TITLE OF THE ARTICLE:

Characterization of the biofilm landscape of *Bacillus subtilis* by spatial microproteomics

RUNNING TITLE:

Spatial proteoform characterization in biofilms

AUTHORS:

Kevin J. Zemaitis[1]*, Mowei Zhou[1,+], Sarah M. Yannarell[3,4], James M. Fulcher[1], Arunima Bhattacharjee[1], Marija Veličković[1], David J. Degnan[2], Elizabeth A. Shank[5], Christopher R. Anderton[1], William Kew[1], Ljiljana Paša-Tolić[1], Dušan Veličković[1]*

AFFILIATIONS:

[1]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington, USA

[2]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington, USA

[3]Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

[4]Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[5]Department of Systems Biology, University of Massachusetts Chan Medical School, Worcester, Massachusetts, USA

[+]Present affiliation: Department of Chemistry, Zhejiang University, Hangzhou, Zhejiang 310058, China

CORRESPONDING AUTHOR(S):

*Kevin Zemaitis – Email: kevin.zemaitis@pnnl.gov

*Dusan Veličković – Email: dusan.velickovic@pnnl.gov

ABSTRACT

Bulk proteomics has been demonstrated to differentiate subpopulations within pleiotropic bacterial colonies, yet advanced analyses by mass spectrometry imaging (MSI) hold even

greater promise for high-throughput phenotyping and differentiation to reshape biological discovery through visualization of biomolecular mechanisms. High mass resolving power and high spatial resolution analyses are now routine and confidently enable proteoform imaging directly from samples with minimal preparation. Pairing those analyses with bulk experimental libraries can provide high confidence in annotations of post-translational modifications (PTMs) and truncations, revealing their localization, unlocking a direct window into unknown biology at the microscale. However, this profiling is not commonplace for microbial species, considering the theoretical proteome of *Bacillus subtilis* was only partially mapped until recently. With little still known about the form and function of many of these proteins – let alone uncharacterized bacterial proteoforms, where PTMs and truncations on the same protein may possess unique physiological roles – there is a wealth of work to be done. With the joint application of top-down proteomics (TDP) and MSI, we outline preliminary results from microscale spatial proteomics on *B. subtilis* to probe a dynamic proteome and demonstrate the feasibility of detecting differentiated subpopulations based on their proteome.

MAIN BODY:

Bacterial biofilms form distinct microenvironments where sporulation, colonization, growth, and survival are governed by underlying mechanisms. Proteins within these biochemical mechanisms can be used to localize cellular phenotypes from the central core to the outer periphery of a colony.(1) These subpopulations are spatiotemporally dynamic and result from differential expression at the genomic level,(2) with mutations further driving heterogeneous responses.(3) Using a model organism for biofilm development,(4) we conducted TDP and matrix-assisted laser desorption/ionization (MALDI)-MSI on thin sections of a *B. subtilis* biofilm to explore the intact protein landscape revealing differentiated subpopulations.(5, 6) Through localization of proteins distributions within distinct mechanisms,(7) these subpopulations are identified in a fashion similar to spatial biotyping.(8) Similar works have focused upon small toxic peptides (e.g., sporulation delay protein, SDP; sporulation killing factor, SKF; and epipeptide, EPE), which traced cannibalistic cell-types in wild-type and knockout mutants, providing insight into the function and interdependence in sporulation and further biofilm development.(9) Here, we demonstrate discovery-based efforts for contextualization of spatial dependencies of intercellular proteins and proteins within the extracellular matrix,(10) providing insight into proteomic mechanisms behind cellular differentiation and biofilm production.(7)

MALDI-MSI experiments were performed on 20 µm cyrosectioned cross-sections of embedded *B. subtilis* biofilms at a spatial resolution of 20 µm **(Supplementary Information)**.

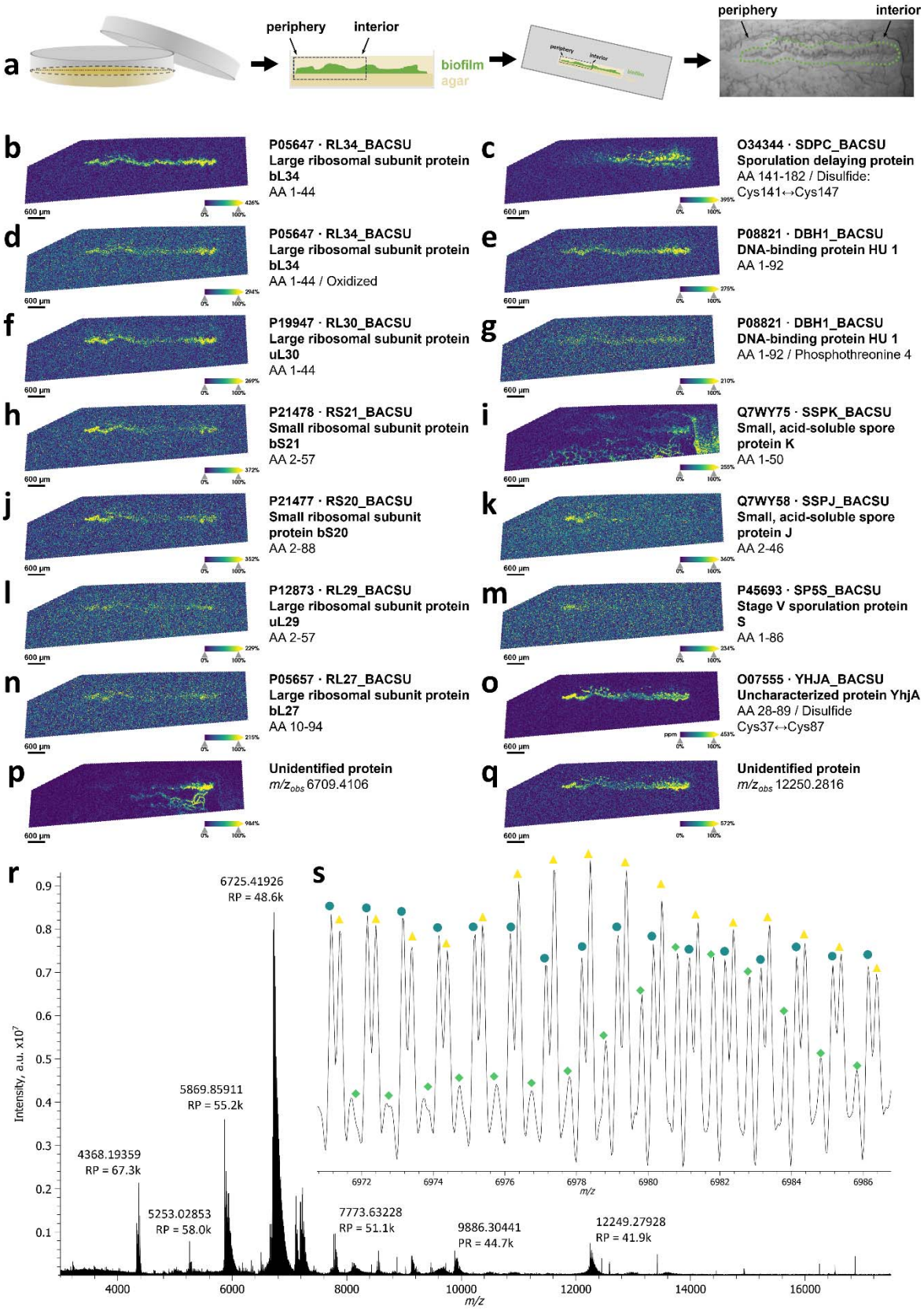
Previous analyses of these samples had identified distinct distributions of lipids and surfactants,(11) and here an optimized sample preparation **(Supplementary Information)** to detect intact protein distributions was applied **(Figure 1a)**. Through spatial proteomics we provide direct mechanistic insights into differentiated subpopulations. Analyses of these samples were enabled by coupling MALDI source to ultrahigh mass range (UHMR) high-field (HF) Orbitrap **(Supplementary Information)**,(12) yielding rich spectra with high intensities of intact proteins and proteoforms, up to 20 kDa in mass **(Figure 1r)**.(13) Subsequently, we performed semi-automated annotation and evaluation of the proteomic spatial heterogeneity within the biofilm.

Characteristic toxic peptides (e.g., SDP) are seen within the active interior of the biofilm **(Figure 1c)**,(14) while sporulating proteins **(Figure 1i, 1k, 1m)** and further uncharacterized proteins **(Figure 1o, 1q)** were found within the outer periphery. Overall, more than 285 unique isotopic envelopes from proteoforms were detected **(Figure 1r)** and annotations were made via UniProt, and experimentally generated TDP libraries – enabling precise annotation of PTMs and truncations.(15) Manual validation was performed after isotopic matching,(16) and sample-specific TDP libraries were imperative for the successful annotation of ion images. The presence of SDP **(Figure 1c)** in the central core of the colony can be indicative of cannibalistic cells recycling older generations, but the absence in the outer periphery and presence of characteristic sporulating proteins highly localized to the outermost periphery **(Figure 1i, 1k, 1m)**, suggests the biofilm is in early phases of sporulation. Stage V sporulation protein S can be overcome by other proteins or mutation,(17) but through colocalization of small acid-soluble spore (SASP) protein J **(Figure 1k)**, the presence of sporulated cells can be confirmed where SASPs are roughly 20% of the outer spore coat of *B. subtilis.*(18) These findings are supported by recent reports of cannibalistic tendencies that prolong and delay sporulation.(9)

TDP also detected sporulating delaying protein (SdpC) and immunity protein (SdpI) **(Supplementary Material)**, which offers immunity through sequestration of SDP.(19) Demonstrating that tandem TDP of microbial colonies complements MALDI-MSI analyses with coverage of larger proteoforms. When paired with laser capture microdissection (LCM), spatial/sub-sample specific information can be integrated.(20) Additionally, if paired with fluorescence microscopy a robust multimodal pipeline to validate subpopulations and phenotypes is created. Several ribosomal proteins, including large subunits **(Figure 1b, 1f, 1l, 1n)** were also routinely annotated with initiator methionine truncations, small subunits **(Figure 1h, 1j)** were also located throughout the main body of the biofilm. The broad detection of these

ribosomal proteins lends credence to spatial biotyper analyses, where speciation and localization of microbial cocultures is feasible based on proteomic signatures.(8)

Dozens of isotopic signatures are resolved by only a few hundred mDa **(Figure 1s)**, highlighting the importance of sample specific proteoform libraries and the high mass-resolving power achievable with the Fourier transform instruments. While current applications were limited by incompatibility of ITO glass slides and LCM-TDP workflows,(20, 21) future work should leverage spatial TDP with MALDI-MSI for proteoform-informed spatial proteomics.(15) For example, several annotations would benefit, such as the 7.1 kDa protein **(Figure 1q)** annotated as either uncharacterized protein YhjA (O07555) with a disulfide, or regulatory protein DegR (P68731) with higher mass error. Protein YhjA possesses a signal peptide (AA 1-27) which is cleaved, and a high confidence disulfide link (C37↔C87) which was annotated due to the mass shift. This is supported by AlphaFold structural predictions **(Supplementary Figure 1)**, even though it is not listed within UniProt.(22, 23) Functionally, DegR stabilizes DegU phosphorylation with links to increased protease activity,(24) and regulates transition of motile cells to sessile forms.(7) This links major pleiotropic mechanisms producing miner cells which degrade extracellular biopolymers and peptides,(25) and potential interactions with cannibalistic toxins noted to prolong sporulation.(9) However, DegR was not detected by TDP, possibly due to analysis of a different culture or growth period, while other proteins such as DegV (26) were identified **(Supplementary Material)**. Indeed, advanced measurement science is primed for application and further use of this pipeline can potentially unravel functions of many uncharacterized or understudied proteins.


a
periphery
interior
biofilm
agar
periphery
interior
b
P05647 · RL34_BACSU
Large ribosomal subunit protein bL34
AA 1-44
c
O34344 · SDPC_BACSU
Sporulation delaying protein
AA 141-182 / Disulfide: Cys141↔Cys147
d
P05647 · RL34_BACSU
Large ribosomal subunit protein bL34
AA 1-44 / Oxidized
e
P08821 · DBH1_BACSU
DNA-binding protein HU 1
AA 1-92
f
P19947 · RL30_BACSU
Large ribosomal subunit protein uL30
AA 1-44
g
P08821 · DBH1_BACSU
DNA-binding protein HU 1
AA 1-92 / Phosphothreonine 4
h
P21478 · RS21_BACSU
Small ribosomal subunit protein bS21
AA 2-57
i
Q7WY75 · SSPK_BACSU
Small, acid-soluble spore protein K
AA 1-50
j
P21477 · RS20_BACSU
Small ribosomal subunit protein bS20
AA 2-88
k
Q7WY58 · SSPJ_BACSU
Small, acid-soluble spore protein J
AA 2-46
l
P12873 · RL29_BACSU
Large ribosomal subunit protein uL29
AA 2-57
m
P45693 · SP5S_BACSU
Stage V sporulation protein S
AA 1-86
n
P05657 · RL27_BACSU
Large ribosomal subunit protein bL27
AA 10-94
o
O07555 · YHJA_BACSU
Uncharacterized protein YhjA
AA 28-89 / Disulfide
Cys37↔Cys87
p
Unidentified protein
m/z obs 6709.4106
q
Unidentified protein
m/z obs 12250.2816
600 µm
r
s
Intensity, a.u. x10^7
m/z
6725.41926
RP = 48.6k
5869.85911
RP = 55.2k
4368.19359
RP = 67.3k
5253.02853
RP = 58.0k
7773.63228
RP = 51.1k
9886.30441
PR = 44.7k
12249.27928
RP = 41.9k

119

Figure 1. The sample preparation is outlined within **(a)** showing colonies grown on MSgg media which were further embedded and cryosectioned at 20 µm as outlined in **Supplementary Information** and described in depth elsewhere.(11) After sample preparation, ion images are shown for each proteoform, the name, UniProt accession ID, and amino acid sequence with modifications (if any) are reported. Theoretical and experimental *m/z* values of the highest abundance isotope with calculated error can be found within **Supplementary Information**. **(b)** Large ribosomal subunit protein bL34 (P05647), **(c)** sporulation delaying protein (SDP; O34344), **(d)** oxidized large ribosomal subunit protein bL34 (P05647), **(e)** DNA-binding protein HU 1 (P08821), **(f)** large ribosomal subunit protein uL30 (P19947;), **(g)** phosphorylated DNA-binding protein HU 1 (P08821), **(h)** small ribosomal subunit protein bS21 (P21478;), **(i)** small acid-soluble spore protein K (Q7WY75), **(j)** small ribosomal subunit protein bS20 (P21477), **(k)** small acid-soluble spore protein J (Q7WY58), **(l)** large ribosomal subunit protein uL29 (P12873), **(m)** stage V sporulation protein S (P45693), **(n)** large ribosomal subunit protein bL27 (P05657), **(o)** a 7.1kDa protein, which is either uncharacterized protein YhjA (O07555) with a disulfide at Cys37↔Cys87 resulting in a [-2.01565 Da] shift, or regulatory protein DegR (P68731) with larger error, **(p)** a 6.1kDa unannotated protein, and **(q)** a 12.4kDa unannotated protein. The average spectrum from MALDI-MSI **(r)** is shown from *m/z* 3000 to 17500, with the complexity shown within an inset zoom **(s)** of the spectrum from *m/z* 6971 to 6987. Three distinct isotopic signatures are contained within this region, resulting from complex ion adduct and PTMs. **SMART** annotation:(27) **S** (step size, spot size, total pixels) = 20 µm, 12 µm x 15 µm, 33084 pixels; **M** (molecular confidence) = MS1, <5 ppm; **A** (annotations) = MS1 matching from TDP LC-MS/MS; **R** (resolving power) = 48.6k at *m/z* 6725.4; **T** (time of acquisition) = 286 min.

ACKNOWLEDGEMENTS:

The authors would also like to acknowledge Dr. Mikhail Belov at Spectroglyph, LLC as well as Dr. Gordon Anderson and Chris Anderson at GAA Custom Electronics, LLC for technical support for the MALDI source. We would also like to acknowledge Drs. Tobias Wörner, Kyle Fort, Maria Reinhardt-Szyba, and Alexander Makarov of Thermo Fisher Scientific for technical guidance and licensing of the instrument. We thank Dr. Matthew E. Monroe of Pacific Northwest National Laboratory for his assistance in depositing the raw proteomic data onto MassIVE. This work was performed at the Environmental Molecular Science Laboratory (EMSL), a Department of Energy (DOE) Office of Science User Facility sponsored by the Office of Biological and Environmental Research. This expanded research on samples from EUP 51105 (doi.org/10.46936/expl.proj.2019.51105/60000139, E.A.S.) was funded by the intramural program on project awards doi.org/10.46936/intm.proj.2019.51159/60000152 (W.K) and at EMSL (grid.436923.9) operated under Contract No. DE-AC05-76RL01830.

COMPETING INTERESTS:

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT:

The data presented within this article is publicly available and was uploaded to MassIVE, partner repository with the MassIVE accession: MSV000100434.

For now, manuscript reviewers can access the data via FTP using:

Server: massive-ftp.ucsd.edu

User: MSV000100434

Password: Spatial7303

SUPPLEMENTARY INFORMATION (50 words):

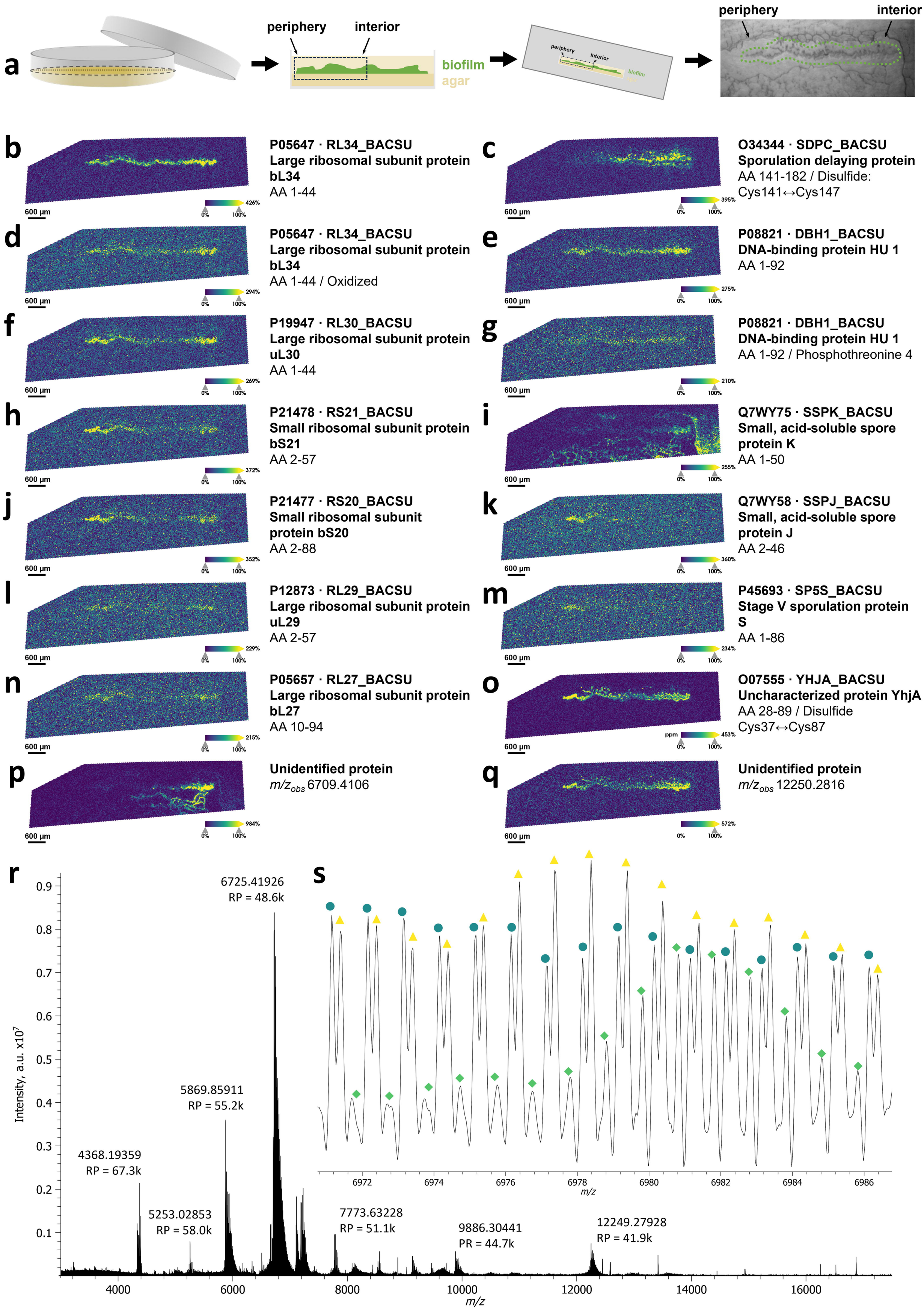
The following is provided as supplementary information for this article includes: (1) "Supplementary Information" outlining sample preparation, mass spectrometry methods, and data analysis, supplementary figure S1, and a supplementary table of highlighted annotations, and (2) "Supplementary Material" in an output table of the TDP results from both mass spectrometry instruments used.

REFERENCES:

1. van Dijl J, Hecker M. Bacillus subtilis: from soil bacterium to super-secreting cell factory. Microbial Cell Factories. 2013;12(1):3.
2. Dergham Y, Le Coq D, Nicolas P, Bidnenko E, Dérozier S, Deforet M, et al. Direct comparison of spatial transcriptional heterogeneity across diverse Bacillus subtilis biofilm communities. Nature Communications. 2023;14(1):7546.
3. Branda SS, González-Pastor JE, Ben-Yehuda S, Losick R, Kolter R. Fruiting body formation by Bacillus subtilis. Proceedings of the National Academy of Sciences. 2001;98(20):11621-6.
4. Vlamakis H, Chai Y, Beauregard P, Losick R, Kolter R. Sticking together: building a biofilm the Bacillus subtilis way. Nature Reviews Microbiology. 2013;11(3):157-68.
5. Bisht K, Wakeman CA. Discovery and Therapeutic Targeting of Differentiated Biofilm Subpopulations. Frontiers in Microbiology. 2019;10.
6. Ravikumar V, Nalpas NC, Anselm V, Krug K, Lenuzzi M, Šestak MS, et al. In-depth analysis of Bacillus subtilis proteome identifies new ORFs and traces the evolutionary history of modified proteins. Sci Rep. 2018;8(1):17246.
7. Mielich-Süss B, Lopez D. Molecular mechanisms involved in Bacillus subtilis biofilm formation. Environ Microbiol. 2015;17(3):555-65.
8. Gekenidis M-T, Studer P, Wüthrich S, Brunisholz R, Drissner D. Beyond the Matrix-Assisted Laser Desorption Ionization (MALDI) Biotyping Workflow: in Search of Microorganism-Specific Tryptic Peptides Enabling Discrimination of Subspecies. Applied and Environmental Microbiology. 2014;80(14):4234-41.

9. Friebel L, Knepper J-P, Becker NS, Abbaszade G, Stückrath K, Soltwisch J, et al. Cannibalism Shapes Biofilm Structure and Composition in Bacillus subtilis. bioRxiv. 2025:2025.03.21.644447.
10. Karygianni L, Ren Z, Koo H, Thurnheer T. Biofilm Matrixome: Extracellular Components in Structured Microbial Communities. Trends Microbiol. 2020;28(8):668-81.
11. Yannarell SM, Veličković D, Anderton CR, Shank EA. Direct Visualization of Chemical Cues and Cellular Phenotypes throughout Bacillus subtilis Biofilms. mSystems. 2021;6(6):e0103821.
12. Zemaitis KJ, Veličković D, Kew W, Fort KL, Reinhardt-Szyba M, Pamreddy A, et al. Enhanced Spatial Mapping of Histone Proteoforms in Human Kidney Through MALDI-MSI by High-Field UHMR-Orbitrap Detection. Analytical Chemistry. 2022;94(37):12604-13.
13. Smith LM, Kelleher NL, Linial M, Goodlett D, Langridge-Smith P, Ah Goo Y, et al. Proteoform: a single term describing protein complexity. Nature Methods. 2013;10(3):186-7.
14. Lamsa A, Liu W-T, Dorrestein PC, Pogliano K. The Bacillus subtilis cannibalism toxin SDP collapses the proton motive force and induces autolysis. Molecular microbiology. 2012;84(3):486-500.
15. Zemaitis KJ, Fulcher JM, Kumar R, Degnan DJ, Lewis LA, Liao Y-C, et al. Spatial top-down proteomics for the functional characterization of human kidney. Clinical Proteomics. 2025;22(1):9.
16. Degnan DJ, Zemaitis KJ, Lewis LA, McCue LA, Bramer LM, Fulcher JM, et al. IsoMatchMS: Open-Source Software for Automated Annotation and Visualization of High Resolution MALDI-MS Spectra. Journal of the American Society for Mass Spectrometry. 2023.
17. Resnekov O, Driks A, Losick R. Identification and characterization of sporulation gene spoVS from Bacillus subtilis. Journal of Bacteriology. 1995;177(19):5628-35.
18. Moeller R, Setlow P, Reitz G, Nicholson Wayne L. Roles of Small, Acid-Soluble Spore Proteins and Core Water Content in Survival of Bacillus subtilis Spores Exposed to Environmental Solar UV Radiation. Applied and Environmental Microbiology. 2009;75(16):5202-8.
19. Ellermeier CD, Hobbs EC, Gonzalez-Pastor JE, Losick R. A Three-Protein Signaling Pathway Governing Immunity to a Bacterial Cannibalism Toxin. Cell. 2006;124(3):549-59.
20. Liao Y-C, Fulcher JM, Degnan DJ, Williams SM, Bramer LM, Veličković D, et al. Spatially Resolved Top-Down Proteomics of Tissue Sections Based on a Microfluidic Nanodroplet Sample Preparation Platform. Molecular & Cellular Proteomics. 2023;22(2).
21. Veličković M, Day LZ, Zemaitis KJ, Attah IK, Burnum-Johnson KE, Anderton CR, et al. Enhanced Spatial Proteomics and Metabolomics from a Single Tissue Section Using MALDI-MSI and LCM-microPOTS Platforms. Analytical Chemistry. 2025;97(44):24715-23.
22. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-9.
23. Fleming J, Magana P, Nair S, Tsenkov M, Bertoni D, Pidruchna I, et al. AlphaFold Protein Structure Database and 3D-Beacons: New Data and Capabilities. Journal of Molecular Biology. 2025;437(15):168967.
24. Mukai K, Kawata-Mukai M, Tanaka T. Stabilization of phosphorylated Bacillus subtilis DegU by DegR. J Bacteriol. 1992;174(24):7954-62.
25. Kobayashi K. Gradual activation of the response regulator DegU controls serial expression of genes for flagellum formation and biofilm formation in Bacillus subtilis. Molecular Microbiology. 2007;66(2):395-409.
26. Knudsen GM, Olsen JE, Dons L. Characterization of DegU, a response regulator in Listeria monocytogenes, involved in regulation of motility and contributes to virulence. FEMS Microbiology Letters. 2004;240(2):171-9.

27. Xi Y, Sohn AL, Joignant AN, Cologna SM, Prentice BM, Muddiman DC. SMART: A data reporting standard for mass spectrometry imaging. Journal of Mass Spectrometry. 2023;58(2):e4904.


a
periphery
interior
biofilm
agar
b
P05647 · RL34_BACSU
Large ribosomal subunit protein bL34
AA 1-44
c
O34344 · SDPC_BACSU
Sporulation delaying protein
AA 141-182 / Disulfide: Cys141↔Cys147
d
P05647 · RL34_BACSU
Large ribosomal subunit protein bL34
AA 1-44 / Oxidized
e
P08821 · DBH1_BACSU
DNA-binding protein HU 1
AA 1-92
f
P19947 · RL30_BACSU
Large ribosomal subunit protein uL30
AA 1-44
g
P08821 · DBH1_BACSU
DNA-binding protein HU 1
AA 1-92 / Phosphothreonine 4
h
P21478 · RS21_BACSU
Small ribosomal subunit protein bS21
AA 2-57
i
Q7WY75 · SSPK_BACSU
Small, acid-soluble spore protein K
AA 1-50
j
P21477 · RS20_BACSU
Small ribosomal subunit protein bS20
AA 2-88
k
Q7WY58 · SSPJ_BACSU
Small, acid-soluble spore protein J
AA 2-46
l
P12873 · RL29_BACSU
Large ribosomal subunit protein uL29
AA 2-57
m
P45693 · SP5S_BACSU
Stage V sporulation protein S
AA 1-86
n
P05657 · RL27_BACSU
Large ribosomal subunit protein bL27
AA 10-94
o
O07555 · YHJA_BACSU
Uncharacterized protein YhjA
AA 28-89 / Disulfide Cys37↔Cys87
p
Unidentified protein
m/z_obs 6709.4106
q
Unidentified protein
m/z_obs 12250.2816
600 µm
0%
100%
r
s
Intensity, a.u. x10^7
m/z
6725.41926
RP = 48.6k
5869.85911
RP = 55.2k
4368.19359
RP = 67.3k
5253.02853
RP = 58.0k
7773.63228
RP = 51.1k
9886.30441
PR = 44.7k
12249.27928
RP = 41.9k