



OPEN

## Ionic liquid-assisted sample preparation mediates sensitive proteomic analysis of *Bacillus subtilis* spores

Masato Taoka<sup>1,5✉</sup>, Ritsuko Kuwana<sup>2,5</sup>, Tatsumi Fukube<sup>3</sup>, Akiko Kashima<sup>4</sup>, Yuko Nobe<sup>1</sup>, Takamasa Uekita<sup>3</sup>, Tohru Ichimura<sup>3✉</sup> & Hiromu Takamatsu<sup>2</sup>

Endospore-forming bacteria are ubiquitous. Bacterial endospores are multilayered proteinaceous structures that protect the bacterial genome during stress conditions. They are also responsible for a wide range of critical clinical infections in humans. Precise analysis of spore-forming pathogens remains a major challenge in the field of proteomics because spore structures are highly resistant to conventional solubilizers and denaturing agents, such as sodium dodecyl sulfate and urea. We present an ionic liquid-assisted (*i*-soln) technique of sample preparation, called pTRUST, which enables shotgun analysis of *Bacillus subtilis* spores even when the starting materials are in the sub-microgram range. In proteomic analysis, this technique shows 50–2000-fold higher sensitivity than other conventional gel-based or gel-free methods (including one-pot sample processing). Using this technique, we identified 445 proteins with high confidence from trace amounts of highly pure spore preparations, including 52 of the 79 proteins (approximately 70%) previously demonstrated to be localized in spores in the SubtiWiki database and detected through direct protein analysis. Consequently, 393 additional proteins were identified as candidates for spore constitutive proteins. Twenty of these newly identified candidates were produced as green fluorescent protein fusion proteins, and each was evaluated for authenticity as a spore constituent using fluorescence microscopy analysis. The pTRUST method's sensitivity and reliability using the *i*-soln system, together with hitherto unreported proteins in spores, will enable an array of spore research for biological and clinical applications.

**Keywords** pTRUST, Ionic liquid, *Bacillus subtilis*, Endospore, Proteomics

Spores are small round or oval body that are produced by certain bacteria such as Gram-positive *Bacillus* and *Clostridium* species in response to cytoplasmic dehydration under stress conditions. They are typically dormant and show no metabolic activity but can respond to nutrients in the environment and quickly become active vegetative cells through a process named germination<sup>1</sup>. Several spore-producing bacteria are pathogenic including *Bacillus* (aerobic) and *Clostridium* (anaerobic) species. Among these, *B. anthracis* and *C. tetani* cause clinical infections, whereas *B. cereus* and *B. weihenstephanensis* cause food poisoning. Furthermore, *B. anthracis* and *C. botulinum*, because of their ability to produce lethal toxins and their potential for mass dissemination, must be effectively controlled in terms of bioterrorism countermeasures. Therefore, developing rapid and sensitive detection (or identification) techniques for spore samples is critical for defending against these pathogens and diseases.

Spores are primarily composed of proteins, along with some DNA and small molecules. These proteins are crucial for the formation, resistance, and germination of spores, as well as for the pathogenicity of spore-forming bacteria. Various proteomic approaches have been developed to characterize these proteins, using the non-pathogenic *B. subtilis* as a model<sup>2–6</sup>. For example, in a gel-based method, proteins isolated from the spore are solubilized in sodium dodecyl sulfate (SDS), separated by SDS-polyacrylamide gel electrophoresis (PAGE) or two-dimensional (2D) separation<sup>2–5</sup>. The protein of interest is then excised, digested in-gel, and analyzed using mass spectrometry (MS). However, from relatively large amounts of purified spore preparations, only 20–30 µg

<sup>1</sup>Department of Chemistry, Tokyo Metropolitan University, Tokyo 192-0397, Japan. <sup>2</sup>Faculty of Pharmaceutical Science, Setsunan University, Osaka 573-0101, Japan. <sup>3</sup>Department of Applied Chemistry, National Defense Academy, Yokosuka, Kanagawa 239-8686, Japan. <sup>4</sup>Carriere Reseau Co., Ltd., Kanagawa 238-0011, Japan. <sup>5</sup>These authors contributed equally: Masato Taoka and Ritsuko Kuwana ✉email: mango@tmu.ac.jp; ichimura@nda.ac.jp

of protein is often required for SDS-PAGE<sup>2,3</sup> and 600–800 µg for 2D-PAGE<sup>4,5</sup> for effective MS analysis, which is labor-intensive and time-consuming. Furthermore, SDS can efficiently solubilize only the proteins present on the coat region or certain portions of the spores<sup>7</sup>, indicating that key proteins and peptides can be missed during extraction and subsequent MS analysis. Recently, a one-pot sample processing proteomics approach was developed for MS analysis, which involves direct digestion of spore proteins with proteases without SDS and gels<sup>6</sup>. In this method, spores are crushed by bead-beating in the presence of 6 M urea, and the protein suspension is directly subjected to double digestion with the lysylendopeptidase, LysC, and trypsin in a single tube. The digested samples are separated into 10 fractions using hydrophobic chromatography, and the resulting fractions are analyzed using liquid chromatography-tandem mass spectrometry (LC-MS/MS). This one-pot method is more efficient in protein identification than conventional gel-based methods and enables the analysis of membrane proteins that are difficult to solubilize with SDS<sup>6</sup>. However, this approach also requires a large amounts of sample (630 µg) as in the typical 2D system described above, suggesting that it is susceptible to the same analytical problems. Furthermore, large-scale preparation may increase the risk of contamination and spread of vegetative cells. Thus, alternative approaches that can be adapted to rare and quantity-limiting spore samples are required for sensitive spore proteomics.

Ionic liquids (ILs), composed of organic cations and weakly coordinating anions, have recently emerged as solvent media for application in many research areas, including biotechnology and the pharmaceutical industry<sup>8</sup>. The solubilization capacity of ILs has been exploited to dissolve highly insoluble biomolecules, such as cellulose<sup>9</sup>, thereby providing a versatile platform for processing and utilizing insoluble materials. We previously demonstrated that a mixture of imidazolium-based IL, 1-butyl-3-methylimidazolium cyanate [bmim][SCN] and NaOH (*i*-soln), completely dissolved highly insoluble aggregates such as heat-aggregated hen egg whites within 10 min<sup>10</sup>. The dissolution efficacy of this solution significantly surpassed those of conventional detergents and chaotropes<sup>10</sup>. We also developed two proteomic methods (named as *i*BOPs<sup>10</sup> and its improved version, pTRUST<sup>11</sup>) using hydrophobic microbead supports to process *i*-soln-solubilized proteins, using trypsin. The analytical properties of these methods have allowed for ultrasensitive proteomic characterization of a variety of insoluble proteins, including inclusion bodies of human TRIM32 ubiquitin ligase<sup>12</sup>, SDS-resistant aggregates deposited in individual senescent *Caenorhabditis elegans* nematodes<sup>10</sup>, and integral membrane proteins prepared from various human cancer cell lines<sup>11</sup>. Proteomic analysis of spore samples remains a major challenge because of their low solubility in conventional detergents and chaotropes; thus, we evaluated the potential of *i*-soln and pTRUST to analyze the spore proteome of *B. subtilis* as a model. The pTRUST technology with the *i*-soln system can support shotgun analysis of *B. subtilis* spores, even in the sub-microgram range.

## Materials and methods

### Strains

*B. subtilis* and *Escherichia coli* strains and plasmids used in this study are listed in Supplementary Tables S1.

### Chemicals and enzymes

[Bmim][SCN] was purchased from Sigma Aldrich (St. Louis, MO, USA). *i*-soln was prepared by mixing [bmim][SCN] and 0.5 M NaOH (in water) in a 40:60 (v/v) ratio<sup>10</sup>. POROS R2 microbeads (diameter, 50 µm) were obtained from PerSeptive Biosystems, Inc. (Framingham, MA, USA). The beads (250 mg) were suspended in 50% methanol at a concentration of 50 mg/mL and stored at room temperature until use. Before use, 10 µL of R2 suspension (2 µL bead volume, 500 µg) was rinsed with 100 µL of 75% acetonitrile (CH<sub>3</sub>CN) in 0.1% trifluoroacetic acid and 100 µL of 100 mM Tris-HCl (pH 8; hereafter, Tris buffer), and resuspended in 200 µL water. StageTips (polystyrene-divinylbenzene copolymer) were obtained from Nikkyo Technos Co., Ltd. (Bunkyo-ku, Tokyo, Japan). Trypsin (platinum-grade) was obtained from Promega (Madison, WI, USA). SDS was acquired from Sigma-Aldrich. Tris (2-carboxyethyl) phosphine (TCEP) and iodoacetamide were purchased from Hampton Research (Aliso Viejo, CA, USA) and Nacalai Tesque (Nakagyo-ku, Kyoto, Japan), respectively. Other materials were purchased as previously described<sup>13</sup>.

### Spore preparation

The *B. subtilis* 168 strains were grown in Schaeffer's medium at 37 °C as previously described<sup>2</sup>. Spores were harvested at 18 h after the cessation of exponential growth, washed once with 10 mM Tris-HCl (pH 7.2), and prepared using a previously described procedure with some modifications<sup>2</sup>. Spores were suspended in deionized water and washed using centrifugation at 12,000×*g* for 4 min at 4 °C until all cell debris and vegetative cells were removed, as verified under a phase-contrast microscope<sup>14</sup>. The pellets were suspended in 0.1 mL lysozyme buffer (10 mM Tris-HCl, pH 7.2, with 1% [w/v] lysozyme) and incubated at 37 °C for 10 min to completely remove cell debris and vegetative cells. The pellets were repeatedly washed with 10 mM Tris-HCl (pH 7.2) and 0.5 M NaCl at 25 °C. This treatment resulted in over 99% refractive spores, and almost no dark or gray spores were visible under a phase-contrast microscope. The samples did not contain vegetative cells or debris. The pellets were resuspended in 10 mM Tris-HCl (pH 7.2) and stored at –80 °C until use.

### pTRUST procedure

Purified *B. subtilis* spores were processed using pTRUST as previously described<sup>11</sup>. Briefly, the spores (2–3 × 10<sup>6</sup> cfp) containing 0.4–0.5 µg protein, were incubated for 10 min in 50 µL *i*-soln ([bmim][SCN]–0.5 M NaOH solution [40% (v/v) IL])<sup>10</sup> containing 20 mM TCEP using three cycles of ultrasonication (2 min) and agitation (1 min) in a water bath sonicator (ASU-10D; AS ONE Corporation, Osaka, Japan) and vortex mixer, respectively. The reduced samples were alkylated with 40 mM iodoacetamide in the dark for 20 min. The rinsed R2 beads (2 µL of bead volume) were directly added to the sample. The bead-protein mixture was agitated four times with

a vortex mixer for 1 min and left to stand for 1 min to allow the protein to concentrate on the beads. For washing and trypsin digestion, the bead-protein mixture was pipetted into a StageTip container and centrifuged at  $2000 \times g$  at  $20^{\circ}\text{C}$  for 30 s to remove any excess *i*-soln from the beads that were retained on the StageTip filter. All subsequent centrifugation steps were performed under the same conditions. The retained beads were rinsed with 100  $\mu\text{L}$  Tris buffer (100 mM Tris-HCl, pH 8.0) using centrifugation, and washed twice with 100  $\mu\text{L}$  acetone via centrifugation to remove impurities in the commercial [bmim] [SCN] product that could bind to the R2 surface<sup>11</sup>. The beads were rinsed with 100  $\mu\text{L}$  Tris buffer and 100  $\mu\text{L}$  water. Trypsin digestion was conducted with 0.5  $\mu\text{g}$  trypsin in 20  $\mu\text{L}$  *i*BOPs digestion buffer (5 mM Tris, 60%  $\text{CH}_3\text{CN}$ , pH 8.8)<sup>10,11</sup> at  $37^{\circ}\text{C}$  in a sealed StageTip container with rotation, overnight.

### LC-MS/MS analysis and database searching

After trypsin digestion, the peptides were recovered from the beads using centrifugation and one additional wash with 20  $\mu\text{L}$  of 60%  $\text{CH}_3\text{CN}$  in 0.1% trifluoroacetic acid. The flowthroughs were combined and diluted tenfold with 0.1% formic acid. The diluted samples were directly subjected to LC-MS/MS analysis using a quadrupole-Orbitrap hybrid mass spectrometer (Q Exactive; Thermo Fisher Scientific, Waltham, MA, USA) as previously described<sup>10,11</sup>. Tandem MS data were converted into a Mascot-compatible format using Proteome Discoverer (version 3.0; Thermo Fisher Scientific). A database search was performed using MASCOT v. 2.3.02 (Matrix Science K.K., Tokyo, Japan) against the UniProt *B. subtilis* proteome database. The search parameters were set as follows: fixed modification for carbamidomethyl (C), variable modifications for acetylation (protein N-terminus) and oxidation (Met), maximum missed cleavage at 1, peptide mass tolerance of  $\pm 25$  ppm, and MS/MS tolerance of  $\pm 0.8$  Da. The peptide identification criteria were based on vendor definitions ( $P < 0.05$ ; Matrix Science K.K.), as used in our previous works<sup>10–13</sup>. Identification was manually confirmed through inspection of the MS/MS spectra.

### Turbidity measurement

For water and *i*-soln, purified spores ( $2–3 \times 10^8$  cfp) were incubated for 10 min at room temperature in 1 mL of water or *i*-soln using three cycles of ultrasonication (2 min) and agitation (1 min) as described above. For SDS, the same number of spores was placed in 1 mL of 1% SDS and boiled in a water bath for 3 min. The turbidity (optical density [OD] at 600 nm) of these solutions was measured using a UV-vs spectrophotometer (SmartSpec™ Plis; Bio-Rad, Hercules, CA, USA) at 0, 1, and 24 h.

### Spore counting and protein quantitation

The *B. subtilis* spore samples ( $\sim 7 \times 10^7$  cfp) were diluted with distilled water, and an appropriate amount of diluted solution was spread on Luria-Broth agar plates and incubated at  $37^{\circ}\text{C}$  overnight. The number of colonies produced on the plates was counted and regarded as the spore number. The protein concentration of the spores lysed in *i*-soln was measured using the Bradford assay (Novexin, Braham, Cambridge, UK) with bovine serum albumin as a standard<sup>11</sup>.

### Predicting protein localization sites within spore regions

Sporulation-related genes are transcribed during sporulation by an RNA polymerase containing developmentally specific sigma factors, SigF, SigE, SigG, and SigK. These sigma factors are temporally and spatially activated and regulate gene expression in a compartment-specific manner<sup>15,16</sup>. Proteins synthesized in the mother cell compartment under the regulation of SigE and SigK are localized in the cortex or coat. In contrast, proteins synthesized in the forespore under the regulation of SigF and SigG are localized in the cortex and core<sup>2</sup>. Additionally, signal sequences or transmembrane regions may facilitate the translocation of proteins to the outer and inner spore membranes<sup>17</sup>. Proteins whose expression is regulated by SigE or SigK that have transmembrane or signal sequences are localized in the outer spore membrane. Proteins whose expression is regulated by SigF or SigG that have transmembrane or signal sequences are localized to the inner spore membrane.

### Constructing green fluorescent protein fusions

Green fluorescent protein (GFP)-fused strains were constructed using overlap extension polymerase chain reaction (PCR)<sup>18</sup>. A *gfp*-*cat* (*chloramphenicol acetyltransferase*) cassette was amplified from the plasmid pCBGFPUV4<sup>18</sup> by PCR using UV4-F and cat-R primers (Supplementary Table S2). To construct strains G-0507, G-0637, G-1230, and G-1533, a *gfp*-*cat* cassette was amplified from the plasmid pCBGFPUV4 by PCR using UV4-F and cat-T-R primers (Supplementary Table S2). Cat-T-R primers contained transcription terminators to prevent read-through. The coding region of the target gene and its downstream region were amplified by PCR using the primer pairs F1/R1 and F2/R2 (Supplementary Table S2). The 5' sequences of R1 and F2 were complementary to those of UV4-F and cat-R, respectively. The three PCR fragments were fused by mixing and used as templates for the second round of PCR with the primers F1 and R2. The resulting PCR products were transformed into strain 168 to generate the *gfp*-fused *cat* strain. GFP fusion was achieved using double-crossover recombination. Transformants were selected on Luria-Broth agar plates containing 5  $\mu\text{g}/\text{mL}$  chloramphenicol.

### Phase contrast and fluorescence microscopy

*B. subtilis* strains harboring GFP fusion-encoding genes on the chromosome were grown in Schaeffer's sporulation medium, and aliquots were transferred to microscope slides. Phase-contrast images of cells and GFP fluorescence were observed using an Olympus BX51 phase-contrast microscope with additional fluorescence attachments and mirror cube units for GFP (U-MGFPHQ; Olympus, Tokyo, Japan). A UPlanApo 100  $\times$  oil Iris

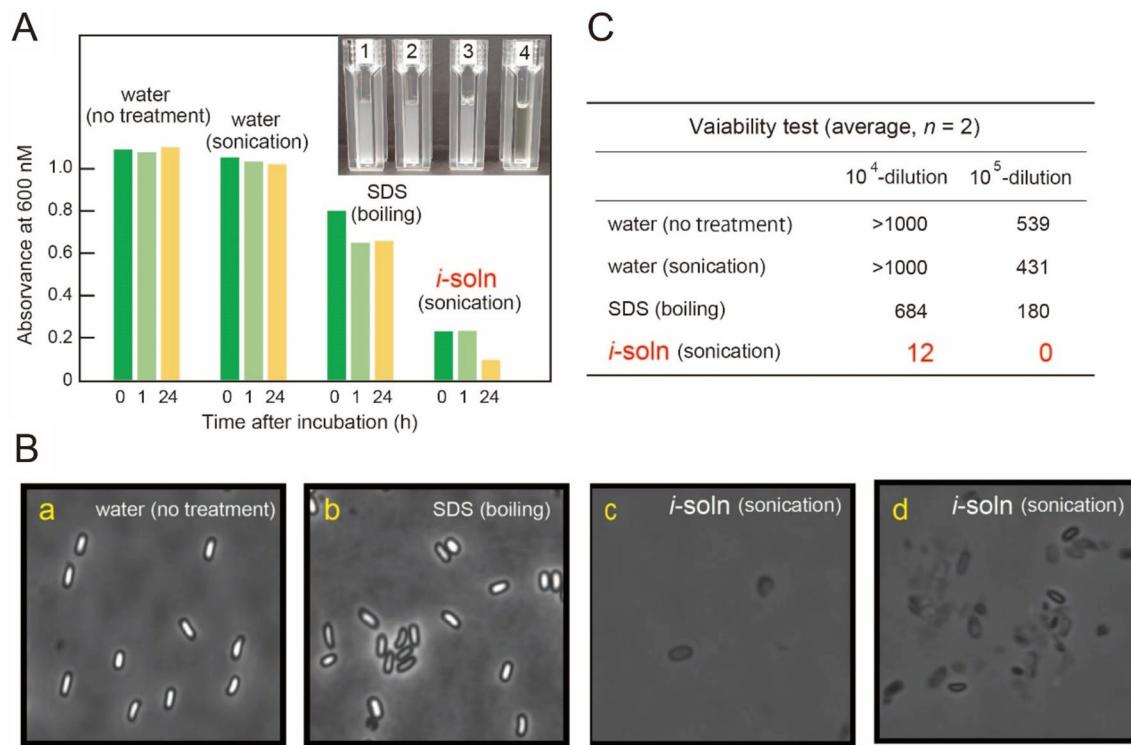
Ph3 objective lens and U-TV0.5XC-3-8 camera adapter were used (Olympus). Images were captured using a backside-illuminated high-sensitivity monochrome CMOS sensor camera DP23M (Olympus) and analyzed using CellSens imaging software (Olympus). The exposure time for image capture of the GFP-fused protein was 0.4–4.0 s. The contrast and tone balance were adjusted using CellSens imaging software (Olympus). For each strain, the images were scaled to the same intensity range.

## Results

### Dissolution of *B. subtilis* spores with *i-soln*

Dormant spores are resistant to a variety of extracellular assaults, including heat, toxic chemicals, lytic enzymes, and other cellular damaging factors. To examine whether *i-soln* (40% [bmim][SCN], 60% 0.5 M NaOH)<sup>10</sup> can disrupt the spore structure, highly purified *B. subtilis* spores (>99% purity, as assessed using phase contrast microscopy) were incubated in *i-soln* for 10 min using sonication and agitation, and the turbidity ( $OD_{600\text{ nm}}$ ) of the resulting solution was measured with a spectrophotometer after 0, 1, and 24 h. As controls for this experiment, the same spore samples were suspended in water (with no treatment) and sonicated in water as described above or boiled in 1% SDS at 100 °C for 3 min. *i-soln* showed the lowest turbidity values at all time points (Fig. 1A). Particularly, the value in *i-soln* was approximately 20% of the no treatment control value after 0 h, whereas the value with 1% SDS was over 60% of the no treatment control value, even after 24 h (Fig. 1A, see also Fig. 1A, inset; compare #3 and #4). The poor disruption efficiency observed after boiling in 1% SDS did not improve following sonication in 1% SDS, 2% SDS containing 5%  $\beta$ -mercaptoethanol, or 5% SDS containing 2.5% TCEP (Supplementary Figure S1).

The spores suspended in water or boiled in 1% SDS had a bright oval morphology typical of intact spores according to phase-contrast microscopy (Fig. 1B [a,b]), although the number of spores in 1% SDS appeared to be slightly lower than that in water. In contrast, no bright spores were detected in the *i-soln* lysate (Fig. 1B [c]). Occasionally, darkened, smaller, and clustered residues were detected in certain fields under the microscope (Fig. 1B [d]). These residual structures may represent damaged spores or debris. The cell viability assay results support this assumption (Fig. 1C); many viable cells were generated from the suspensions in water or 1% SDS, but few were observed from the *i-soln* lysate. Thus, *i-soln* is a superior solubilizer for spore proteomics compared



**Figure 1.** Dissolution of *Bacillus subtilis* spores with *i-soln*. (A) Purified spores ( $2-3 \times 10^8$  cfp) were incubated in water and *i-soln* with or without sonication or boiled in 1% SDS as indicated, and the turbidity ( $OD_{600\text{ nm}}$ ) of the solutions was measured with a UV-vis spectrophotometer at 0, 1, and 24 h. The absorbance values at the indicated time points represent the average of each duplicate experiment. The inset shows the solution images after 24 h: 1. Water (no treatment), 2. water (sonication), 3. SDS (boiling), 4. *i-soln* (sonication). (B) Each solution of (A) was observed under a phase contrast microscope at 1000 $\times$  magnification after 24 h. The scale bar represents 5  $\mu\text{m}$ . (C) Purified spores ( $\sim 7 \times 10^7$  cfp) were diluted with distilled water as indicated, and the diluted solutions were spread on Luria-Broth agar plates and incubated at 37 °C overnight. This experiment was performed twice, and the number of colonies on the plate was counted and expressed as an average number.

with SDS. This conclusion is consistent with those previously reported for heat-denatured hen egg whites and SDS-resistant aggregates deposited in aged nematodes<sup>10</sup>.

### Analysis of spore proteins using pTRUST and LC-MS/MS

To identify proteins,  $2-3 \times 10^6$  cfu of the purified spores (corresponding to 0.4 µg proteins) were lysed in *i*-soln, and the solubilized proteins were S-carbamoylmethylated to cleave intra- and inter-molecular disulfide bridges. The alkylated sample was digested with trypsin according to the procedure described for pTRUST and directly analyzed using one-dimensional LC-MS/MS. We repeated this analysis four times with four independent alkylated samples and obtained 54,903 MS/MS spectra that were assigned to 5,575 peptides, leading to the identification of 445 unique proteins (average of 3.1 corresponding peptides per protein per single assay). In these experiments, approximately 260–300 proteins were consistently identified in each MS run, with good overlap in the analysis, confirming reproducibility even at the low quantity level. The complete datasets from these experiments, together with Mascot's E-values for all identifications, are available in Supplementary Table S3.

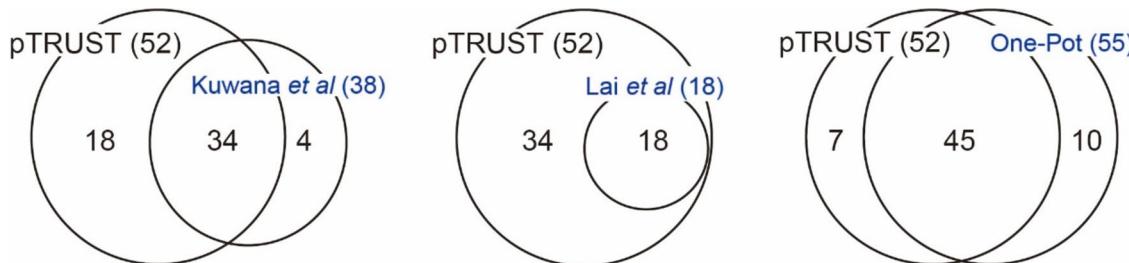
The 445 identified proteins were searched against the SubtiWiki and UniProt *B. subtilis* databases for their molecular mass, isoelectric point, GRAVY (grand average of hydropathicity) value, regulon, function, and subcellular localization. This retrieval revealed a wide range of biochemical and biological properties for the proteins (Supplementary Table S4). The identified proteins also included 52 (70%) of the 79 proteins previously annotated as "sporulation proteins" in the SubtiWiki database and found in direct protein analysis (marked in yellow in the Supplementary Table S4). Accordingly, 393 other proteins were identified as candidates for new spore-constituting proteins, reflecting that the pTRUST method can be used for sensitive global proteomic screening to identify novel components.

### Analytical performance of the pTRUST method is superior to that of conventional approaches

Various approaches such as gel-based and gel-free methods have been used to analyze the *B. subtilis* spore proteome<sup>2–6</sup>. These methods include 1D-PAGE<sup>2</sup> or 2D-PAGE<sup>4</sup> and the gel-free one-pot approach<sup>6</sup>. However, we could not directly compare these approaches with pTRUST using the respective datasets because of the several critical differences in sample processing, MS instrumentation, sample amount requirements, and sample purity. Thus, we used the 79 sporulation proteins mentioned above as queries to search previous datasets. The number and types of proteins included in the previous analyses were not significantly different from those identified using pTRUST, although the pTRUST data fully covered the 2D-PAGE data (Fig. 2, Supplementary Table S5). However, the clear advantage of pTRUST over conventional techniques is that the sample amount (0.4 µg) was 50–2,000-fold less than that used in previous approaches (20–800 µg<sup>2–6</sup>). This result indicates that pTRUST significantly improves the efficiency of spore sample preparation for high-resolution MS analysis compared to that of conventional techniques. Furthermore, pTRUST does not require additional sample purification steps, such as PAGE or hydrophobic chromatography necessary in traditional methods; these steps are time-consuming and result in protein loss. Thus, pTRUST using the *i*-soln system provide a simple and direct platform for highly sensitive proteome analysis of *B. subtilis* spores compared to previous approaches.

### Predicting localization sites of identified proteins in the spore layer

Although the *B. subtilis* gene sequencing project has identified approximately 4,100 protein-coding genes, it is largely unknown which proteins are located in the spores, particularly in the spore layer. We previously proposed a mechanism by which the localization of a protein in spore structures can be predicted based on information regarding transcriptional regulation by sporulation-specific sigma factors (SigF, SigE, SigG, and SigK) and the presence or absence of signal sequences or transmembrane regions in the amino acid sequences (2, 15–17; see "Materials and methods"). On the basis of this prediction, a total of 98 proteins were selectively extracted from the complete dataset presented here, which were indeed downstream of the sigma factors mentioned above, and approximately 20% of these proteins shared signal sequences, membrane anchor sequences, or transmembrane segments (Supplementary Table S6). The extracted 98 proteins included most of the above-mentioned 52 authentic sporulation proteins (49 proteins) and other 49 proteins newly identified in this study as candidates for spore constitutive proteins. Furthermore, this prediction suggested that the proteins identified by pTRUST may be components of all spore-forming layers, including the core, inner membrane, cortex, outer membrane, and spore coat (Supplementary Table S6). To further characterize the identified proteins, we examined the spore



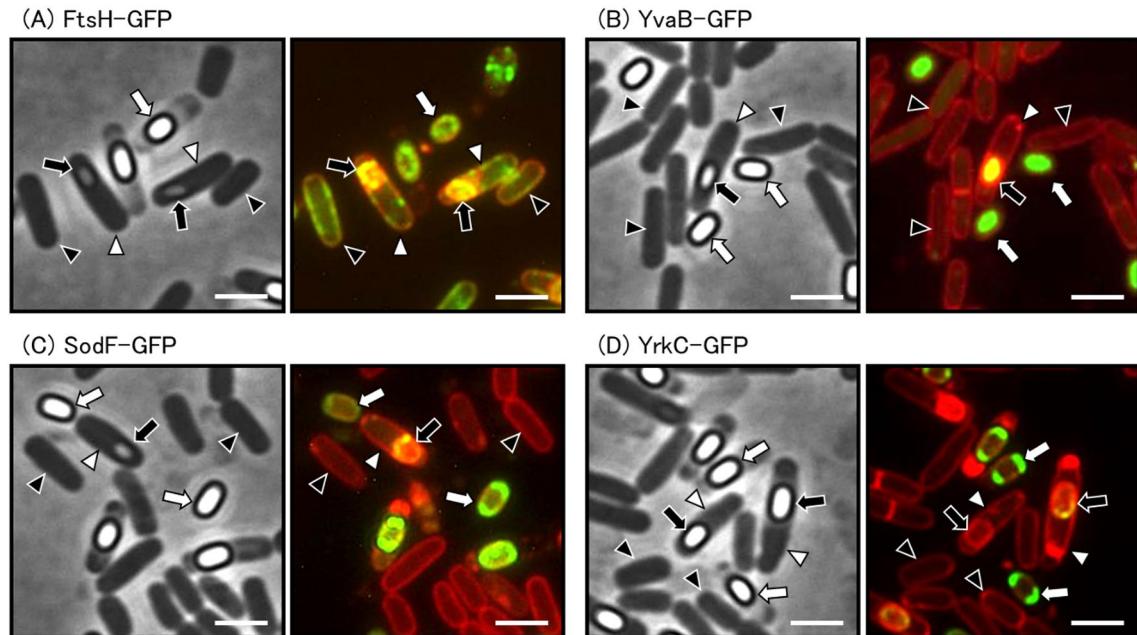
**Figure 2.** Comparing *Bacillus subtilis* spore proteins identified using pTRUST and conventional approaches. Numbers in parentheses indicate the number of proteins identified using each method. Kuwana et al.<sup>2</sup> represent the 1D-PAGE method, Lai et al.<sup>4</sup> represent the 1D- and 2D-PAGE, and Swarge et al.<sup>6</sup> represent the one-pot approach.

localization of newly identified candidates listed in Supplementary Table S6 as well as the remaining candidates not included in this list that are not known to be expressed under the control of these sigma factors (see below).

### Demonstrating location of identified proteins in the spores using GFP fusions

To characterize the identification process described above, we initially used four GFP fusions generated with the newly identified candidates listed in Supplementary Table S6 (FtsH-GFP, YvaB-GFP, SodF-GFP, and YrkC-GFP). All four expressed fusion proteins were localized to the spores (Fig. 3, Supplementary Table S7). The observed localization in the spore region perfectly matched the deduced distribution of the corresponding proteins. FtsH is a membrane protein produced in a SigA-, SigF-, and SigM-dependent manner, and FtsH-GFP fluorescence was detected in the spore membrane as well as in the cytoplasmic membrane of vegetative cells and mother cells (Fig. 3A). YvaB is produced in a SigA- and SigG-dependent manner. YvaB-GFP fluorescence was detected in the spore core and weakly in the cytoplasm of vegetative cells (Fig. 3B). This result suggests that the expression level of the *yvaB* gene dependent on SigA was low in vegetative cells. SodF and YrkC are proteins produced within the mother cells and are dependent on SigE and SigK, respectively. As SodF- and YrkC-GFP fluorescence were detected in the outer layer of the spore, these proteins were thought to be localized in the spore coat (Fig. 3C, D). Subsequently, we produced another GFP fusion with a total of 16 different candidate proteins with three listed in Supplementary Table S6 (PhoA, YfkD, and YqfQ) and 13 candidates (Alr, Hbs, PdhB, PdhD, Pfk, Pfk, Tsf, TufA, YloB, YojB, YpfD, YtcI, and YuiE) unlisted in Supplementary Table S6. All these fusion proteins were localized in the spores (in the core, cortex, or membrane) (Supplementary Figure S2, Supplementary Table S7). Thus, 20 of the newly identified candidate proteins examined here (i.e., 4 and 16 shown in Fig. 3 and Supplementary Figure S2, respectively) were considered to represent genuine spore components.

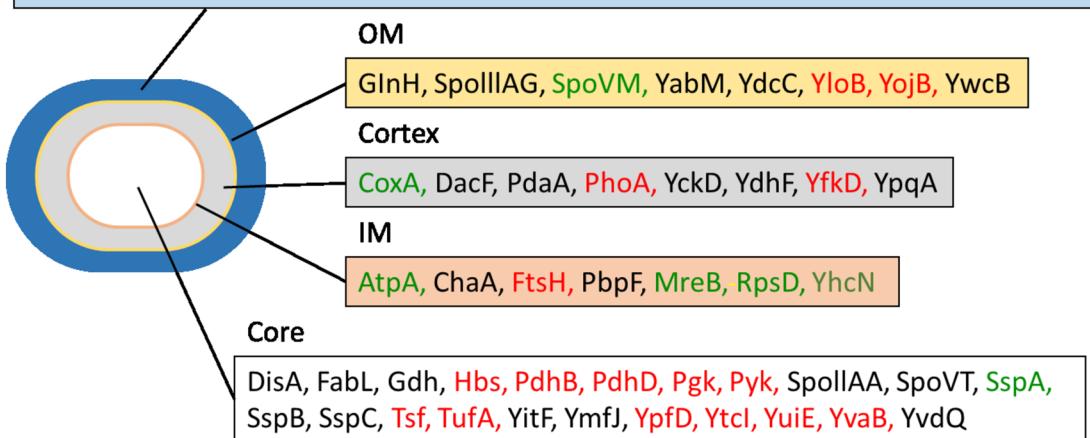
To further assess the impact and reliability of the present identification protocol, we surveyed previous literature on case-by-case data on GFP fusion proteins. This survey identified 63 proteins that were included in our complete dataset (Supplementary Table S4) and demonstrated that 41 (or 65%) were indeed localized within the spores (Supplementary Table S8, marked in red), supporting our identification. These 41 proteins included, in addition to 31 of the authentic 52 sporulation proteins mentioned above, 10 novel proteins identified here as candidates for spore components (AtpA, LonA, MreB, RpsD, SleL, SspA, YncN, YodI, YusA, and YvdP). None of these candidates overlapped with the 20 proteins characterized above with the GFP assay. In contrast, the remaining 22 of the 63 proteins identified in this survey were localized to vegetative cells or mother cells (Supplementary Table S8). However, the distribution of these 22 proteins reported in previous literature has been investigated primarily using vegetative cells or mother cells, rather than spores, and should be reexamined directly using purified spore samples. The spore locations of the identified proteins revealed by these analyses are summarized in Fig. 4. To the best of our knowledge, this study represents the first report in the field of spore proteomics describing the use of a cell imaging technique to systematically confirm proteomic data reliability. Although further examination of these and other samples is needed, these findings validate our identification



**Figure 3.** Verifying the localization of identified proteins within spores using green fluorescent protein (GFP)-fusion constructs. Four GFP fusion constructs (FtsH-GFP, YvaB-GFP, SodF-GFP, and YrkC-GFP) were generated, and their localization in spores was observed under a phase contrast microscope with fluorescence attachments. White and black arrows indicate mature spores and forespores, respectively. White and black arrowheads indicate mother and vegetative cells, respectively. Scale bar, 2  $\mu$ m.

## Coat

Alr, CgeA, CotA, CotB, CotC, CotD, CotE, CotF, CotG, CotH, CotI, CotJA, CotJB, CotJC, CotO, CotQ, CotR, CotS, CotSA, CotU, CotW, CotY, CotZ, Cse15 (YknT), CwlJ, DapA, DapG, GerPC, GerQ, GerT, GlgP, KamA, LipC, MmgD, OxdD, PrkA, RmlA, SafA, SleL, SodF, SpolVA, SpoVID, SpoVIF, SpsA, SpsE, Tgl, YaaH, YabG, YdhD, YhbB, YheC, YhjR, YhxC, YisY, YisZ, YjdH, YjqC, YjzB, YkuD, YkvP, YkzQ, YnzH, YodI, YqfC, YqfQ, YrkC, YtcB, YtcC, YtxO, YurS, YusA, YuzC, YvdP, YxeE, Yybl



**Figure 4.** Schematic of *Bacillus subtilis* spores and estimated distribution of identified proteins. Proteins marked in black indicate those deduced from the sigma factors and transmembrane/anchor sequences (taken from Supplementary Table S6). Proteins marked in red indicate those whose localization was demonstrated in this study (taken from Supplementary Table S7). Proteins marked in green indicate those whose localization has already been localized with GFP assay (taken from Supplementary Table S8). OM, outer membrane. IM, inner membrane.

and suggest that the proteins identified using pTRUST and LC–MS/MS may become valuable resources for future studies on the spore and spore layer proteome.

## Discussion

We used pTRUST and MS to identify 445 proteins in trace amounts of *B. subtilis* spore samples. The localization of 52 of these 445 proteins in spores is listed in the SubtiWiki database (Supplementary Table S5). Localization of the remaining 393 proteins in the spore was not fully confirmed in the database. The identified proteins included key proteins and enzymes belonging to all functional categories in the database, including a series of Cot proteins in the spore coat architecture and ribosomal subunits in protein synthesis. However, several important spore proteins were not identified in our analysis, which may be attributed to several factors: (i) the protein content was below the detection limit; (ii) the proteins were very small (< 10 kDa) and did not produce tryptic fragments of suitable length; and (iii) the proteins negligibly solubilized in *i*-soln and were easily precipitated or aggregated during the extraction step, preventing trypsin digestion. We recently found that several known mediators or regulators of sporulation and germination localized in the inner membrane (such as SpoVAD, SleB, YpeB, and GerD) can be identified using this method only when the spore germination agent (AGFK) is added to the spore preparation (37 °C for 3 h, unpublished data). Therefore, alternative conditions may be required to identify additional spore proteins.

It is known that one peptide hits are less robust and have a higher chance of being false-positives. However, we did not remove the identification with one-hit peptides from our dataset because the identification criteria ( $p < 0.05$ ) employed in this study were commonly used and were practically proven by our previous works<sup>10–13</sup>. In fact, at least 19 proteins identified with a single peptide in this study (DapG, ChaA, CotD, CotO, YisZ, FabI, FtsH, GerT, PbpF, SpoVT, SpsA, YabM, YckD, YdhF, YkuD, YrkC, YtcB, YvdQ, YwcB) were considered to be true spore components as revealed by another approach, as shown in Fig. 4.

Recently, a new proteomic method (called SPEED) using trifluoroacetic acid (TFA) as a solubilizing agent has been published<sup>19,20</sup>. We assume that in addition to those of the reported bacterial species (e.g., *E. coli*, *S. aureus*, and *B. cereus*), TFA could be effective in lysing *B. subtilis* spores as well. However, the described SPEED protocols typically require TFA-solubilized sample neutralization and dilution prior to trypsin digestion, resulting in a tenfold or greater sample volume scale-up (> 100  $\mu$ L). This may reduce the generation of tryptic peptides when applied to small sample quantities (for example, less than 1  $\mu$ g), as in-solution digestion becomes ineffective with decreasing protein concentration (for example, at submicromolar protein concentration, Ref<sup>13</sup> and references therein). From this perspective, we note that the current pTRUST protocol displays several benefits even in processing low-abundance (i.e., low-concentration) samples. For example, the inclusion of 60% acetonitrile in the trypsin digestion buffer (termed *i*BOPs buffer<sup>11</sup>) stimulates the catalytic activity of trypsin and tryptic

peptide solubility during the digestion reaction and reduces the potential for the nonspecific adsorptive loss of polypeptides to the walls. Furthermore, the small-sized container of the StageTip pipette used for trypsin digestion facilitates small-scale enzymatic digestion (< 20  $\mu$ L), thereby reducing the surface area involved in the proteolytic peptide adsorption loss<sup>11</sup>. Therefore, we believe that the pTRUST method employing the *i*-soln system is one of the most efficient and reliable platforms in high-sensitivity spore proteome analysis.

Proteomic techniques are widely used in spore biology to analyze the mechanisms of stress tolerance, sporulation, germination, and elongation of spores at the molecular level and to identify biological markers that support efficient clinical settings<sup>21</sup>. Previous techniques of analysis used relatively large amounts of spore preparations, which may pose various complications when applied to rare and quantity-limited samples, including those of bacteria with extremely low sporulation rates<sup>22</sup>. Such large quantity of preparations may reduce the purity of spore samples for subsequent MS analysis, thereby reducing the reliability of the MS data, and may increase the risk of pathogenic transmission. The sample preparation efficiency of the pTRUST method involving the *i*-soln system is superior to that of traditional approaches; therefore, this technique may be useful for more sensitive and important studies of spore biology, and in the pharmaceutical and food industries for detecting and controlling spores to prevent contamination and infection by pathogenic bacteria. The application of this method to spores of various bacterial strains other than *B. subtilis* is currently being evaluated in our laboratory.

## Conclusions

The pTRUST method involving the *i*-soln system is simple and allows for high-efficiency shotgun analysis of the *B. subtilis* spore proteome in trace amounts of samples. Excluding trypsin digestion, this technique requires less than 45 min<sup>11</sup>, which is shorter than the conventional gel-based or gel-free approaches that typically take several hours. The identification accuracy of pTRUST was also evaluated using 20 of the newly identified candidates, all of which were confirmed to be localized to spores using fluorescent microscopic analysis. Furthermore, this technology may be applicable to sensitive spore-protein dynamic analysis without substantial modification when used in combination with conventional quantitative techniques, such as stable isotope labeling<sup>23</sup> and label-free methods<sup>24</sup>. Therefore, pTRUST has enormous potential to facilitate a wide spectrum of spore research for biological and therapeutic applications.

## Data availability

The proteomics data have been deposited to the ProteomeXchange Consortium via the jPOST partner repository (<https://repository.jpostdb.org/>) with the dataset identifier PXD053180 (JPST003175).

Received: 14 February 2024; Accepted: 8 July 2024

Published online: 29 July 2024

## References

1. Paidhungat, M. & Setlow, P. Spore germination and outgrowth. In *Bacillus subtilis and its Closest Relatives from Genes To Cells*, 537–548 (eds Sonenshein, A. L., Hoch, J. A. & Losick, R.) (Washington, DC: American Society for Microbiology, (2001)).
2. Kuwana, R. *et al.* Proteomics characterization of novel spore proteins of *Bacillus subtilis*. *Microbiology (Reading)* **148**, 3971–3982. <https://doi.org/10.1099/00221287-148-12-3971> (2002).
3. Chen, Y. *et al.* Membrane proteomes and ion transporters in *Bacillus anthracis* and *Bacillus subtilis* dormant and germinating spores. *J. Bacteriol.* **201**, e00662-e718. <https://doi.org/10.1128/JB.00662-18> (2019).
4. Lai, E. M. *et al.* Proteomic analysis of the spore coats of *Bacillus subtilis* and *Bacillus anthracis*. *J. Bacteriol.* **185**, 1443–1454. <https://doi.org/10.1128/JB.185.4.1443-1454.2003> (2003).
5. Mao, L. *et al.* Protein profile of *Bacillus subtilis* spore. *Curr. Microbiol.* **63**, 198–205. <https://doi.org/10.1007/s00284-011-9967-4> (2011).
6. Swarge, B. N. *et al.* 'One-pot' sample processing method for proteome-wide analysis of microbial cells and spores. *Proteomics Clin. Appl.* **12**, e1700169. <https://doi.org/10.1002/prca.201700169> (2018).
7. Zheng, L. *et al.* *Bacillus subtilis* spore inner membrane proteome. *J. Proteome Res.* **15**, 585–594. <https://doi.org/10.1021/acs.jproteome.5b00976> (2016).
8. Patel, R., Kumari, M. & Khan, A. B. Recent advances in the applications of ionic liquids in protein stability and activity: A review. *Appl. Biochem. Biotechnol.* **172**, 3701–3720. <https://doi.org/10.1007/s12010-014-0813-6> (2014).
9. Swatloski, R. P., Spear, S. K., Holbrey, J. D. & Rogers, R. D. Dissolution of cellulose [correction of cellose] with ionic liquids. *J. Am. Chem. Soc.* **124**, 4974–4975. <https://doi.org/10.1021/ja025790m> (2002).
10. Taoka, M. *et al.* An ionic liquid-based sample preparation method for next-stage aggregate proteomic analysis. *Anal. Chem.* **91**, 13494–13500. <https://doi.org/10.1021/acs.analchem.9b02555> (2019).
11. Taoka, M. *et al.* An ionic liquid-assisted sample preparation method for sensitive integral-membrane proteome analysis. *Anal. Biochem.* **683**, 115349. <https://doi.org/10.1016/j.ab.2023.115349> (2023).
12. Tozawa, T. *et al.* Ubiquitination-coupled liquid phase separation regulates the accumulation of the TRIM family of ubiquitin ligases into cytoplasmic bodies. *PLOS ONE* **17**, e0272700. <https://doi.org/10.1371/journal.pone.0272700> (2022).
13. Taoka, M., Fujii, M., Tsuchiya, M., Uekita, T. & Ichimura, T. A sensitive microbead-based organic media-assisted method for proteomics sample preparation from dilute and denaturing solutions. *ACS Appl. Mater. Interfaces* **9**, 42661–42667. <https://doi.org/10.1021/acsmami.7b16095> (2017).
14. Nicholson, W. L. & Setlow, P. Sporulation, germination and out-growth in ( (ed.s) Harwood, C. R. & Cutting, S. M.). *Molecular Biological Methods for Bacillus* (Chichester, United Kingdom: Wiley (1990)), 391–450.
15. Errington, J. *Bacillus subtilis* sporulation: regulation of gene expression and control of morphogenesis. *Microbiol. Rev.* **57**, 1–33. <https://doi.org/10.1128/mr.57.1.1-33.1993> (1993).
16. de Hoon, M. J., Eichenberger, P. & Vitkup, D. Hierarchical evolution of the bacterial sporulation network. *Curr. Biol.* **20**, R735–R745. <https://doi.org/10.1016/j.cub.2010.06.031> (2010).
17. Tjalsma, H., Bolhuis, A., Jongbloed, J. D., Bron, S. & van Dijl, J. M. Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol. Mol. Biol. Rev.* **64**, 515–547. <https://doi.org/10.1128/MMBR.64.3.515-547.2000> (2000).
18. Nishikawa, M. & Kobayashi, K. Calcium prevents biofilm dispersion in *Bacillus subtilis*. *J. Bacteriol.* **203**, e0011421. <https://doi.org/10.1128/JB.00114-21> (2021).

19. Doellinger, J., Schneider, A., Hoeller, M. & Lasch, P. Sample preparation by easy extraction and digestion (SPEED): A universal, rapid, and detergent-free protocol for proteomics based on acid extraction. *Mol. Cell. Proteomics* **19**, 209–222. <https://doi.org/10.1074/mcp.TIR119.001616> (2020).
20. Abele, M. *et al.* Unified workflow for the rapid and in-depth characterization of bacterial proteomes. *Mol. Cell. Proteomics* **22**, 100612. <https://doi.org/10.1016/j.mcpro.2023.100612> (2023).
21. Abhyankar, W., de Koning, L. J., Brul, S. & de Koster, C. G. Spore proteomics: the past, present and the future. *FEMS Microbiol. Lett.* **358**, 137–144. <https://doi.org/10.1111/1574-6968.12568> (2014).
22. Kuwana, R., Yamazawa, R., Ito, K. & Takamatsu, H. The study of diversity in sporulation among closely genetically related *Bacillus cereus* strains. *Biocontrol Sci.* **27**, 143–151. <https://doi.org/10.4265/bio.27.143> (2022).
23. Blagoev, B. *et al.* A proteomics strategy to elucidate functional protein–protein interactions applied to EGF signaling. *Nat. Biotechnol.* **21**, 315–318. <https://doi.org/10.1038/nbt790> (2003).
24. Anand, S., Samuel, M., Ang, C. S., Keerthikumar, S. & Mathivanan, S. Label-based and label-free strategies for protein quantitation. *Methods Mol. Biol.* **1549**, 31–43. [https://doi.org/10.1007/978-1-4939-6740-7\\_4](https://doi.org/10.1007/978-1-4939-6740-7_4) (2017).

## Acknowledgements

We thank Dr. Kazuo Kobayashi of the Nara Institute of Science and Technology for his assistance in the construction of several GFP fusions.

## Author contributions

M. T. and R. K. performed most of the experiments. T.F., A.K., Y.N., and T.U. assisted with the mass spectrometry and other experiments. T.I. and H.T. performed experiments and wrote the manuscript. All authors were involved in data discussions and preparation of the final version of the manuscript.

## Funding

This work was supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan through Grants-in-Aid for Scientific Research (grant no. 21K05322 to T. Ichimura).

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-67010-z>.

**Correspondence** and requests for materials should be addressed to M.T. or T.I.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024