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

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 J. Zemaitis – Email: kevin.zemaitis@pnnl.gov

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

ABSTRACT

Bulk proteomics has been demonstrated to differentiate subpopulations within bacterial colonies, yet advanced analyses by mass spectrometry imaging (MSI) hold even greater promise for the future. This technology can enable high-throughput spatial phenotyping that can reshape biological discovery by providing visualization of components of various biomolecular mechanisms. With high mass resolving power and high spatial resolution analyses being routine, we can confidently enable intact protein 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 PTM localization within the samples unlocks a direct window into unknown biology at the microscale. However, top-down proteomics (TDP) is not commonplace for microbial species, largely due to challenges in identifying detected peptides and proteins; considering the theoretical proteome of even the well-studied model bacterium *Bacillus subtilis* was only partially mapped recently. With little still known about the form and function of many of these proteins – let alone proteoforms, where PTMs and truncations of the same protein may possess unique physiological roles – there is a wealth of work to be done. Here we jointly apply TDP and MSI to describe the microscale spatial proteomic landscape within *B. subtilis* and further demonstrate the feasibility of detecting differentiated subpopulations through proteoforms across the biofilm landscape.

INTRODUCTION:

Bacillus subtilis biofilms form distinct spatial microenvironments where unique biological functions such as sporulation, colonization, and metabolite production are governed by underlying mechanisms.[1,2] The proteins within these biochemical mechanisms can be used as proxies to localize cellular phenotypes from the central core to the outer periphery of a colony.[3] Not only are these subpopulations spatiotemporally dynamic, but they directly result from differential genetic expression,[4] with mutations driving further heterogeneous responses.[5] Using *B. subtilis* as a model organism for biofilm development,[6] we conducted TDP and matrix-assisted laser desorption/ionization (MALDI) mass spectrometry imaging (MSI) on thin sections of the biofilm to explore the intact protein landscape revealing differentiated

58 subpopulations.[7,8] While MSI has been broadly applied within microbiology for the detection
59 of primary and secondary metabolites, and highly abundant lipids.[9] The detection of intact
60 proteins distributions hold unique promise;[10] where species, subpopulations, and molecular
61 phenotypes can be identified by differentially abundant proteins within unique biochemical
62 mechanisms in a manner similar to spatial biotyping.[11]

63 Similar works have focused upon MALDI-MSI of small toxic peptides within *B. subtilis*
64 (e.g., sporulation delay protein, SDP; sporulation killing factor, SKF; and epipeptide, EPE),[12]
65 which traced cannibalistic cell-types in wild-type and knockout mutants. In addition to MALDI-
66 MSI mapping the localization of these cannibalistic toxins, further insight has been drawn into
67 the function and interdependence of these toxic peptides in the control of sporulation and
68 further biofilm development through genetic knockouts.[13] Demonstrating that in parallel
69 application MALDI-MSI is an enabling tool for microbiology, where the direct localization of
70 lipids, primary and secondary metabolites, and toxic peptides enhance more traditional
71 spectroscopy completed on cultured biofilms. Here, we demonstrate discovery-based efforts
72 for contextualization of spatial dependencies of intercellular and extracellular proteins,[14]
73 providing the location of larger abundant proteomic mechanisms responsible for cellular
74 differentiation and biofilm production.[10]

75 **METHODOLOGY:**

76 *Thin Bacterial Biofilms Production for Imaging*

77 For MALDI-MSI *B. subtilis* were cultured on MSgg growth medium as previously described,[2]
78 and a thin sectioning protocol was applied as described in greater depth elsewhere.[15,16]
79 Briefly, the biofilm agar blocks were quartered and transferred to molds (22-363-553, Fisher
80 Scientific). These molds were then snap frozen at -80 °C, overlaid with 4% (wt/vol) agarose
81 (Lonza Bioscience) and frozen at -80 °C for 15 min. Sectioning was completed as previously
82 reported using a Cryostar NX70 cryotome (Thermo Scientific, Runcorn, UK) and cross-sections
83 were taken at 20 µm thickness onto ITO slides (Bruker Daltonics, Bremen, DE). These sections
84 were stored at -80 °C until analysis.

85 *Top-Down Proteomics of B. subtilis*

86 For TDP, *B. subtilis* was cultured similarly to as previously described before on MSgg agar
87 plates.(1) After 48 hours of growth isolated *B. subtilis* colonies were scraped with a cell scraper
88 and collected in Eppendorf tubes. These tubes were subsequently frozen
89 and lyophilized. The lyophilized sample was bead beaten with a 3 mm tungsten carbide bead at
90 1,400 c.p.m. for 2 min using a HG-600 Geno/Grinder 2010 Tissue Homogenizer and
91 Cell Lyser (Cole Palmer, Metuchen, NJ), using cooled cryoblocks to inhibit sample
92 heating. Briefly, powdered sample was resuspended in Type 1 water and cold (−20 °C)
93 chloroform:methanol mixture (2:1, *v/v*) in chloroform-compatible 2 mL Sorenson MulTI
94 SafeSeal microcentrifuge tubes (Sorenson Bioscience, Salt Lake City, UT). The sample was vortex
95 mixed for 1 min, followed by incubation on ice for 5 min, then vortexed again for 1 min and
96 centrifuged at 10,000 g at 4 °C for 10 min. The top and bottom phases were discarded. The
97 remaining protein interlayer was washed with cold 100% methanol, vortexed, and centrifuged
98 for 5 min to pellet the protein. Supernatant was discarded. The protein interlayer was
99 resuspended in 8 M urea prepared in 50 mM ammonium bicarbonate (pH 8). Sample was
100 cleaned up using Amicon Ultra centrifugal filtration tubes, 100K followed by 10k molecular
101 weight cutoff (Millipore, Bedford, MA). Protein concentration was determined using a
102 bicinchoninic acid (BCA) assay (Thermo Scientific, Waltham, MA, USA) and samples were diluted
103 to 0.5 µg/µL in water with 0.1% FA.

104 Liquid chromatography mass spectrometry (LC-MS) was completed on both an Exploris 480
105 Orbitrap MS (Thermo Scientific, Bremen, DE) and a Lumos Tribrid Orbitrap MS (Thermo
106 Scientific, San Jose, CA) injecting protein onto a NanoAcquity (Waters, Millford, MA) with dual
107 pumps configuration. Binary solvents for mobile phase (MP) A and B were 0.2% formic acid in
108 water (MPA), and 0.2% formic acid in acetonitrile (MPB) for both pumps. The stationary phase
109 material for in-house packed reversed phase LC columns was C2 particle (SMTC2MEB2-3-300,
110 Separation Methods Technologies, Newark, DE). 10 µl injections were loaded onto a short trap
111 column (150 µm i.d., 5 cm long) and desalted with a flow of 3 µl/min with 5% MPB for 5 min.
112 Then the trap column was switched to the analytical column (100 µm i.d., 50 cm long) with flow
113 rate of 0.3 µl/min. Separation gradient started at 5% MPB to 15% MPB over 1 minute and was
114 linearly ramped to 90% MPB over 89 min.

115 The Exploris 480 Orbitrap MS (Thermo Scientific, Bremen, DE) was operated with the following
116 source parameters including a spray voltage of 2.2 kV, transfer capillary temperature of 275 ˚C,
117 and ion funnel RF amplitude set to 60%. The instrument was set to "intact protein" application
118 mode with standard pressures, and data was collected as full profile. MS1 acquisitions were
119 collected with 5 microscans at a mass resolution setting of 120k over a scan range of *m/z* 500 to
120 2000 with a normalized AGC target of 400%. Data dependent MS2 scans were collected with 2
121 microscans over a scan range of *m/z* 300 to 2000 with a normalized AGC target of 1000%. MS2
122 used a dynamic exclusion after 1 time over a duration of 40s with an isolation window of ±2
123 *m/z* with fragmentation by HCD at a fixed normalized collision energy of 35%.

124 The Lumos Tribrid Orbitrap MS (Thermo Scientific, San Jose, CA) was operated with the
125 following source parameters including a spray voltage of 2.2 kV, transfer capillary temperature
126 of 275 ˚C, and ion funnel RF amplitude set to 50%. The instrument was set to "peptide"
127 application mode with standard pressures, and data was collected as full profile. MS1
128 acquisitions were collected with 3 microscans at a mass resolution setting of 120k over a scan
129 range of *m/z* 400 to 2000 with a normalized AGC target of 200%. Data dependent MS2 scans
130 were collected with 1 microscans over a scan range of *m/z* 300 to 2000 with a normalized AGC
131 target of 500% and maximum injection time of 118 ms. MS2 used a dynamic exclusion after 1
132 time over a duration of 60s with an isolation window of ±1 *m/z* with fragmentation by HCD at a
133 stepped collision energy of 20V, 30V, and 40V.

134 MS instrument ".RAW" files were first converted into ".mzML" format with MSConvert before
135 processing with the open-search software TopPIC Suite (version 1.4.13).[17] For TopFD
136 deconvolution, a precursor window of 1 and 2 *m/z* corresponding to the Lumos and Exploris
137 respectively, error tolerance of 0.02 *m/z*, MS1 signal to noise ratio (SNR) of 3, MS2 SNR of 1,
138 maximum charge of 30, and maximum monoisotopic mass of 50,000 Da were applied. Specific
139 settings for the TopPIC search included an error tolerance of 15 ppm, PrSM cluster tolerance of
140 3.2 Da, FDR of 0.01, and a maximum and minimum unknown mass shift of ±500 Da, and a
141 maximum of one unknown mass shift.

142 *Preparation for Spatial Proteomics*

143 Slides were prepared for MALDI by vacuum desiccating at -15 inHg for 30 min with subsequent
144 washing in a two-part series for prior to matrix deposition. Washing was completed with
145 submersion in a solution of 70% ethanol for 30 sec, then submersion in a solution of 100%
146 ethanol for 30 sec. After which, the biofilms were then dried by a gentle stream of nitrogen for
147 30 sec. A M5 Sprayer (HTX Technologies, Chapel Hill, NC) was then used to spray the surface
148 with 5 crisscross (CC) passes of a solution of 5% acetic acid (v/v) in 50% ethanol with a flow rate
149 of 150 µL/min, a nozzle temperature of 30.0 °C, and the velocity set to 1,250 mm/min with 10
150 PSI of nitrogen gas to boost proteoform signal intensity.[18] A 5 s drying period was added
151 between each of the passes with a 3 mm track spacing, the same sprayer was then immediately
152 used to deposit 15 mg/mL 2,5-DHA in 90% acetonitrile with 0.2% TFA. Only the supernatant of
153 the dissolved matrix was sprayed after sonication and centrifugation. The flow rate of the
154 matrix was 150 µL/min with a nozzle temperature of 30.0 °C, and the velocity was set to 1,300
155 mm/min with 10 PSI of nitrogen gas. The matrix was applied with a 2 mm track spacing and a
156 crisscross pattern, and the matrix coverage was calculated to be 277 µg/cm^2. The matrix was
157 then recrystallized with 5% acetic acid solution (v/v) in water at 38.5 °C and dried for 3.5 min,
158 using an apparatus similar to that previously reported.[19]

159 *MALDI Mass Spectrometry Imaging of Sectioned Biofilm*s

160 An elevated pressure (EP) MALDI source with an atmospheric pressure ionization inlet
161 (Spectroglyph, LLC, Kennewick, WA) was mounted on a research grade Q Exactive HF Orbitrap
162 MS (Thermo Scientific, Bremen, DE), this instrument was upgraded with ultrahigh mass range
163 (UHMR) boards and operated under custom privileges licenses. This platform is described in
164 greater depth elsewhere,[20] briefly, the MALDI source was operated at 7.8 Torr, with fore
165 vacuum of the spectrometer at 1.0 mbar, and an ultrahigh vacuum adjusted to 2.54e-10 mbar
166 with argon gas. Spectra were acquired over a *m/z* range of 2,000 to 20,000, and acquisitions
167 were carried out with preset resolution of 240k, resulting in a transient of 512 ms, and
168 observed mass resolving power of 45k at *m/z* 6,725. The laser diode of the Explorer One 349
169 nm Nd:YLF (Spectra-Physics, Milpitas, California) was set to output an average of 1.1 µJ/pulse at
170 1 kHz, and ion were accumulated for 500 ms – resulting in 500 laser shots per pixel. Imaging
171 was carried out at 20 µm spatial resolution, and the resultant desorption of the crater

172 measured by light microscopy was found to be roughly 12 µm × 15 µm, which did not
173 oversample the surface.

174 Visualization was completed with automatic import of the .RAW acquisition file and .xml of the
175 pixel coordinates into SCiLS Lab Pro (v.2021c, Bruker Daltonics, Bremen, DE). Ion images were
176 generated by extracting singular isotopic peaks for the proteoforms. For proteoform
177 assignments, an experimental library of proteoforms from TDP LC-MS/MS were used as custom
178 databases and isotopic profiles were manually matched and mass shifts annotated. Mozaic
179 (v.2023.4.0.b3, Spectroswiss, Lausanne, CH) was used to average the 34,698 acquisitions within
180 the MALDI-MSI experiment, after averaging the spectra was baseline corrected using the
181 model-free method with a median size of 250, and other default parameters for number of
182 points, sigma, iterations and level (11, 5, 1, 1). A developed package IsoMatchMS, was used to
183 generate candidate matches based on isotopic distribution matching, this package is explained
184 in greater depth elsewhere.[21] Briefly, an entire average of the acquisition file was exported as
185 a .mzML file after processing. This file was imported with annotations and accurate masses with
186 <5 ppm error were considered, and further filtering of the matches occurred with Pearson
187 correlation score of ≥0.7. Further manual verification of annotated and non-annotated
188 proteoforms were searched against multiple adducts.

189 **RESULTS AND DISCUSSION:**

190 Through spatial proteomics we provide direct insights into differentiated
191 subpopulations. Previous analyses of these samples had identified distinct distributions of lipids
192 and surfactants,[2] and here an optimized sample preparation which included mild ethanol
193 fixation enabled the detection of intact protein distributions **(Figure 1a)**, and further desalting
194 was found to not be necessary unlike within mammalian and plant tissues previously analyzed
195 on this platform, or within colonies dried directly onto ITO slides and analyzed from the top
196 down [22]. Analyses of these samples were enabled by coupling a MALDI source to a research
197 grade UHMR HF Orbitrap MS,[18] this combination yields rich spectra with high intensities of
198 intact proteins and proteoforms, up to 20 kDa in mass **(Figure 1r)**, although the highest
199 detected species from the biofilm was roughly 13.5 kDa.[23] Subsequently, we performed semi-
200 automated annotation and evaluation of the proteomic spatial heterogeneity within the

201 biofilm. Overall, more than 285 unique isotopic envelopes from proteoforms were detected
202 **(Figure 1r)** and can be annotated via UniProt or experimentally generated TDP libraries –
203 enabling precise annotation of PTMs and truncations.[24] Manual validation was performed
204 after isotopic matching,[21] and sample-specific TDP libraries were imperative for the
205 successful annotation of ion images.

206 Of the proteoform distributions, characteristic toxic peptides (e.g., SDP) are seen within
207 the active interior of the biofilm **(Figure 1c)**,[25] while sporulating proteins **(Figure 1i, 1k, 1m)**
208 and further uncharacterized proteins **(Figure 1o, 1q)** were found within the outer periphery.
209 The presence of SDP **(Figure 1c)** in the central core of the colony can be indicative of
210 cannibalistic cells recycling older generations, but the absence within the outer periphery and
211 presence of characteristic sporulating proteins highly localized to the outermost periphery
212 **(Figure 1i, 1k, 1m)**, suggests the biofilm is in early phases of sporulation. Stage V sporulation
213 protein S can be overcome by other proteins or mutation,[26] but through colocalization of
214 small acid-soluble spore (SASP) protein J **(Figure 1k)**, the presence of sporulated cells can be
215 confirmed where SASPs are roughly 20% of the outer spore coat of *B. subtilis.*[27] These
216 findings are supported by recent reports of cannibalistic toxins tend to prolong and delay
217 sporulation,[13] further demonstrating the utility of the detection of intact proteoforms for
218 spatial subtyping of populations.

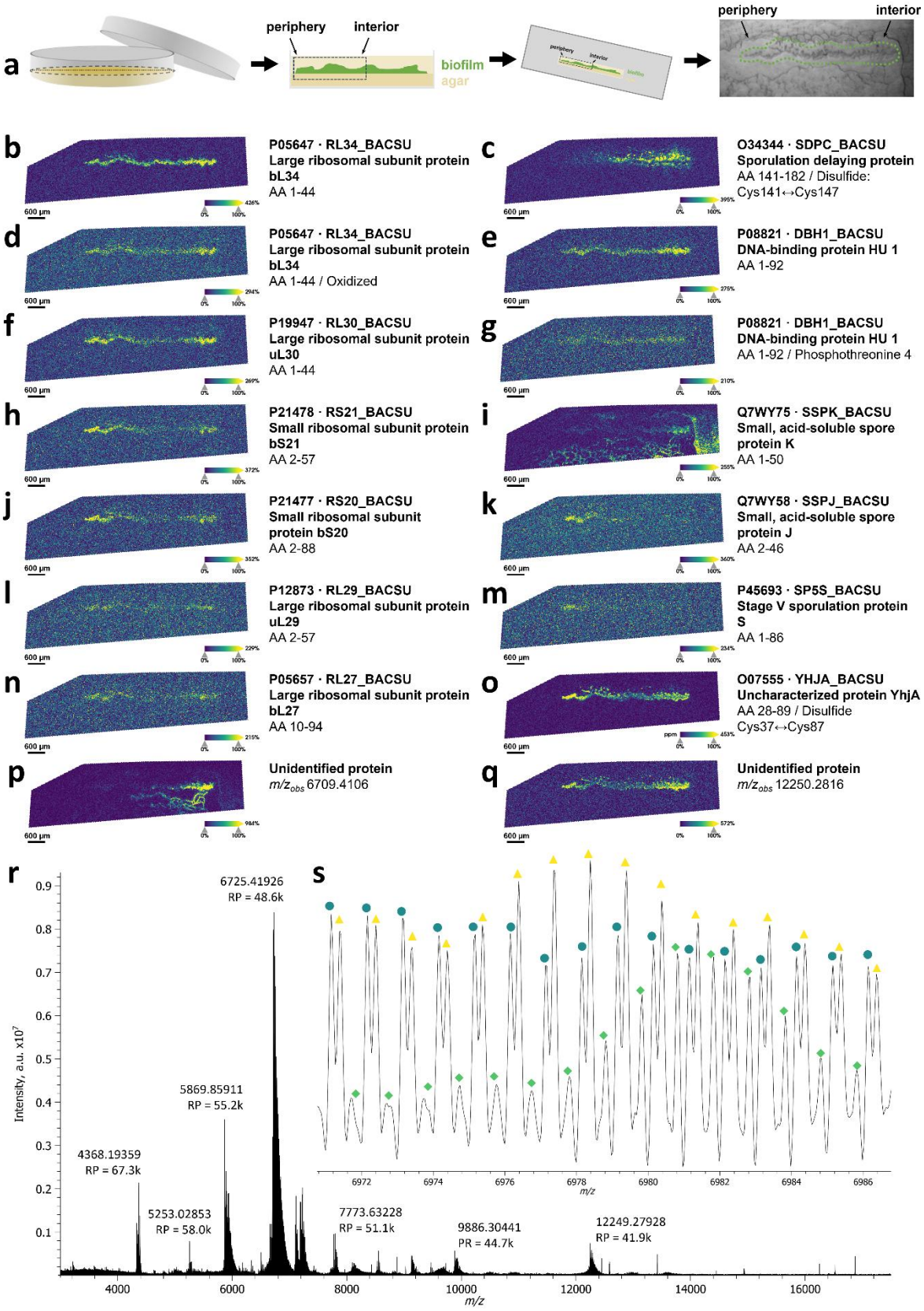
219 TDP also detected sporulating delaying protein (SdpC) and immunity protein (SdpI)
220 **(Supplementary Material)**, which offers immunity to cells through the sequestration of
221 SDP.[28] Demonstrating that tandem TDP of microbial colonies can complement MALDI-MSI
222 analyses with coverage of larger proteoforms. When paired with laser capture microdissection
223 (LCM), spatial/sub-sample specific proteoform information can be integrated readily.[29]
224 Additionally, if paired with fluorescence microscopy a robust multimodal pipeline to validate
225 subpopulations and phenotypes is created. Several ribosomal proteins, including large subunits
226 **(Figure 1b, 1f, 1l, 1n)** were also routinely annotated with initiator methionine truncations, and
227 small subunits **(Figure 1h, 1j)** were also located throughout the main body of the biofilm. The
228 broad detection of these ribosomal proteins lends credence to spatial biotyper analyses,[11]
229 where speciation and localization of microbial cultures is feasible based on proteomic

230 signatures. If these cocultured colonies of bacteria and fungi are analyzed, or when working
231 upon mutant knockouts or bioengineered strains, this information can be vital to understanding
232 intermicrobial colonization and organization while further informing upon the cellular
233 processes and phenotypes.

234 We identified dozens of isotopic signatures which are resolved being separated by only
235 a few hundred mDa throughout the mass range **(Figure 1s)**, highlighting the importance of
236 sample specific proteoform libraries and high mass-resolving power achievable with the Fourier
237 transform (FT)-MS instruments. While current applications were limited by incompatibility of
238 ITO glass slides and LCM-TDP workflows,[29,30] future works should leverage spatial LCM-TDP
239 with MALDI-MSI for proteoform-informed spatial proteomics.[24] For example, several
240 annotations would benefit from such serial analyses, such as the 7.1 kDa protein **(Figure 1q)**
241 annotated as either uncharacterized protein YhjA (O07555) with a putative disulfide, or
242 regulatory protein DegR (P68731) with higher mass error. While protein YhjA possesses a signal
243 peptide (AA 1-27) which is cleaved, and a high confidence disulfide link (C37↔C87) which was
244 annotated due to the mass shift, and is supported by AlphaFold structural predictions
245 **(Supplementary Figure 1)** even though it is not listed within UniProt.[31,32] Functionally, DegR
246 stabilizes DegU phosphorylation with links to increased protease activity,[33] and regulates
247 transition of motile cells to sessile forms.[10] While mass spectral error is higher than current
248 thresholds, this demonstrates the limits of high-resolution accurate mass isotopic pattern
249 matching. In either case, these links major mechanisms producing miner cells which degrade
250 extracellular biopolymers and peptides[34] and potential interactions with cannibalistic toxins
251 noted to prolong sporulation.[13] It should be noted DegR was not detected by TDP, possibly
252 due to analysis of a different culture or growth period, while other proteins such as DegV [35]
253 were identified **(Supplementary Material)**. Overall, these results demonstrate that advanced
254 measurement science is primed for future microbial applications and further use of this
255 instrumentation and pipeline can potentially unravel functions of many uncharacterized or
256 understudied proteins with proper experimental design.

257 **CONCLUSIONS:**

258 This study demonstrates that proteoform-informed spatial proteomics can resolve the
259 biochemical heterogeneity that governs microbial differentiation within *B. subtilis* biofilms. By
260 integrating MALDI-MSI with sample specific TDP libraries, we detected more than 285
261 proteoforms and revealed several distinct spatial distributions corresponding to cannibalistic,
262 sporulating, and metabolically active subpopulations. These results show that
263 proteoform-informed MSI can directly map functional microenvironments, offering mechanistic
264 insight into the temporal progression of biofilm development. While these experiments were
265 completed within a model system of *B. subtilis*, these methods for MALDI-MSI and TDP are
266 universal and can be broadly applied to any microbial system that can be cultured on solid
267 media. Additionally, while this is functionally limited to small proteoforms (< 20 kDa) for MALDI-
268 MSI, the information gathered at fine spatial resolutions significantly enhances our
269 understanding of the metabolic gradients and biomolecular mechanisms that shape community
270 structure. The ability to localize toxic peptides, sporulation-associated proteins, ribosomal
271 subunits, and uncharacterized proteins underscores the power of this approach to uncover
272 both known and previously unappreciated contributors to microbial physiology.


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

273

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.[2] 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:[36] **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

303 performed at the Environmental Molecular Science Laboratory (EMSL), a Department of Energy
304 (DOE) Office of Science User Facility sponsored by the Office of Biological and Environmental
305 Research. This expanded research on samples from EUP 51105
306 (doi.org/10.46936/expl.proj.2019.51105/60000139, E.A.S.) was funded by the intramural
307 program on project awards doi.org/10.46936/intm.proj.2019.51159/60000152 (W.K) and at
308 EMSL (grid.436923.9) operated under Contract No. DE-AC05-76RL01830.

309 **COMPETING INTERESTS:**

310 The authors declare no competing interests.

311 **DATA AVAILABILITY STATEMENT:**

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

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

315 Server: massive-ftp.ucsd.edu

316 User: MSV000100434

317 Password: Spatial7303

318 **SUPPLEMENTARY INFORMATION (50 words):**

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

324 **REFERENCES:**

325 1. Yannarell SM, Beaudoin ES, Talley HS, Schoenborn AA, Orr G, Anderton CR, Chrisler WB, Shank EA
326 (2023) Extensive cellular multi-tasking within Bacillus subtilis biofilms. mSystems 8 (4):e00891-
327 00822. doi:doi:10.1128/msystems.00891-22

2. Yannarell SM, Veličković D, Anderton CR, Shank EA (2021) Direct Visualization of Chemical Cues and Cellular Phenotypes throughout Bacillus subtilis Biofilms. mSystems 6 (6):e0103821. doi:10.1128/mSystems.01038-21
3. van Dijl J, Hecker M (2013) Bacillus subtilis: from soil bacterium to super-secreting cell factory. Microbial Cell Factories 12 (1):3. doi:10.1186/1475-2859-12-3
4. Dergham Y, Le Coq D, Nicolas P, Bidnenko E, Dérozier S, Deforet M, Huillet E, Sanchez-Vizuete P, Deschamps J, Hamze K, Briandet R (2023) Direct comparison of spatial transcriptional heterogeneity across diverse Bacillus subtilis biofilm communities. Nature Communications 14 (1):7546. doi:10.1038/s41467-023-43386-w
5. Branda SS, González-Pastor JE, Ben-Yehuda S, Losick R, Kolter R (2001) Fruiting body formation by Bacillus subtilis. Proceedings of the National Academy of Sciences 98 (20):11621-11626
6. Vlamakis H, Chai Y, Beauregard P, Losick R, Kolter R (2013) Sticking together: building a biofilm the Bacillus subtilis way. Nature Reviews Microbiology 11 (3):157-168
7. Bisht K, Wakeman CA (2019) Discovery and Therapeutic Targeting of Differentiated Biofilm Subpopulations. Frontiers in Microbiology 10. doi:10.3389/fmicb.2019.01908
8. Ravikumar V, Nalpas NC, Anselm V, Krug K, Lenuzzi M, Šestak MS, Domazet-Lošo T, Mijakovic I, Macek B (2018) In-depth analysis of Bacillus subtilis proteome identifies new ORFs and traces the evolutionary history of modified proteins. Sci Rep 8 (1):17246. doi:10.1038/s41598-018-35589-9
9. Dunham SJB, Ellis JF, Li B, Sweedler JV (2017) Mass Spectrometry Imaging of Complex Microbial Communities. Acc Chem Res 50 (1):96-104. doi:10.1021/acs.accounts.6b00503
10. Mielich-Süss B, Lopez D (2015) Molecular mechanisms involved in Bacillus subtilis biofilm formation. Environ Microbiol 17 (3):555-565. doi:10.1111/1462-2920.12527
11. Gekenidis M-T, Studer P, Wüthrich S, Brunisholz R, Drissner D (2014) 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 80 (14):4234-4241. doi:doi:10.1128/AEM.00740-14
12. Si T, Li B, Zhang K, Xu Y, Zhao H, Sweedler JV (2016) Characterization of Bacillus subtilis Colony Biofilms via Mass Spectrometry and Fluorescence Imaging. J Proteome Res 15 (6):1955-1962. doi:10.1021/acs.jproteome.6b00127
13. Friebel L, Knepper J-P, Becker NS, Abbaszade G, Stückrath K, Soltwisch J, Müller S, Dreisewerd K, Mascher T (2025) Cannibalism Shapes Biofilm Structure and Composition in Bacillus subtilis. bioRxiv:2025.2003.2021.644447. doi:10.1101/2025.03.21.644447
14. Karygianni L, Ren Z, Koo H, Thurnheer T (2020) Biofilm Matrixome: Extracellular Components in Structured Microbial Communities. Trends Microbiol 28 (8):668-681. doi:10.1016/j.tim.2020.03.016
15. Marlow VL, Cianfanelli FR, Porter M, Cairns LS, Dale JK, Stanley-Wall NR (2014) The prevalence and origin of exoprotease-producing cells in the Bacillus subtilis biofilm. Microbiology 160 (Pt 1):56
16. Vlamakis H, Aguilar C, Losick R, Kolter R (2008) Control of cell fate by the formation of an architecturally complex bacterial community. Genes & development 22 (7):945-953
17. Kou Q, Xun L, Liu X (2016) TopPIC: a software tool for top-down mass spectrometry-based proteoform identification and characterization. Bioinformatics 32 (22):3495-3497. doi:10.1093/bioinformatics/btw398
18. Zemaitis KJ, Veličković D, Kew W, Fort KL, Reinhardt-Szyba M, Pamreddy A, Ding Y, Kaushik D, Sharma K, Makarov AA, Zhou M, Paša-Tolić L (2022) Enhanced Spatial Mapping of Histone Proteoforms in Human Kidney Through MALDI-MSI by High-Field UHMR-Orbitrap Detection. Analytical Chemistry 94 (37):12604-12613. doi:10.1021/acs.analchem.2c01034
19. Yang J, Caprioli RM (2011) Matrix Sublimation/Recrystallization for Imaging Proteins by Mass Spectrometry at High Spatial Resolution. Analytical Chemistry 83 (14):5728-5734. doi:10.1021/ac200998a

20. Zemaitis KJ, Veličković D, Kew W, Fort KL, Reinhardt-Szyba M, Pamreddy A, Ding Y, Kaushik D, Sharma K, Makarov AA, Zhou M, Paša-Tolić L (2022) Enhanced Spatial Mapping of Histone Proteoforms in Human Kidney Through MALDI-MSI by High-Field UHMR-Orbitrap Detection. Analytical Chemistry. doi:10.1021/acs.analchem.2c01034

21. Degnan DJ, Zemaitis KJ, Lewis LA, McCue LA, Bramer LM, Fulcher JM, Veličković D, Paša-Tolić L, Zhou M (2023) IsoMatchMS: Open-Source Software for Automated Annotation and Visualization of High Resolution MALDI-MS Spectra. Journal of the American Society for Mass Spectrometry. doi:10.1021/jasms.3c00180

22. Veličković D, Zemaitis KJ, Bhattacharjee A, Anderton CR (2024) Mass spectrometry imaging of natural carbonyl products directly from agar-based microbial interactions using 4-APEBA derivatization. mSystems 9 (1):e00803-00823. doi:doi:10.1128/msystems.00803-23

23. Smith LM, Kelleher NL, Linial M, Goodlett D, Langridge-Smith P, Ah Goo Y, Safford G, Bonilla* L, Kruppa G, Zubarev R, Rontree J, Chamot-Rooke J, Garavelli J, Heck A, Loo J, Penque D, Hornshaw M, Hendrickson C, Pasa-Tolic L, Borchers C, Chan D, Young* N, Agar J, Masselon C, Gross* M, McLafferty F, Tsybin Y, Ge Y, Sanders* I, Langridge J, Whitelegge* J, Marshall A, The Consortium for Top Down P (2013) Proteoform: a single term describing protein complexity. Nature Methods 10 (3):186-187. doi:10.1038/nmeth.2369

24. Zemaitis KJ, Fulcher JM, Kumar R, Degnan DJ, Lewis LA, Liao Y-C, Veličković M, Williams SM, Moore RJ, Bramer LM, Veličković D, Zhu Y, Zhou M, Paša-Tolić L (2025) Spatial top-down proteomics for the functional characterization of human kidney. Clinical Proteomics 22 (1):9. doi:10.1186/s12014-025-09531-x

25. Lamsa A, Liu W-T, Dorrestein PC, Pogliano K (2012) The Bacillus subtilis cannibalism toxin SDP collapses the proton motive force and induces autolysis. Molecular microbiology 84 (3):486-500. doi:10.1111/j.1365-2958.2012.08038.x

26. Resnekov O, Driks A, Losick R (1995) Identification and characterization of sporulation gene spoVS from Bacillus subtilis. Journal of Bacteriology 177 (19):5628-5635. doi:doi:10.1128/jb.177.19.5628-5635.1995

27. Moeller R, Setlow P, Reitz G, Nicholson Wayne L (2009) 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 75 (16):5202-5208. doi:10.1128/AEM.00789-09

28. Ellermeier CD, Hobbs EC, Gonzalez-Pastor JE, Losick R (2006) A Three-Protein Signaling Pathway Governing Immunity to a Bacterial Cannibalism Toxin. Cell 124 (3):549-559. doi:10.1016/j.cell.2005.11.041

29. Liao Y-C, Fulcher JM, Degnan DJ, Williams SM, Bramer LM, Veličković D, Zemaitis KJ, Veličković M, Sontag RL, Moore RJ, Paša-Tolić L, Zhu Y, Zhou M (2023) Spatially Resolved Top-Down Proteomics of Tissue Sections Based on a Microfluidic Nanodroplet Sample Preparation Platform. Molecular & Cellular Proteomics 22 (2). doi:10.1016/j.mcpro.2022.100491

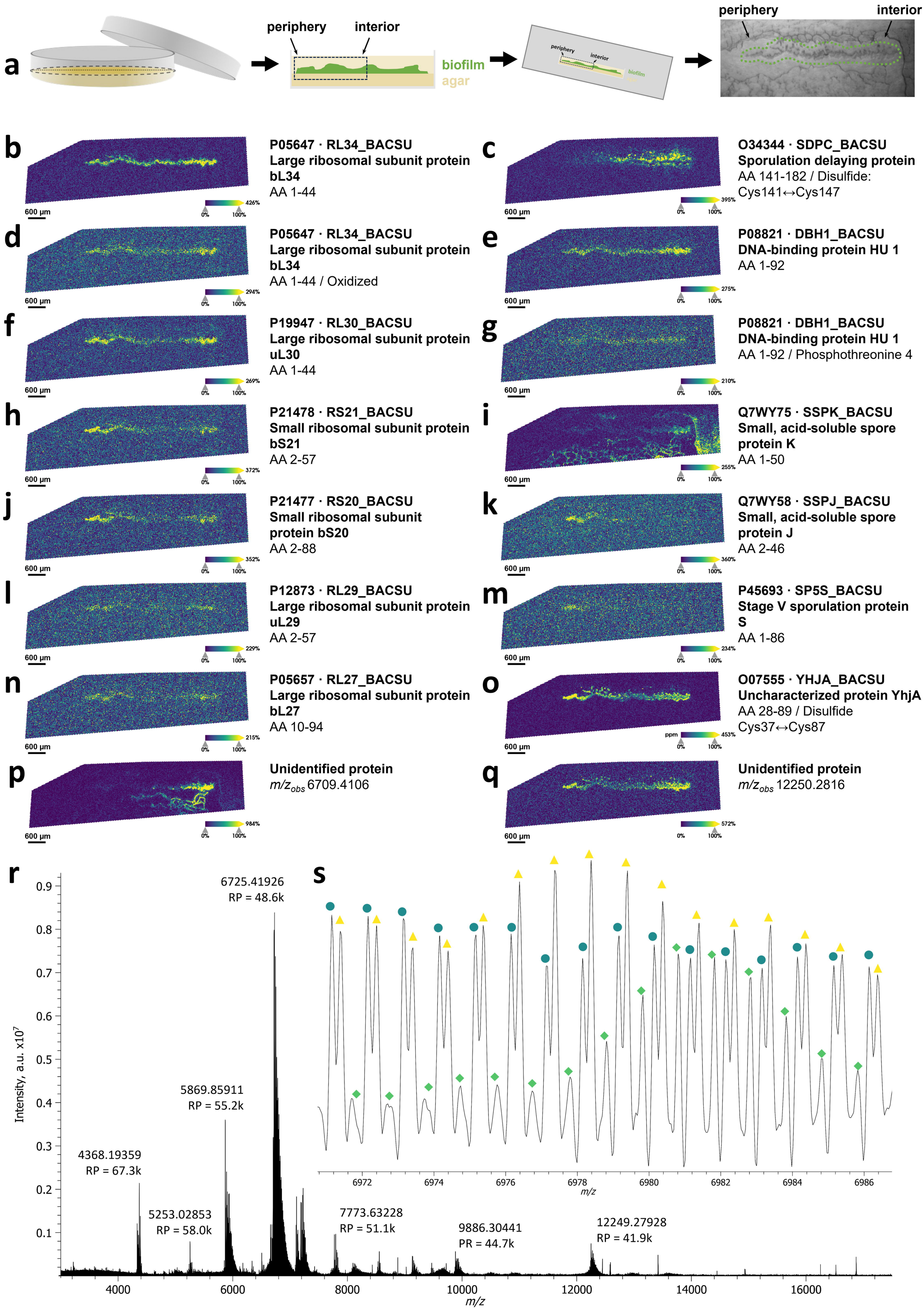
30. Veličković M, Day LZ, Zemaitis KJ, Attah IK, Burnum-Johnson KE, Anderton CR, Veličković D (2025) Enhanced Spatial Proteomics and Metabolomics from a Single Tissue Section Using MALDI-MSI and LCM-microPOTS Platforms. Analytical Chemistry 97 (44):24715-24723. doi:10.1021/acs.analchem.5c05005

31. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D (2021) Highly accurate protein structure prediction with AlphaFold. Nature 596 (7873):583-589. doi:10.1038/s41586-021-03819-2

32. Fleming J, Magana P, Nair S, Tsenkov M, Bertoni D, Pidruchna I, Lima Afonso MQ, Midlik A, Paramval U, Žídek A, Laydon A, Kovalevskiy O, Pan J, Cheng J, Avsec Ž, Bycroft C, Wong LH, Last M, Mirdita M,

424 Steinegger M, Kohli P, Váradi M, Velankar S (2025) AlphaFold Protein Structure Database and 3D-
425 Beacons: New Data and Capabilities. Journal of Molecular Biology 437 (15):168967.
426 doi:https://doi.org/10.1016/j.jmb.2025.168967
427 33. Mukai K, Kawata-Mukai M, Tanaka T (1992) Stabilization of phosphorylated Bacillus subtilis DegU by
428 DegR. J Bacteriol 174 (24):7954-7962. doi:10.1128/jb.174.24.7954-7962.1992
429 34. Kobayashi K (2007) Gradual activation of the response regulator DegU controls serial expression of
430 genes for flagellum formation and biofilm formation in Bacillus subtilis. Molecular Microbiology 66
431 (2):395-409. doi:https://doi.org/10.1111/j.1365-2958.2007.05923.x
432 35. Knudsen GM, Olsen JE, Dons L (2004) Characterization of DegU, a response regulator in Listeria
433 monocytogenes, involved in regulation of motility and contributes to virulence. FEMS Microbiology
434 Letters 240 (2):171-179. doi:10.1016/j.femsle.2004.09.039
435 36. Xi Y, Sohn AL, Joignant AN, Cologna SM, Prentice BM, Muddiman DC (2023) SMART: A data reporting
436 standard for mass spectrometry imaging. Journal of Mass Spectrometry 58 (2):e4904.
437 doi:https://doi.org/10.1002/jms.4904

438


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