Chapter 2

Analysis of Bacterial Surface Interactions with Mass Spectrometry-Based Proteomics

Christofer Karlsson, Johan Teleman, and Johan Malmström

Abstract

Host–pathogen protein–protein interaction networks are highly complex and dynamic. In this experimental protocol we describe a method to isolate host proteins attached to the bacterial surface followed by quantitative mass spectrometry based proteomics analysis. This technique provides an overview of the host–pathogen interaction network, which can be used to guide directed perturbations of the system, and to select target of specific interest for further studies.

Key words Bacteria, Surface absorption, Mass spectrometry, Proteomics, Trypsin digestion, Peptide solid phase extraction, Bioinformatics

1 Introduction

Microbial pathogenesis is the result of complex molecular interactions between the host and a microbial pathogen. Nonspecific and specific pathogen recognition results in the coating of the pathogen surface by immune system proteins derived from several different biochemical processes such as complement deposition and antibody binding. These processes aid the pathogen killing and clearance. However, pathogens have evolved mechanisms to interfere with the host immune reactions by for example expressing surface proteins that specifically bind host proteins, to facilitate immune evasion and bacterial dissemination.

A specific example of a pathogen that can bind many different host proteins to the bacterial surface is *Streptococcus pyogenes*. The major virulence factor on the *S. pyogenes* surface is the cell wall anchored M-protein that can bind several human host proteins [1–4]. The M-protein, together with other streptococcal host binding surface proteins, forms a complex host–pathogen protein interaction network on the bacterial surface [5–11]. Investigating binary interactions between host and pathogen proteins is not sufficient to describe the topology of the protein interaction network.

Pontus Nordenfelt and Mattias Collin (eds.), *Bacterial Pathogenesis: Methods and Protocols*, Methods in Molecular Biology, vol. 1535, DOI 10.1007/978-1-4939-6673-8_2, © Springer Science+Business Media New York 2017

Steric hinders, degree of affinity, secondary binding, competitive interactions, and protein abundances are factors that affect which proteins adhere to the bacterial surface. The comprehensive measurement of these interactions requires analytical techniques capable of identifying and quantifying the majority of the proteins involved in the network.

In this protocol we provide a method for quantitative MS analysis of both surface bound host proteins and the complete bacterial protein content in one experimental setup. The protocol includes the use of whole bacteria as affinity probes to isolate host proteins that attach to the bacterial surface (Fig. 1). Whole bacteria and the proteins adhered to the bacterial surface are isolated using centrifugation followed by quantitative mass spectrometry analysis. The rapid development of mass spectrometry (MS) based proteomics has made MS an important technology within life science [12–14]. The prevailing bottom-up MS based techniques analyze digested proteins (peptides), separated based on hydrophobicity using online liquid chromatography, which are then eluted via electrospray to form gas-phase ions. The chromatographic separation reduces the sample complexity, but numerous peptide ions still enter the MS instrument simultaneously. These peptide ions are first mass analyzed (MS1), after which the most abundant peptide ions are selected for collision-induced dissociation (CID) followed by a second mass analysis (MS2) of the derived fragment ions. Subsequent data analysis strategies attempts to match all acquired MS2 spectra computationally to one of all theoretically derived peptide MS spectra from the organisms analyzed [15–20]. From the identified peptides, proteins are inferred using statistical

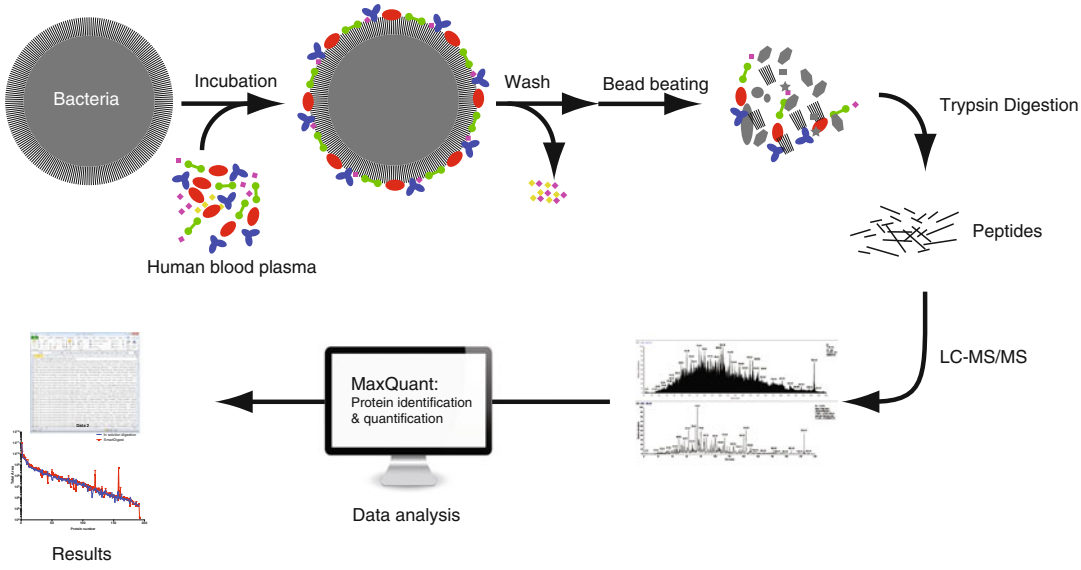

Fig. 1 Outline of the method to identify and quantify bacterial surface interacting host proteins

methods [21, 22]. The intensities of the individual MS1 features are integrated and the area under this curve is used to infer peptide and protein abundance using one of several published software programs [23–27]. In this protocol we use the MaxQuant software [26] as an example, which can be freely downloaded and installed on a standard Microsoft Windows computer.

The protocol outlines how bacterial cellular and surface proteins together with surface attached host proteins can be identified and quantified using MS and label-free quantification. The summed bacterial protein quantity can be utilized to normalize results for uneven sample loss during sample preparation, to remove confounding factors while comparing differential individual protein abundances between different strains or biological conditions. In addition, the quantification of the attached host proteins allows characterization of the host–pathogen protein interaction network topology.

2 Materials

2.1 Bacterial Plasma Adsorption and Sample Homogenization

1. Wash buffer (WB): 150 mM NaCl, 20 mM Tris–HCl pH 7.6.
2. Pooled Normal Human Blood Plasma (Innovative Research) (*see* **Note 1**).
3. LC-grade water.
4. 90 mg silica beads 0.1 μm ∅ (Biospec) in 0.5 ml tubes with an O-ring screw cap.
5. Beadbeater (Fastprep 96, MpBio).

2.2 Trypsin Digestion

1. Urea buffer (UB): 8 M Urea, 0.1 M NH_4HCO_3 in LC-grade water (*see* **Note 2**).
2. Sequence grade trypsin (Promega).
3. 100 mM NH_4HCO_3 (ABC) in LC-grade water (*see* **Note 2**).
4. 500 mM tris(2-carboxyethyl)phosphine) (TCEP).
5. 500 mM 2-iodoacetamide in LC-grade water (*see* **Note 2**).

2.3 Peptide C18 Solid Phase Extraction

1. 10% formic acid (FA) (*see* **Note 3**).
2. UltraMicro Spin Silica C18 300 Å columns (Harvard Apparatus).
3. LC-grade methanol.
4. LC-grade acetonitrile (ACN).
5. LC-grade water.
6. Buffer A: 2% ACN, 0.2% FA in LC-grade water.
7. Buffer B: 50% ACN, 0.2% FA in LC-grade water.
8. Vacuum concentrator.
9. Ultrasonic water bath.

2.4 Shotgun Mass Spectrometry

1. High-resolution, accurate-mass (HR/AM) mass spectrometer with nano-flow UHPLC.

2.5 Data Analysis

1. MaxQuant, http://www.coxdocs.org/doku.php?id = maxquant:start.
2. Protein database in FASTA format, describing the expected protein contents of the samples. This typically includes the proteome of both the bacterium and host/plasma (*see* **Note 4**).

3 Methods

3.1 Bacterial Plasma Adsorption

1. Grow the bacteria to desired growth phase for the interaction analysis.
2. Harvest the bacteria by centrifugation and wash by resuspending the pellet in the wash buffer.
3. Wash the bacteria for a second time and dissolve the pellet in WB to a concentration of 1% w/v.
4. Mix 150 µl of 1% bacteria solution with 450 µl plasma, vortex bacteria briefly just before adding to plasma.
5. Incubate for 30 min at 37 °C on thermal block with 500 rpm shaking.
6. Wash three times with 1 ml WB (5000 × *g*, 5 min, swing-out, soft).
7. Transfer 100 µl of the solution to a 0.5 ml tube with an O-ring screw cap containing 90 mg silica beads.
8. Centrifuge for 5 min 5000 × *g*, remove the supernatant and add 100 µl LC-grade water.
9. Lyse the bacteria with a bead beater for 2 × 3 min at 1600 oscillations/min and 1.5-inch stroke speed.
10. Dry samples completely using a vacuum concentrator.

3.2 Sample Homogenization and Trypsin Digestion

1. Add 50 µl UB to the dried sample.
2. Incubate for 30 min on shaker.
3. Add 1 µl TCEP and incubate at 37 °C for 60 min.
4. Add 2 µl IAA and incubate for 30 min at room temperature in a dark environment.
5. Add 500 µl ABC to the sample.
6. Add 2 µg trypsin to the sample and Incubate for >6 h at 37 °C.
7. Add 100 µl 10% FA to stop the digestion.
8. Ensure that the pH is ≥3.

3.3 Peptide C18 Solid Phase Extraction

1. Place the C18 column in 2 ml collection tube. Add 300 µl methanol for column wash and centrifuge at 200 × *g* for 1 min. Discard the flow-through liquid.

2. Add 300 µl Buffer A to the column and centrifuge at 200 × *g* for 1 min, repeat three times. Discard the flow-through liquid after the second and third centrifugation.
3. Dry the column tip on a lint-free paper towel and place the column in a new collection tube. Add 450 µl digested sample to the column and centrifuge at 200 × *g* for 1.5 min. Reapply the flow-through liquid to the column and centrifuge as above. Repeat twice (totally three centrifugations). Discard the final flow-through liquid.
4. Add 300 µl Buffer A to the column and centrifuge at 200 × *g* for 1.5 min. Repeat three times. Discard the flow-through liquid after the second centrifugation.
5. Dry the column tip on a lint-free paper towel and place the column in a new collection tube.
6. Add 100 µl Buffer B to the column and centrifuge at 200 × *g* for 1 min. Do not discard the flow-through. Repeat three times. Then briefly centrifuge at 1000 × *g*. The final elution volume is 300 µl.
7. Dry the samples to complete dryness using a vacuum concentrator.
8. Add 50 µl Buffer A, resuspend the peptides by incubating for 5 min in a ultrasonic water bath.

For a recent detailed overview of sample preparations methods for MS, *see* ref. [28].

3.4 Shotgun Mass Spectrometry

1. This protocol is optimized for the LC-MS/MS analysis of 1 µl sample corresponding to ~1 µg protein (*see* **Note 5**).
2. Separate the peptides on a 2 h gradient and run the mass spectrometer in data dependent acquisition (DDA) mode according to the instrument vendor's recommendations.

3.5 Data Analysis

1. Launch MaxQuant by double-clicking on "MaxQuant.exe".
2. Click "load" and select the MS data files in the file dialog (*see* **Note 6**).
3. Under the "Group-specific parameters" tab:
   (a) Click "Label-free quantification". In the dropdown menu, select "LFQ".
   (b) Click "Digestion". Ensure that "Trypsin/P" is the only entry in the right list.
   (c) Click "Instrument". Ensure that the instrument type is matching the used instrument (*see* **Note 7**).
   (d) Click "Modifications". Ensure that the right-hand list consists of "Oxidation (M)" and "Acetyl (Protein N-term)".

4. Under the "Global parameters" tab:
   (a) Click "Sequences":
      - Click "Add file" and select the FASTA protein database.
      - Ensure that the right-hand list consists of "Carbamidomethyl (C)".
   (b) Click "Identification". Set the "PSM FDR" to 0.01, and "Protein FDR" to 0.01.
5. Under the "Configuration" tab:
   (a) Click "Sequence databases":
      - Click "Add". On the right hand side, click "Select" and choose the fasta protein database. Type in the fasta file source in the "Source" field. Replace "Homo sapiens" for the appropriate host and pathogen species. Finally click "Modify table" to save this entry.
      - Click "Save changes".
6. Under the "Raw files" tab:
   (a) Click "Start" to start the analysis. Depending on the number of sample and size of the protein database, the analysis might take several hours.
7. Results are found in the tab-separated file combined/proteinGroups.txt.
   (a) The measured relative quantity of each protein is given in the "Intensity" column. This is very precise for comparing the concentration of a given protein between samples, but should not be used to compare levels between different proteins.
   (b) Protein IDs starting with "CON__" or "REV__" are known contaminants and mock proteins respectively. This status is also shown in the "Potential contaminant" and "Reverse" columns. Such proteins should not be used in the following analysis.
   (c) Many proteomics scientists consider proteins with only one supporting peptide dubious, these proteins should be used with caution.

4 Notes

1. Other proteinous fluids can also be used, for example saliva.
2. These solutions should be made fresh and used the same day.
3. Prepare 10% working solution in LC-grade water. Do not use plastics (tips, beakers or bottles) when handling concentrated FA.
4. Translated bacterial genomes can be found both in Uniprot (http://uniprot.org), but also in the Human Microbiome Project

(http://hmpdacc.org/), PANTHER (http://pantherdb.org) and Patric (http://patricbrc.org) databases. For host-translated genomes (human, mouse, etc.) we suggest using the UniProt KB reference proteomes.

5. The injection volume is dependent on the amount of bacteria and absorbed proteins. The total protein concentration of the sample homogenate can be estimated with protein assays, for example bicinchoninic acid (BCA) assay kits.
6. MaxQuant support the native data formats of several vendors. If the used instrument vendor is not in this list. MSConvert [29] might be used to convert the data files to the generic format mzXML, that is also supported by MaxQuant.
7. To maximize mass spectrometry search results, the search parameters and especially precursor and fragment tolerances should be adapted to the used method and instrument. If unsure, please consult with the instrument operator on the appropriate settings.

Acknowledgments

This work was supported by the Swedish Research Council (project 621-2012-3559).

References

1. Macheboeuf P, Buffalo C, Fu C-Y et al (2011) Streptococcal M1 protein constructs a pathological host fibrinogen network. Nature 472:64–68. doi:10.1038/nature09967
2. Cunningham MW (2000) Pathogenesis of group A streptococcal infections. Clin Microbiol Rev 13:470–511
3. Herwald H, Cramer H, Mörgelin M et al (2004) M protein, a classical bacterial virulence determinant, forms complexes with fibrinogen that induce vascular leakage. Cell 116:367–379
4. Carlsson F, Berggård K, Stålhammar-Carlemalm M, Lindahl G (2003) Evasion of phagocytosis through cooperation between two ligand-binding regions in Streptococcus pyogenes M protein. J Exp Med 198:1057–1068. doi:10.1084/jem.20030543
5. Sjöholm K, Karlsson C, Linder A, Malmström JA (2014) A comprehensive analysis of the Streptococcus pyogenes and human plasma protein interaction network. Mol Biosyst 10:1698–1708. doi:10.1039/c3mb70555b
6. Cole JN, Barnett TC, Nizet V, Walker MJ (2011) Molecular insight into invasive group A streptococcal disease. Nat Rev Microbiol 9:724–736. doi:10.1038/nrmicro2648
7. Malmström JA, Karlsson C, Nordenfelt P et al (2011) Streptococcus pyogenes in human plasma: adaptive mechanisms analyzed by mass spectrometry based proteomics. J Biol Chem 287:1415–1425. doi:10.1074/jbc.M111.267674
8. Nyberg P, Rasmussen M, Bjorck L (2004) Alpha2-macroglobulin-proteinase complexes protect Streptococcus pyogenes from killing by the antimicrobial peptide LL-37. J Biol Chem 279:52820–52823. doi:10.1074/jbc.C400485200
9. Kronvall G, Simmons A, Myhre EB, Jonsson S (1979) Specific absorption of human serum albumin, immunoglobulin A, and immunoglobulin G with selected strains of group A and G streptococci. Infect Immun 25:1–10
10. Kahn F, Mörgelin M, Shannon O et al (2008) Antibodies against a surface protein of Streptococcus pyogenes promote a pathological inflammatory response. PLoS Pathog 4:e1000149. doi:10.1371/journal.ppat.1000149
11. Nordenfelt P, Waldemarson S, Linder A et al (2012) Antibody orientation at bacterial surfaces is related to invasive infection. J Exp Med 209:2367–2381. doi:10.1084/jem.20120325

12. Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. Nature 422:198–207. doi:10.1038/nature01511
13. Gillette MA, Carr SA (2013) Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. Nat Methods 10:28–34. doi:10.1038/nmeth.2309
14. Picotti P, Bodenmiller B, Aebersold R (2013) Proteomics meets the scientific method. Nat Methods 10:24–27. doi:10.1038/nmeth.2291
15. Xu C, Ma B (2006) Software for computational peptide identification from MS-MS data. Drug Discov Today 11:595–600. doi:10.1016/j.drudis.2006.05.011
16. Nesvizhskii AI, Vitek O, Aebersold R (2007) Analysis and validation of proteomic data generated by tandem mass spectrometry. Nat Methods 4:787–797. doi:10.1038/nmeth1088
17. Käll L, Storey JD, MacCoss MJ, Noble WS (2008) Assigning significance to peptides identified by tandem mass spectrometry using decoy databases. J Proteome Res 7:29–34. doi:10.1021/pr700600n
18. Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods 4:207–214. doi:10.1038/nmeth1019
19. Craig R, Beavis RC (2004) TANDEM: matching proteins with tandem mass spectra. Bioinformatics 20:1466–1467. doi:10.1093/bioinformatics/bth092
20. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551–3567. doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2
21. Nesvizhskii AI, Keller A, Kolker E, Aebersold R (2003) A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem 75:4646–4658
22. Reiter L, Claassen M, Schrimpf SP et al (2009) Protein identification false discovery rates for very large proteomics data sets generated by tandem mass spectrometry. Mol Cell Proteomics 8:2405–2417. doi:10.1074/mcp.M900317-MCP200
23. Häkkinen J, Vincic G, Månsson O et al (2009) The proteios software environment: an extensible multiuser platform for management and analysis of proteomics data. J Proteome Res 8:3037–3043. doi:10.1021/pr900189c
24. Weisser H, Nahnsen S, Grossmann J et al (2013) An automated pipeline for high-throughput label-free quantitative proteomics. J Proteome Res 12:1628–1644. doi:10.1021/pr300992u
25. Mueller LN, Rinner O, Schmidt A et al (2007) SuperHirn - a novel tool for high resolution LC-MS-based peptide/protein profiling. Proteomics 7:3470–3480. doi:10.1002/pmic.200700057
26. Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26:1367–1372. doi:10.1038/nbt.1511
27. Zhang R, Barton A, Brittenden J, Huang JT, Crowther D (2010) Evaluation for computational platforms of LC-MS based label-free quantitative proteomics: A global view. J Proteomics Bioinform 3:260–265. doi:10.4172/jpb.1000149
28. Vowinckel J, Capuano F, Campbell K et al (2013) The beauty of being (label)-free: sample preparation methods for SWATH-MS and next-generation targeted proteomics. F1000Res 2:272. doi:10.12688/f1000research.2-272.v2
29. Kessner D, Chambers M, Burke R et al (2008) ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics 24:2534–2536. doi:10.1093/bioinformatics/btn323

http://www.springer.com/978-1-4939-6671-4

Bacterial Pathogenesis
Methods and Protocols
Nordenfelt, P.; Collin, M. (Eds.)
2017, XI, 357 p. 71 illus., 48 illus. in color., Hardcover
ISBN: 978-1-4939-6671-4
A product of Humana Press