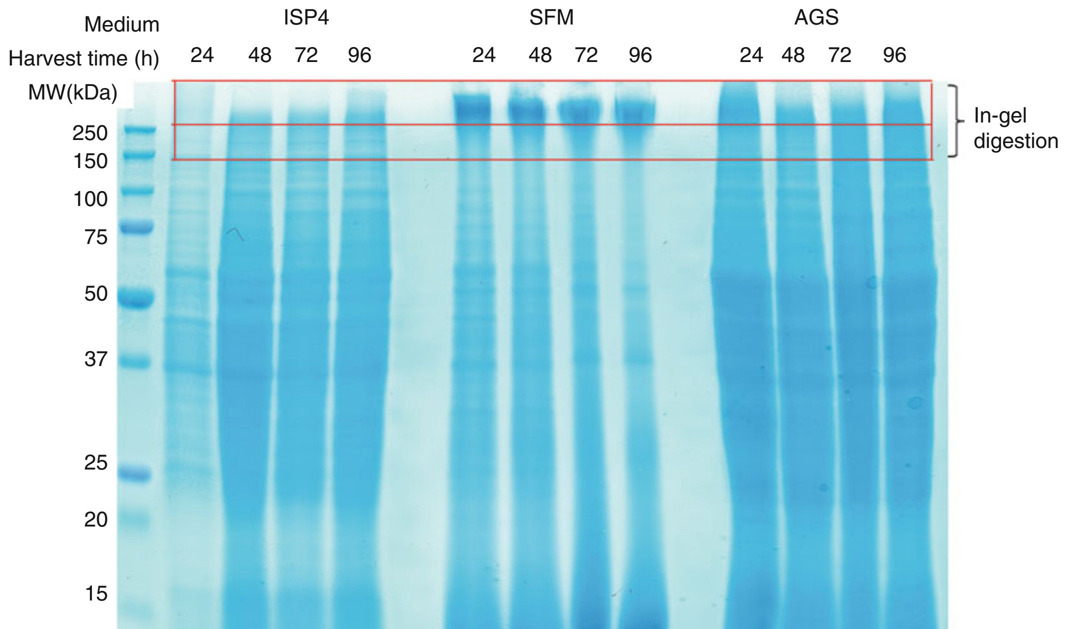

Fig. 1 PrISM screening for an Actinobacteria strain grown in three different media (ISP4, SFM, and AGS) and harvested at 4 time points (24, 48, 72, and 96 h). The proteome samples are separated on an SDS-PAGE gel and stained by Coomassie. The high molecular weight regions (>150 kDa, shown within *red* box) are subjected to in-gel trypsin digestion

5. Add 100 μL acetonitrile to dehydrate the gel pieces for 5 min.
6. Discard the supernatant. Dry the gel pieces in a SpeedVac for 5 min. The samples can be stored at −80 °C until ready to move onto the next step.
7. Reduce the disulfide bonds of proteins by adding 50 μL of 10 mM TCEP in 50 mM NH_4HCO_3, incubate at room temperature for 30 min. Discard the supernatant.
8. Add 50 μL of 50 mM iodoacetamide in 50 mM NH_4HCO_3 (freshly prepared). Incubate at room temperature in the dark for 30 min. Discard the supernatant.
9. Wash the gel pieces with 300 μL of 100 mM NH_4HCO_3 for 15 min. Discard the supernatant.
10. Wash the gel pieces with 300 μL of 50 mM NH_4HCO_3–acetonitrile (50:50 vol/vol) for 15 min. Discard the supernatant.
11. Add 100 μL of acetonitrile to dehydrate the gel pieces for 5 min. Discard the supernatant.
12. Dry the gel pieces in a SpeedVac for 5 min. The samples can be stored at −80 °C until ready to move onto the next step.
13. Add 20 μL of the trypsin solution (12.5 μg/mL) to each tube. Allow the gel bands to rehydrate in trypsin solution for 30 min on ice.
14. If the digestion solution is fully absorbed after 30 min, add an additional 10 μL trypsin solution to cover the gel pieces. Let the gel sit on ice for an additional 1 h.
15. Remove the excessive trypsin solution and add 50 mM NH_4HCO_3 (sufficient volume to cover the gel pieces). Keep a record of the quantity of liquid added to the gel pieces.
16. Incubate at 30 °C overnight.
17. Cool the sample to room temperature and add 2 μL of 100 % formic acid to quench the digestion.
18. Add acetonitrile (an equal volume as added from 15) to the tube and shake vigorously on a vortexer for 30 min.
19. Transfer the supernatant to a new clean Eppendorf LoBind tube. This contains the extracted peptides.
20. Add 100 μL of 1:2 5 % formic acid–acetonitrile to the gel pieces and let shake for 20 min. Transfer the supernatant to the tube in **step (19)**.
21. Repeat **step (20)**.
22. Add 100 μL of acetonitrile to the gel pieces. Shake for 10 min. Transfer the supernatant to the tube in **step (19)**.
23. Repeat **step(22)**.
24. Dry the peptides in the tube at **step (19)** completely using the SpeedVac. Store the sample in −80 °C freezer until ready for LC-MS/MS analysis.

3.5 LC-MS/MS

1. Add 20 μL of LC-MS buffer A to each tube. To solubilize the peptides, place the tubes in an ultrasonic cleanser containing ice water and sonicate for 10 min.
2. Centrifuge the samples at 20,000 × *g* for 10 min to precipitate any insoluble materials. Transfer the supernatant to an LC sample vial.
3. The nanoLC method is configured as follows: the samples are initially loaded using 100 % LC-MS buffer A to the nanoLC trap column with a flow rate of 3 μL/min for 10 min. The eluent is diverted to the waste at this step.
4. Switch the valve so that the eluent is diverted to the nanoLC analytical column. The LC gradient is set as follows, with a flow rate at 300 nL/min:

Time (min)	% B
0	0
55	45
63	80
67	0
90	0

5. The eluted peptides are subjected to positive nanoelectrospray ionization (nanoESI) for mass spectrometric analysis using a Thermo LTQ-Orbitrap instrument (*see* **Note 11**). The instrument settings are configured as follows:

   FTMS full scan Automatic Gaining Control (AGC): 1e6.

   FTMS full scan maximum ion time: 2000 ms.

   Ion trap MS^n AGC: 1e4.

   Ion trap MS^n maximum ion time: 300 ms.

   Mass spectrometry method setup:

   FTMSfull scan from *m/z* 400–2000 using a 30,000 resolving power. For each full scan, the top five most intense ions are selected for data dependent fragmentation using collision induced dissociation (CID) at 35 eV. The fragments are detected in the ion trap. Minimum signal required for MS^2 is 500. Dynamic exclusion is set to 60 s, repeat count = 2. 1+ and unassigned charge states are rejected for MS^2 events.

3.6 Data Analysis

1. To process the LC-MS data, extract the peak file lists from the .raw data files by using the DTA generator module from COMPASS software suite (http://www.chem.wisc.edu/~coon/software.php). Assume the precursor charge state to be 2+ through 4+.
2. Search the data using a proteomics software. In this protocol we use the Mascot software (Matrix Science, UK) for demon-

stration but other software can be used. If the target strain has a sequenced genome or draft genome, the target protein database should be used for the search. Otherwise, the publicly available NCBI nonredundant protein database (NCBInr) can be used, selecting bacteria as the taxonomy.

3. The following searching parameters are recommended: 10 ppm precursor mass error tolerance; 0.8 Da MS^2 error tolerance; carbamidomethylation of cysteines as fixed modification; oxidation of methionines as variable modifications; maximal number of missed cleavages is 2; allowed peptide charge states are 2+, 3+, and 4+; data format is Mascot generic; and instrument type is ESI-TRAP.
4. After the Mascot search has finished, open the search results and perform filtering. We recommend the "ion score or expect cut-off" to be set to 30.
5. Browse the "protein family summary" and look for proteins that are associated with NRPS or PKS biosynthesis. For strains with sequenced genomes, the protein identification list will inform which high molecular weight NRPS and/or PKS proteins are expressed. The number of peptides identified from each proteomic sample is generally correlated with the expression level of the NRPS/PKS proteins under each growth conditions (Table 1).

Table 1
An example of protein identification list (partial) from Mascot search for an Actinobacteria strain using its own sequenced genome as the database

Protein description	Protein mass (kDa)	#Matches (sig)	#Sequence (sig)	%Coverage
NAD-glutamate dehydrogenase	185	262	68	48.60
GAF sensor hybrid histidine kinase	194	35	15	10.33
Protein of unknown function DUF3686	175	30	16	12.54
Hypothetical protein Sfla_5835	185	27	14	9.96
Hypothetical protein Sfla_5465	169	18	15	11.81
DNA-directed RNA polymerase, beta subunit	144	18	12	11.39
Amino acid adenylation domain protein	265	15	8	3.80
Bacterioferritin	18	8	6	39.62
Amino acid adenylation domain protein	340	8	5	1.93
Amino acid adenylation domain protein	511	8	5	1.60

Three "amino acid adenylation proteins" (i.e., NRPSs, bolded), 511 kDa, 340 kDa, and 265 kDa each, are identified, indicating these three NRPS proteins are expressed under this growth condition

6. If the genome sequence of the target strain is not available and the Mascot search is performed using NCBInr as the database, the protein identification list usually contains NRPS/PKS proteins from different organisms (Table 2). Depending on the sequence homology between the actual protein sequence in your sample and the sequence in the database, you may get one or two protein identifications that contain many peptide hits, or you may get many protein identifications each containing only one or two peptide hits. In the latter case and when the LC-MS/MS data are of high quality (as shown in Table 2), it is likely that the NRPS/PKS in your sample is a novel gene cluster that is not represented in the database.

4 Notes

1. For a general PrISM screening of Actinobacteria strains, the typical medium for starting culture is Malt Yeast Glucose (MYG) medium or ATCC 172 medium, in which most strains grow vigorously. For the screening media, the researchers can use whichever media they prefer. In our practice, we have used ISP4, Arginine Glycerol Salts (AGS), Soy Flour Mannitol (SFM), 4× R2A, and MSB media. All media are prepared using tap deionized water (do not use ultrapure water) and autoclaved.
2. This growth condition (30 °C, 3 day, 250 rpm) is suitable for most Actinobacteria strains. However, if the strain of interest has special growth needs (e.g., higher/lower temperature, special growth factors, light/dark environment, or faster/slower growth rate), the specific growth conditions should be adjusted accordingly.
3. For some strains, the cell pellet can be very loose. Be careful when pouring out the culture supernatant. Use a pipette when necessary. If the cells do not pellet down completely, repeat the centrifugation step at a higher speed then remove the medium completely.
4. Estimate the volume of the cell pellet. Then add an equal volume of the lysis buffer, e.g., 25–100 μL.
5. Bead beating has been demonstrated as the most powerful and universal method for lysis of cells and tissue. We have tested that carbide bead with 0.25 mm diameter is best for the lysis of bacteria samples. Placing the tubes in the horizontal orientation increases lysis efficiency. If the cells are not lysed using bead beating, other methods should be considered, sonication, freeze–thaw, French press, manual grinding, etc.

Table 2
An example of Mascot search result (partial) using NCBInr as the database for an Actinobacteria strain which has no sequenced genome available

Peptide sequence	Protein description	Accession #
DAEALVAYcDR	Nonribosomal peptide synthetase [Streptomyces kitasatoensis	BAH68474
FVADPFGEPGER	Nonribosomal peptide synthetase [Streptomyces griseolus]	BAH68409
ADGAVEYIGR	Pyoverdine sidechain peptide synthetase I, epsilon-Lys module [Pseudomonas syringae pv. tabaci ATCC 11528]	ZP_05637427
GVGPESVVGVAVPR	Nonribosomal peptide synthetase [Rhodococcus opacus B4]	YP_002779301
ADGNVDFLGR	Amino acid adenylation domain-containing protein [Opitutus terrae PB90-1]	YP_001818855
FVADPYGAPGSR	Nonribosomal peptide synthetase [Streptomyces gibsonii]	BAH68663
LGAGDDIPIGTPVAGR	Peptide synthetase 2 [Streptomyces roseosporus NRRL 11379]	AAX31558
DVVFGTTVSGR	Amino acid adenylation domain protein [Acetivibrio cellulolyticus CD2]	ZP_07326042
FTADPYGPAGSR	Putative NRPS [Streptomyces griseus subsp. griseus NBRC 13350]	YP_001824777
GAGPETLVAVALPR	Mannopeptimycin peptide synthetase MppB [Streptomyces hygroscopicus]	AAU34203
TVAALAALAR	Putative nonribosomal peptide synthetase [Nocardia farcinica IFM 10152]	YP_119006
LGAGDDIPIGSPVAGR	Amino acid adenylation domain-containing protein [Frankia sp. EAN1pec]	YP_001510190
VVAIALPR	Peptide synthetase ScpsA [Saccharothrix mutabilis subsp. capreolus]	AAM47272
VAGPALTDFLADQR	Nonribosomal peptide synthetase [Streptomyces griseolus]	BAH68411
LVLITDEAR	Nonribosomal peptide synthetase [Streptomyces griseolus]	BAH68411
DIATAYAAR	Amino acid adenylation domain-containing protein [Streptomyces roseosporus NRRL 15998]	ZP_04696845
IPLSYAQR	Nonribosomal peptide synthetase [Streptomyces fungicidicus]	ABD65957

All the peptides related to NRPS/PKS are extracted. These peptides match to NRPS proteins from different organisms, indicating the detected NRPS share a low sequence homology with any known NRPS gene clusters in the database and likely represent a new NRPS biosynthetic pathway

6. This short sonication step is to disrupt DNA in the cells. The undisrupted DNA molecules will cause streaking bands on the SDS-PAGE.
7. The Coomassie staining step is primarily to check the efficiency of cell lysis. A successfully lysed sample should show an abundant amount of proteins on the SDS-PAGE gel. If the cells are not lysed properly, little or no protein bands will show on the gel. In this case, there is no need to proceed to the following steps. Other cell lysis protocols should be attempted to extract the proteome. If the proteome sample is much diluted (>100 μL volume), and higher sensitivity of detection is desired, the sample can be separated on a preparative SDS-PAGE gel, which contains a single lane with 250 μL loading capacity.
8. We recommend performing in-gel trypsin digestion to all samples, whether they show visible protein bands at the high molecular weight region by Coomassie stain. Coomassie stain has a much lower sensitivity of detection than nanoLC-MS. We also suggest cutting the high molecular weight region into two separate gel slices, one contains proteins >250 kDa and one contains 150–250 kDa proteins. Additional size based fractionation will result in more protein identifications.
9. In-gel digestion procedure is very sensitive to the contamination of keratin. To minimize contamination, always wear gloves when handling gels and performing in-gel digestion. Always use LoBind tubes and use clean spatula when preparing reagents.
10. If the band pieces are still blue after this step, repeat the washing steps using 300 μL of 100 mM NH_4HCO_3 for 15 min and then 300 μL of 100 mM NH_4HCO_3–acetonitrile (50:50 vol/vol) for 15 min.
11. Other types of mass spectrometers that are capable of proteomic analysis can also be used, Q-Exactive, QqTOF, Ion trap, FT-ICR, etc. Users need to adjust the mass spectrometric method according to the type of instrument used.

Acknowledgement

This work was supported by National Institute of General Medical Sciences of the National Institutes of Health under award number R01 GM067725. We thank the Agricultural Research Service, US Department of Agriculture for providing the bacterial strains.

References

1. Walsh CT (2004) Polyketide and nonribosomal peptide antibiotics: modularity and versatility. Science 303:1805–1810
2. Fischbach MA, Walsh CT (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. Chem Rev 106:3468–3496
3. Koehn FE (2008) High impact technologies for natural products screening. Prog Drug Res 65:177–210
4. Baltz RH (2006) Marcel Faber Roundtable: is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? J Ind Microbiol Biotechnol 33:507–513
5. Nett M, Ikeda H, Moore BS (2009) Genomic basis for natural product biosynthetic diversity in the actinomycetes. Nat Prod Rep 26: 1362–1384
6. Lautru S, Deeth RJ, Bailey LM et al (2005) Discovery of a new peptide natural product by Streptomyces coelicolor genome mining. Nat Chem Biol 1:265–269
7. Corre C, Challis GL (2009) New natural product biosynthetic chemistry discovered by genome mining. Nat Prod Rep 26:977–986
8. Gross H, Stockwell VO, Henkels MD et al (2007) The genomisotopic approach: a systematic method to isolate products of orphan biosynthetic gene clusters. Chem Biol 14:53–63
9. Scherlach K, Hertweck C (2009) Triggering cryptic natural product biosynthesis in microorganisms. Org Biomol Chem 7:1753–1760
10. van Wezel GP, McDowall KJ (2011) The regulation of the secondary metabolism of *Streptomyces*: new links and experimental advances. Nat Prod Rep 28:1311–1333
11. Bumpus SB, Evans BS, Thomas PM et al (2009) A proteomics approach to discovering natural products and their biosynthetic pathways. Nat Biotechnol 27:951–956
12. Evans BS, Ntai I, Chen Y et al (2011) Proteomics-based discovery of koranimine, a cyclic imine natural product. J Am Chem Soc 133:7316–7319
13. Chen Y, McClure RA, Zheng Y et al (2013) Proteomics guided discovery of flavopeptins: anti-proliferative aldehydes synthesized by a reductase domain containing non-ribosomal peptide synthetase. J Am Chem Soc 135: 10449–10456
14. Chen Y, Ntai I, Ju KS et al (2012) A proteomic survey of nonribosomal peptide and polyketide biosynthesis in Actinobacteria. J Proteome Res 11:85–94
15. Chen Y, Unger M, Ntai I et al (2013) Gobichelin A and B: mixed-ligand siderophores discovered using proteomics. Medchemcomm 4:233–238
16. Albright JC, Goering AW, Doroghazi JR et al (2014) Strain-specific proteogenomics accelerates the discovery of natural products via their biosynthetic pathways. J Ind Microbiol Biotechnol 41:451–459