EXCRETE enables deep proteomics of the microbial extracellular environment

David A. Russo[1*†], Denys Oliinyk[2†], Georg Pohnert[1], Florian Meier[2], Julie A. Z. Zedler[3]

[1] Bioorganic Analytics, Institute for Inorganic and Analytical Chemistry, Friedrich Schiller University Jena, Lessingstr. 8, 07743 Jena, Gemany
[2] Functional Proteomics, Jena University Hospital, Am Klinikum 1, 07747, Jena, Germany
[3] Synthetic Biology of Photosynthetic Organisms, Matthias Schleiden Institute for Genetics, Bioinformatics and Molecular Botany, Friedrich Schiller University Jena, Dornburgerstr. 159, 07743 Jena, Germany

[*] corresponding author
[†] these authors contributed equally

Abstract

Extracellular proteins play a significant role in shaping microbial communities which, in turn, can impact ecosystem function, human health, and biotechnological processes. Yet, for many ubiquitous microbes, there is limited knowledge regarding the identity and function of secreted proteins. Here, we introduce EXCRETE (enhanced exoproteome characterization by mass spectrometry), a workflow that enables comprehensive description of microbial exoproteomes from minimal starting material. Using cyanobacteria as a case study, we benchmark EXCRETE and show a significant increase over current methods in the identification of extracellular proteins. Subsequently, we show that EXCRETE can be miniaturized and adapted to a 96-well high-throughput format. Application of EXCRETE to cyanobacteria from different habitats (*Synechocystis* sp. PCC 6803, *Synechococcus* sp. PCC 11901, and *Nostoc punctiforme* PCC 73102), and in different cultivation conditions, identified up to 85% of all predicted secreted proteins. Finally, functional analysis reveals that cell envelope maintenance and nutrient acquisition are central functions of the cyanobacterial secretome. Collectively, these findings challenge the general belief that cyanobacteria lack secretory proteins and point to a functional conservation of the secretome across freshwater, marine, and terrestrial species.

Introduction

Bacteria produce and release a remarkable repertoire of small molecules and proteins to communicate and respond to stimuli [1]. Secreted molecules mediate interactions with the environment and the surrounding microbiome, and numerous biotechnological applications take advantage of the secretion machinery. Bacterial protein secretion has mainly been studied in pathogenic and polymer-degrading bacteria due to their importance in human health and the environment [2]. However, so far, protein secretion remains understudied in many environmentally relevant bacteria due to their complex, intractable, habitats where secreted proteins are dilute and difficult to extract. One such example are cyanobacteria.

Cyanobacteria are ubiquitous oxygenic photosynthetic bacteria that contribute at least 10% of global net primary production [3,4], and can be found in a range of environments rich in salt [5], lipids [6], polysaccharides [7] , or in biofilms and microbial mats [8]. The traditional

view of a cyanobacterium is that of a self-sufficient photoautotroph that lacks secretory enzymes and is surrounded by a highly impermeable outer membrane which only allows the diffusion of small molecules [9–11]. Challenging this paradigm, emerging evidence suggests that proteins in the cell envelope and extracellular space play key roles in the cyanobacterial lifestyle.

In cyanobacteria, only four out of the eleven bacterial protein secretion systems have so far been identified: type I (T1SS), type IV (T4SS), type V (T5SS) and type IV pili (T4P) [12]. Structurally, cyanobacterial secretion systems resemble their counterparts in other bacteria. Functionally, however, they are involved in remarkably different processes. The T1SS, for example, plays a role in heterocyst development and deposition of the S-layer [13–16]. And T4P are involved in phototaxis, DNA uptake, aggregation, and flotation [17–22]. Interestingly, the recent discovery of multiple protein secretion mechanisms that do not involve the aforementioned systems suggests that there may be a myriad of secreted proteins which remain largely unexplored [23–26].

Mass spectrometry (MS)-based proteomics has become the preferred method to study protein secretion and characterize the exoproteome [27]. To date, MS-based exoproteomics studies in cyanobacteria have identified up to a few hundred proteins with only a fraction of these considered truly secreted [28–30]. This low number is in stark contrast with the abundance of secretory processes already described in cyanobacteria, thus suggesting there is potential for further exploration. To handle the high dilution and sample complexity of the cyanobacterial exoproteome, current workflows for MS-based exoproteomics start with concentration and clean-up steps [21,31,32]. Protein precipitation and centrifugal concentrators are widely used, however these approaches require excessive handling steps and are prone to protein loss. Therefore, the full description of the cyanobacterial exoproteome remains a challenge.

To address this challenge, we developed EXCRETE (enhanced exoproteome characterization by mass spectrometry), an MS-based workflow for direct analysis of microbial exoproteomes adapted from solid-phase enhanced protein aggregation protocols (e.g., SP3 [33], protein aggregation capture (PAC) [34]). Utilizing cyanobacteria as a test system, we show that EXCRETE can robustly characterize the exoproteome across species, media, and growth conditions, independent of extracellular matrix, with minimal time and sample handling. Our workflow should be broadly applicable to microbes from a wide range of habitats, with the potential to open new avenues of investigation in microbial exoproteomics.

Materials and methods

Cyanobacterial strains and growth conditions

Synechocystis sp. PCC 6803 substrain PCC-M (hereafter *Synechocystis* wild-type (WT)) and the *Synechocystis* Δ*hfq* mutant [35] were kindly donated by Annegret Wilde. *Synechococcus* sp. UTEX 3153 (also known as *Synechococcus* sp. PCC 11901) was obtained from the UTEX culture collection. *Nostoc punctiforme* PCC 73102 was kindly donated by Elke Dittmann. *Synechocystis* was maintained on BG-11 medium [36] supplemented with 10 mM 2-tris(hydroxymethyl)-methyl-2-amino 1-ethanesulfonic acid (TES) buffer (pH 8.0) and 1.5% (w/v) Kobe I agar. *Synechococcus* was maintained on AD7 medium [37] supplemented with 1.5% (w/v) Kobe I agar. Both strains were kept at 30 °C with continuous illumination of 40 µmol photons m^{-1} s^{-1}. *Nostoc* was maintained in BG-11 medium supplemented with 10 mM TES buffer (pH 8.0) at 22 °C with continuous illumination of 30 µmol photons m^{-1} s^{-1}. For the proteomics experiments, *Synechocystis* WT and the Δ*hfq* mutant were grown in 100 mL conical flasks with 20 mL of BG-11 TES medium at 30 °C, continuous illumination

of 50 µmol photons m^{-1} s^{-1} and shaking at 150 rpm. The precultures of the Δ*hfq* mutant were supplemented with 15 µg mL^{-1} chloramphenicol. The antibiotic was omitted during experiments to prevent distorting effects. To induce the formation of aggregates, *Synechocystis* was grown in a CellDEG high-density system (HDC 6.10 starter kit; CellDEG) [38]. The CellDEG system was set up as previously described [39] and shaking was set to 100 rpm. *Synechococcus* was grown in BA+ medium (DSMZ ID 1677) in a CellDEG high-density system as previously described [39] with shaking set to 250 rpm. *Nostoc* was grown in T-75 tissue flasks (Sarstedt) with 20 mL of BG-11 TES medium, supplemented with 10 times higher concentration of Na_2CO_3, at 21 °C, continuous illumination of 50 µmol photons m^{-1} s^{-1} and no shaking. All pre-cultures were washed twice in their respective media. *Synechococcus* and *Synechocystis* cultures inoculated to a starting $OD_{750\ nm}$ of 0.4 and *Nostoc* cultures to a wet weight of 2.5 g L^{-1}. All cultures were grown for 72 h before harvesting for protein extraction. All experiments were conducted on a Unimax 1010 orbital shaker (Heidolph Instruments).

Preparation of endoproteome and exoproteome fractions

To obtain the exoproteome fraction, cultures were centrifuged for 10 min at 5000 × g. The supernatant was removed, centrifuged again for 10 min at 10 000 × g and transferred to a fresh microcentrifuge tube. Samples were kept on ice until processing. Endoproteome fractions were obtained by centrifuging an $OD_{750\ nm}$ equivalent of 3 for 10 min at 10 000 × g. The pellet was then washed and resuspended in 300 µL of lysis buffer (25 mM Tris-HCl, 5% (w/v) glycerol, 1% (v/v) Triton X-100, 1% (w/v) sodium deoxycholate, 0.1% (w/v) sodium dodecyl sulfate (SDS) and 1 mM EDTA). Cells were broken with zirconium oxide beads (diameter 0.15 mm) using a Bullet Blender Storm 24 (Next Advance) with three cycles of 5 min. Cell lysates were then centrifuged at 10 000 × g for 10 min and the resulting supernatant transferred to fresh microcentrifuge tubes for protein content determination. Protein content of all samples was determined with a Pierce™ BCA assay kit (Thermo Fisher Scientific).

Ultrafiltration and in-solution digestion workflow

Exoproteome fractions were concentrated to 1 mg mL^{-1} total protein (approximately 10 times concentration) using an Amicon Ultra-0.5 Centrifugal Filter Unit with a molecular weight cutoff of 3 kDa (Merck Millipore). A 10 µL aliquot of the concentrated exoproteome fraction was reduced with 5 mM tris (2-carboxyethyl) phosphine (TCEP) and alkylated with 5.5 mM chloroacetamide (CAA) at room temperature for 5 min. Subsequently, 90 µL of 25 mM ammonium bicarbonate containing 0.5 µg of MS grade Trypsin/LysC (Promega) (enzyme/protein ratio of 1:20 (w/w)) were added to the sample for in-solution digestion and incubated for 4 h at 37 °C and 1000 rpm. Peptide purification and desalting was performed as described below.

EXCRETE workflow

Supernatant volumes equivalent to 10 µg (microcentrifuge tube) or 3 µg (microplate) of protein were harvested and transferred to 2 mL microcentrifuge tubes or 96-well microplates, respectively. Sample volumes were normalized to the sample with the highest volume to ensure similar mixing dynamics. NaCl and SDS were added to a final concentration of 10 mM and 1% (w/v), respectively. Samples were then reduced with 5 mM TCEP and alkylated with 5.5 mM CAA at room temperature for 5 min. To induce protein aggregation, LC-MS grade ethanol was added to a final concentration of 50% (v/v) followed by the addition of SiMAG-Carboxyl magnetic particles (product no. 1201, Chemicell) to a final concentration of 0.5 µg $µL^{-1}$. Samples were then incubated for 10 min with shaking

at 1000 rpm (750 rpm for microplates) on a Biometra TS1 ThermoShaker (Analytik Jena). Subsequently, magnetic particles were separated on custom made magnetic racks for 60 s. Supernatants were removed and the magnetic particles were washed, on-magnet, 3 times with 80% (v/v) ethanol. Washes were discarded. Following the washing steps, samples were removed from the magnetic racks and air dried for 10 min at room temperature to remove residual ethanol. The magnetic particles were then resuspended in 100 µL of 25 mM ammonium bicarbonate containing 0.5 µg of MS grade Trypsin/LysC (Promega) (enzyme/protein ratio of 1:20 (w/w)) for on-bead digestion. Samples were sonicated for 1 min to reconstitute the magnetic particles and incubated overnight (microcentrifuge tubes) or for 4 h (microplate) at 37 °C and 1000 rpm. Following protein digestion, samples were sonicated for 2 min, magnetic particles were separated for 60 s and supernatants were recovered.

Peptide purification and desalting

Recovered supernatants were diluted into 300 µL wash buffer 1 (1% (v/v) trifluoroacetic acid (TFA) in isopropanol), transferred to SDB-RPS StageTips [40] and centrifuged for 10 min at 1500 × g. StageTips were washed with 100 µL of wash buffer 1 and, subsequently, with 100 µL of wash buffer 2 (0.2% (v/v) TFA in 5% (v/v) acetonitrile (ACN)) for 8 min at 1500 × g. Peptides were eluted by adding 60 µL of freshly prepared SDB-RPS elution buffer (0.2% (v/v) NH_4OH in 60% (v/v) ACN) and centrifuging for 10 min at 1000 × g. Eluates were immediately dried under vacuum at 45 °C and stored at -20 °C until analysis. In the microplate workflow, StageTips were processed using a Spin96 device [41]. Before analysis, peptides were resuspended in MS loading buffer (0.1% (v/v) TFA in 2% (v/v) ACN) and the concentration was measured with a NanoDrop Spectrophotometer (Thermo Fisher Scientific) in order to normalize injections to 200 ng of peptides.

Liquid chromatography-MS analysis

Purified and desalted peptides were separated by nanoflow reversed-phase liquid chromatography in a nanoElute system (Bruker Daltonics) within 60 min at a flow rate of 0.5 µl min^{-1} on a 15 cm x 75 µm column packed, in-house, with 1.9 µm C_{18} beads. Mobile phase A consisted of 0.1% (v/v) formic acid and B of 0.1% (v/v) formic acid in ACN. Peptides, upon elution from the column, were electrosprayed with a CaptiveSpray (Bruker Daltonics) into a trapped ion mobility spectrometry (TIMS) quadrupole time-of-flight mass spectrometer (timsTOF HT, Bruker Daltonics). Data were acquired in data-independent acquisition with parallel accumulation–serial fragmentation (diaPASEF) mode [42] with an equidistant window scheme in the *m/z* and ion mobility dimensions. The ion mobility range was set from $1/K_0$ = 0.6 to 1.6 Vs cm^{-2} and equal ion accumulation time and ramp times in the dual TIMS analyzer of 100 ms each. The collision energy was linearly decreased from 59 eV at $1/K_0$ = 1.4 Vs cm^{-2} to 20 eV at $1/K_0$ = 0.6 Vs cm^{-2}. For all experiments the TIMS elution voltages were calibrated by known $1/K_0$ values from at least two out of three ions from Agilent ESI LC/MS tuning mix (*m/z*, $1/K_0$: 622.0289, 0.9848 Vs cm^{-2}; 922.0097, 1.1895 Vs cm^{-2}; and 1221.9906, 1.3820 Vs cm^{-2}).

Raw data processing

Raw files from diaPASEF were processed in Spectronaut v17 [43] (Biognosys) with spectrum libraries built directly from the DIA experiments with the directDIA+ workflow. False discovery rates were set to <1% at peptide and protein levels using a target-decoy based approach. Cysteine carbamidomethylation was set as a fixed modification and protein N-terminal acetylation and methionine oxidation were set as variable modifications. All spectra were matched against *Synechocystis* (GenBank: CP003265.1), *Synechococcus*

(UniProt: UP000304321) and *Nostoc* (UniProt: UP000001191) reference proteomes (accessed September 2022). Quantification values were filtered by q-value and the 'Automatic' normalization mode was defined for cross run normalization, while the imputation strategy was set to 'None'. Report data matrices were exported on protein group level for downstream bioinformatic analysis.

Bioinformatic analysis

Proteins were annotated using EggNOG v5.0 [44], PsortB v3.0 [45], SignalP 6.0 [46] and UniProtKB [47]. Proteins were considered secreted when PsortB or UniProtKB predicted the location as "Periplasm", "Outer Membrane", or "Extracellular". In addition, proteins with a location predicted as "Unknown" were considered secreted when a signal peptide was detected by SignalP. Data analysis and visualization were performed using custom scripts in R (4.0.1) and Python (3.9.4) with packages data.table (1.14.2), dplyr (1.0.7), ggplot2 (3.3.5), tidyR (1.1.14), pandas (1.1.15), numpy (1.22.2), plotly (5.4.0), scipy (1.7.3).

Statistical analysis

All experiments were performed with three to five replicates. Data are presented as the mean and individual data points are plotted unless stated otherwise. Detailed information is given in the respective figure legend. A protein group was considered identified when it was present in at least 70% of the replicates with a minimum of three replicates. Missing values were then imputed using the k-nearest neighbors algorithm. In Fig. 1d and Fig. 2a, the number of protein groups identified is given before filtering and imputation. Differential protein analysis in Fig. 5d was done using a Student's two-sample unpaired t-test with permutation-based multiple test correction. Differentially produced proteins were selected from all identified proteins with a cutoff criterion of fold change = 2 and adjusted p-value < 0.05.

Data availability

The raw MS data and associated tables have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository and are publicly available with the identifier PXD047856.

Code availability

Python code to reproduce the analysis and figures is publicly available via the GitHub repository (https://github.com/MeierLab).

Results

Design of a solid-phase enhanced protein aggregation workflow for cyanobacterial exoproteomics

Exoproteomic samples from cyanobacteria typically suffer from high dilution and the presence of substantial amounts of other compounds such as inorganic salts, lipids, and polysaccharides. To date, these challenges have been approached with labor-intensive concentration methods such as protein precipitation, ultrafiltration, and dialysis [21,30–32,48]. These techniques are time-consuming, prone to loss of material, and, ultimately, lead to unsatisfactory results (Fig. 1a). Therefore, we sought to develop a fast, robust, and unbiased workflow that allows easy exoproteome concentration and subsequent removal of contaminants. To this end, we took inspiration from solid-phase enhanced protein aggregation protocols (e.g., SP3, PAC) that have already been applied to exoproteomic

samples, albeit of human model systems [34,49]. These protocols rely on a phenomenon where precipitated and aggregated proteins bind to a solid phase irrespective of its surface chemistry. Paramagnetic beads are the most widely used solid phase, however a variety of other microparticles, including glass beads, have also been successfully applied [33,34,50,51]. Protein aggregation and binding is considered to be unbiased and can be induced with a variety of solvents (e.g., acetonitrile, ethanol) and conditions (e.g., high temperature, high salt) [34,50]. Solid-phase enhanced protein aggregation workflows are scalable and can be performed in less than 1 h in a single container, thus allowing for minimal sample loss and same-day analysis (Fig. 1a). In our conceptualization of these protocols, termed EXCRETE (Fig. 1b), chloroacetamide (CAA) and tris(2-carboxyethyl)phosphine (TCEP) are used to allow the alkylation and reduction steps to be combined with the protein aggregation step (Fig. 1b). Then, we chose carboxylate-modified paramagnetic beads as an economic solid-phase that can be added at a high concentration (0.5 µg mL^{-1}) to facilitate the capture of proteins in typically dilute extracellular samples [34]. Protein aggregation is induced by adding ethanol to a final concentration of 50% (v/v). Importantly, the paramagnetic beads can then be extensively washed with 80% ethanol (v/v) to remove co-aggregating contaminants (Fig. 1b). Protein digestion is then performed on-bead in as little as 4 h. Following digestion, peptides are recovered by collecting the supernatant following bead collection on a magnetic rack.

To benchmark our bead-based workflow, we compared the performance of EXCRETE against an ultrafiltration-based method, as reported previously [16,52]. For the comparison, we chose exoproteomic samples of the marine cyanobacterium *Synechococcus* sp. UTEX 3154 (hereafter *Synechococcus*). *Synechococcus* is a recently isolated unicellular marine cyanobacterium of biotechnological interest due to its fast, sustained, biomass accumulation [53]. This also allowed us to test how EXCRETE performs on salt-rich samples. Regarding the number of identified peptides and proteins, EXCRETE clearly outperformed ultrafiltration. On average, 3974 peptides and 639 proteins were identified using EXCRETE (Fig. 1c, d, Supplementary Data 1). In the case of ultrafiltration, we identified, on average, 2549 peptides and 339 proteins (Fig. 1c, d, Supplementary Data 2). On-bead protein digestion has been shown to reduce the number of missed tryptic cleavages in comparison to in-solution digestion [34], therefore we investigated whether this was also the case in our benchmarking experiment. EXCRETE and ultrafiltration followed by in-solution digestion showed similar number of missed cleavages (Supplementary Fig. 1a). We then proceeded to assess the precision of each workflow by calculating the coefficient of variation (CV) amongst the replicates. and found a median CV of 14% for the ultrafiltration workflow and 27% for EXCRETE (Supplementary Fig. 1b, c). This is in line with previous reports showing that bead-based methods can present higher CVs [54]. Finally, regarding protein identifications, we observed an overlap of 47% of proteins with an additional 49% only identified with EXCRETE (Fig. 1e). The additional proteins identified included several low abundance proteins, thus demonstrating that EXCRETE is robust across a wide dynamic range (Supplementary Fig. 1b).

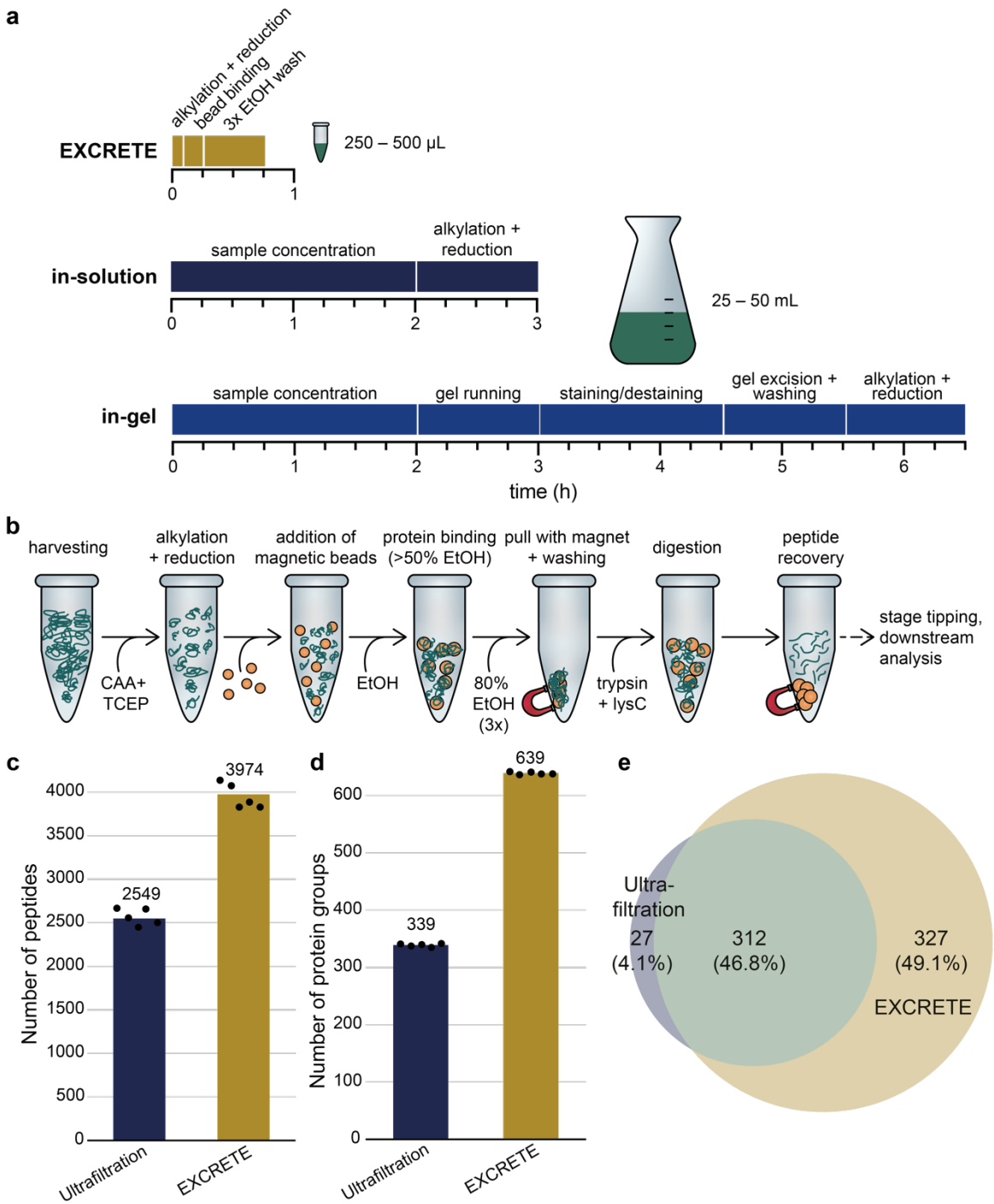

Fig. 1. EXCRETE outperforms traditional workflows for exoproteomic sample preparation. **a** Comparison of workflows, time required and input volume between the classic in-gel and in-solution approaches and EXCRETE. **b** Detailed EXCRETE sample preparation workflow. **c**, **d** Number of peptides (c) and protein groups (d) identified by ultrafiltration and EXCRETE. **e** Venn diagram representing the overlap of protein identifications with ultrafiltration and EXCRETE.

Miniaturization of the EXCRETE workflow allows for increased throughput without loss of information

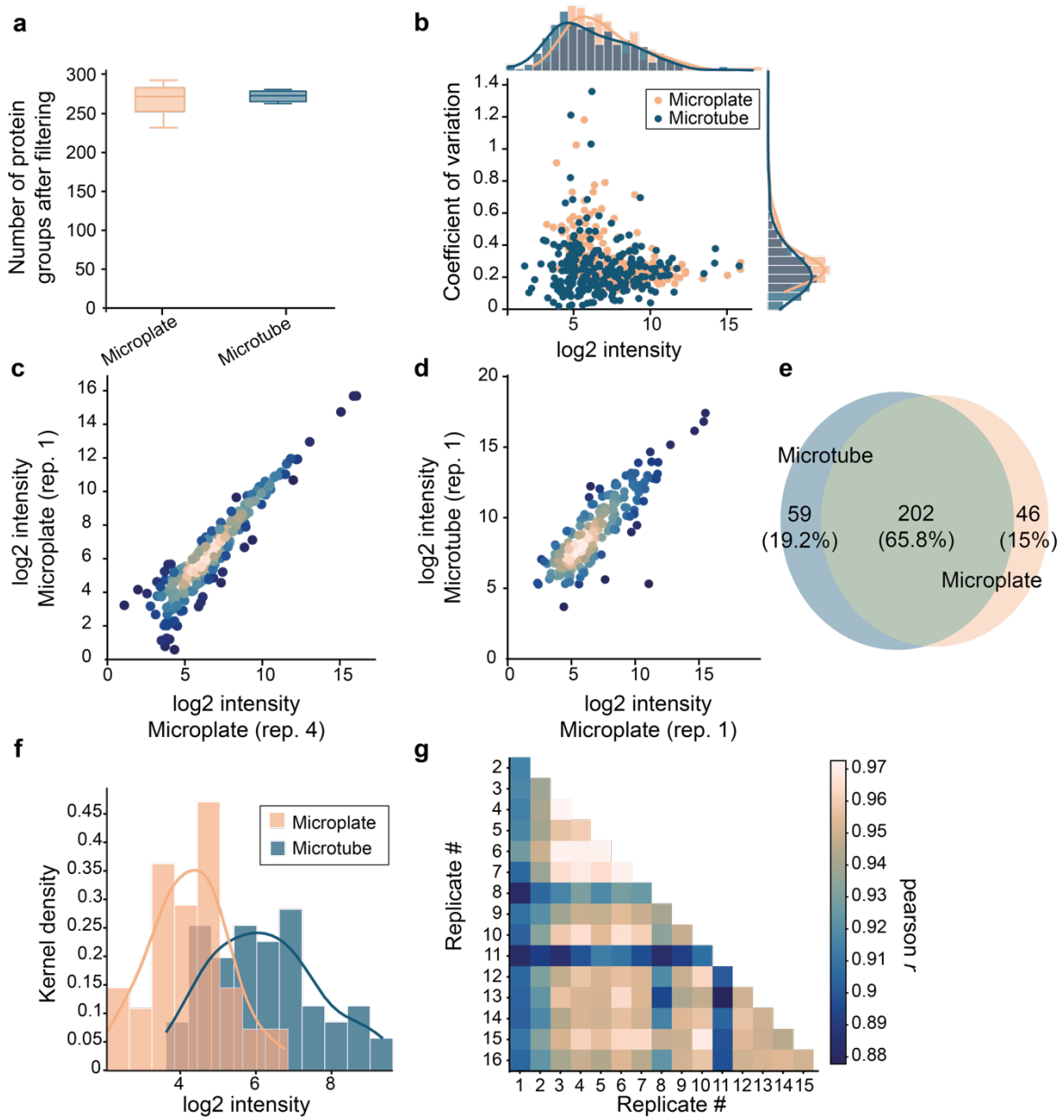

Fig. 2 EXCRETE can be miniaturized to a microplate format without loss of information **a** Number of protein groups identified by microtube- and microplate-based workflows following filtering and imputation. Centre line of boxplots, median; box limits, upper and lower quartiles; whiskers, minimum to maximum values. **b** Coefficients of variation (CVs) ordered by protein intensity. On the secondary x-axis histogram and density plots representing the frequency distribution of protein intensities are shown. On the secondary y-axis histogram and density plots representing the frequency distribution of CVs are shown. Dots represent means of biological replicates. **c, d** Correlation of intensities of identified proteins in between biological replicates processed with (**c**) the microtube- based method or (**d**) the microtube- and microplate-based methods. **e** Venn diagram representing the overlap of protein identifications with microtubes and microplates. **f** Intensity distribution of unique proteins identified with microtubes and microplates. **g** Pairwise Pearson correlations of sixteen microplate workflow replicates.

Having successfully benchmarked EXCRETE, we aimed to miniaturize the workflow to allow for high-throughput analysis in 96-well microplates. To adapt to a microplate format, we decreased all initial volumes of culture supernatant to a maximum of 150 µL, consequentially decreasing the starting protein amount threefold from ~10 to ~3 µg. To evaluate the efficiency of our miniaturized high-throughput workflow in comparison with

its microtube-based analogue, we analyzed the *Synechococcus* exoproteome with both approaches. Regarding protein identification, both approaches were quantitatively similar with 272 ± 40 and 273 ± 7 protein groups identified after filtering and imputation, using the microplate and microtube workflows, respectively (Fig. 2a, Supplementary Data 3, 4). Our data revealed good quantitative reproducibility with median CV of 24% in the microtube workflow and 22% for the microplate workflow (Fig. 2b). In addition, the microtube workflow covered a marginally higher fraction of low abundance proteins (Fig. 2b). Quantitative precision was high for both approaches with a Pearson correlation coefficient of r = 0.94 within replicates processed with microplates (Fig. 2c) and r = 0.91 between replicates processed with microplates and microtubes (Fig. 2d). We then sought to compare the accuracy of protein identification between the microplate and microtube workflows. More than 75% of total proteins identified were found using both approaches (Fig. 2e). The unique proteins identified in each workflow were generally of lower abundance, in particular for the microplate workflow (Fig. 2f). Finally, to evaluate how the reproducibility of the microplate-workflow scales towards higher throughput, we measured sixteen replicates processed in parallel with the same workflow. The pairwise Pearson correlation coefficient was higher than 0.88 across all injections (Fig. 2g). Together, these results show that, despite a lower protein input, the microplate workflow is qualitatively and quantitatively equivalent to the microtube-based workflow. Miniaturization of EXCRETE will allow for the analysis of dozens of secretomes per day and opens the door to automation of exoproteome analysis on robotic liquid handling platforms [55].

EXCRETE allows for deep exoproteomic profiling of cyanobacteria from a range of habitats

Solid-phase enhanced protein aggregation protocols have already been successfully applied to analyze the exoproteome of human tissue cultures [34,49]. However, in contrast to these, bacteria grow in a wide range of environments that can be rich in salt, lipids, and polysaccharides. Given that EXCRETE performed well with high-salt exoproteome samples from a marine cyanobacterium, we decided to test our workflow on another two cyanobacteria with different extracellular matrices. To this end, we compared the performance of EXCRETE on three cyanobacteria from distinct habitats: the aforementioned marine *Synechococcus*; *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*), a motile unicellular freshwater cyanobacterium; and *Nostoc punctiforme* PCC 73102 (hereafter *Nostoc*), a terrestrial N_2-fixing filamentous cyanobacterium which retains the developmental complexity of field isolates and is a model for plant-cyanobacteria symbiosis [56,57]. *Synechocystis* and *Nostoc* are also known to secrete large amounts of polysaccharides [57,58], thus allowing us to test EXCRETE on polysaccharide-rich matrices. Analysis of 10 µg exoproteomic samples (approximately 600 µL of harvested supernatant for *Synechocystis* and 300 µL for *Synechococcus* and *Nostoc*) led to the identification of 4985, 2459, and 3200 peptides and 1092, 557 and 839 protein groups, at 1% false discovery rate (FDR) at both peptide and protein level, in *Synechocystis*, *Synechococcus*, and *Nostoc*, respectively (Fig. 3a, b). We observed that 80% data completeness (i.e., values present in 80% of replicates) was achieved for 97.5%, 91.2% and 62.8% of all protein groups identified in Synechocystis, Synechococcus, and Nostoc, respectively (Fig. 3c). Following data filtering and imputation, 885, 261 and 668 proteins, in *Synechocystis*, *Synechococcus*, and *Nostoc*, respectively, were kept for further analysis (Supplementary Data 4, 5, 6). Compared to previous studies using TCA precipitation, this is a 7.5-fold increase for *Synechocystis*, a 2.3-fold increase for *Nostoc*, and the first report for this *Synechococcus* species (Supplementary Fig. 2). All three species showed good biological reproducibility with median CVs of 43%, 34% and 22% for *Synechocystis*, *Synechococcus*, and *Nostoc*, respectively, and a general trend of lower dispersal at higher

intensities (Fig. 3d, e). Altogether, these results demonstrate that our approach is robust and performs well for all species and extracellular matrices tested here.

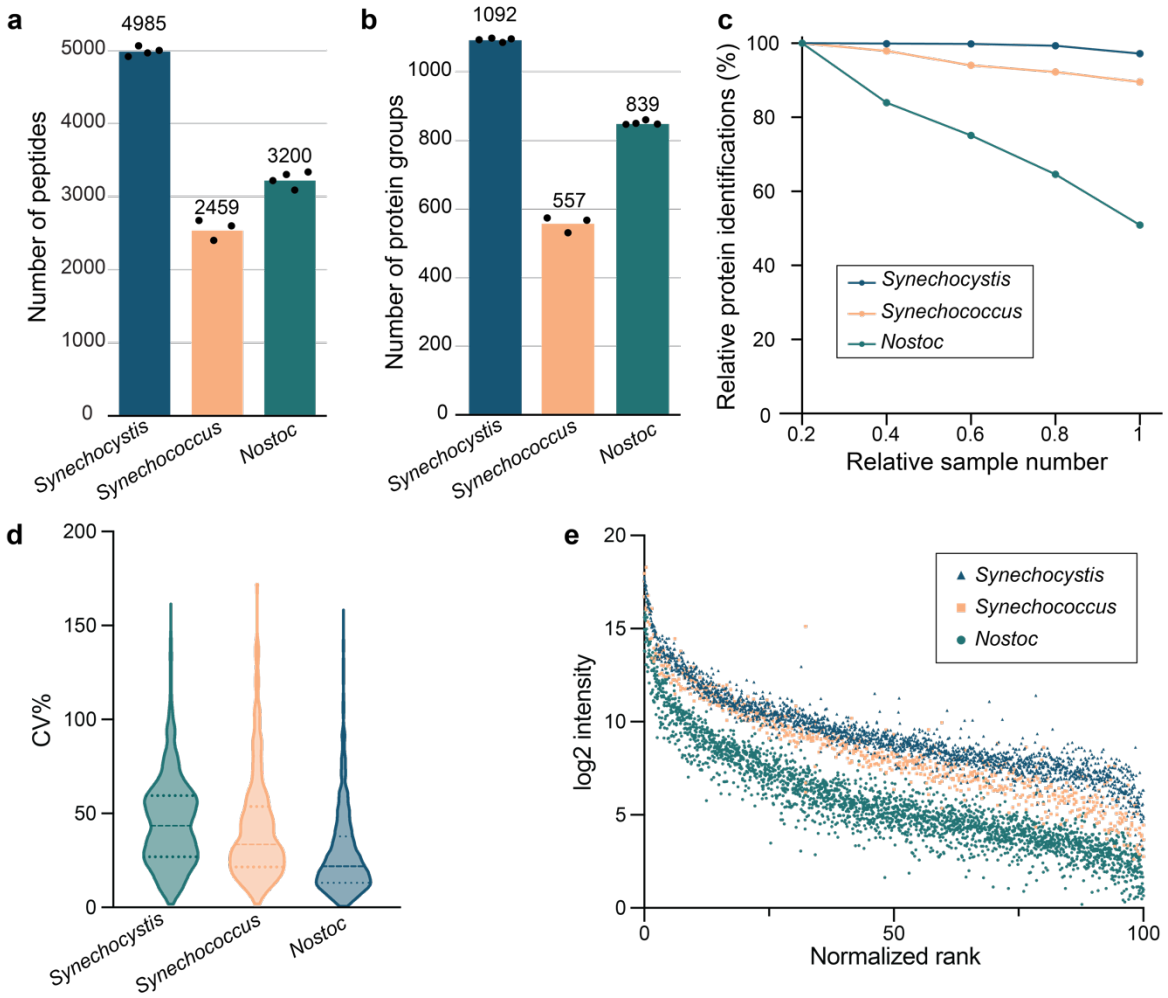

Fig. 3 The EXCRETE workflow enables robust and direct sampling of exoproteomes from *Synechocystis*, *Synechococcus* and *Nostoc* exoproteins cultures. **a, b** Number of peptides (**a**) and protein groups (**b**) identified in the exoproteome of each species. Means are shown above the bars. Black dots represent biological replicates (n = 4 for *Synechocystis* and *Nostoc*, n = 3 for *Synechococcus*). **c** Coefficient of variation (CV) of proteins identified across all biological replicates of each species. **d** Protein intensities ordered by rank. Dots represent individual protein replicates. Protein ranks have been normalized to a range between 0 and 100. **e** Percentage of identified proteins with an increasing fraction of samples for each species.

The cyanobacterial secretome is functionally conserved across species

Besides secretory routes, proteins can be released to the extracellular milieu through unspecific mechanisms such as cell lysis and defective cell division [59,60]. Therefore, having identified a large number of exoproteins in this work, we sought to identify which proteins were likely secreted (i.e., the secretome). To define the secretome of each cyanobacterium we combined PsortB [45], SignalP [46] and UniprotKB [47] for the prediction of subcellular location and signal peptide recognition. Regarding location, the secretome was defined as proteins predicted to be translocated to the periplasm, outer membrane or extracellular milieu [61]. Regarding the latter, in bacteria the presence of a signal peptide typically suggests a protein is secreted. However, some secretion pathways (e.g., T1SS) do not utilize classical N-terminal signal peptides, therefore the presence of a signal peptide was

not required for a protein to be considered secreted. In addition, signal peptides directing proteins to the thylakoid membrane cannot be distinguished from those directing proteins to the exterior of the cell [12,62]. Therefore, proteins predicted to have signal peptides, but also predicted to be localized in the thylakoid membrane, were removed from the final set of secreted proteins.

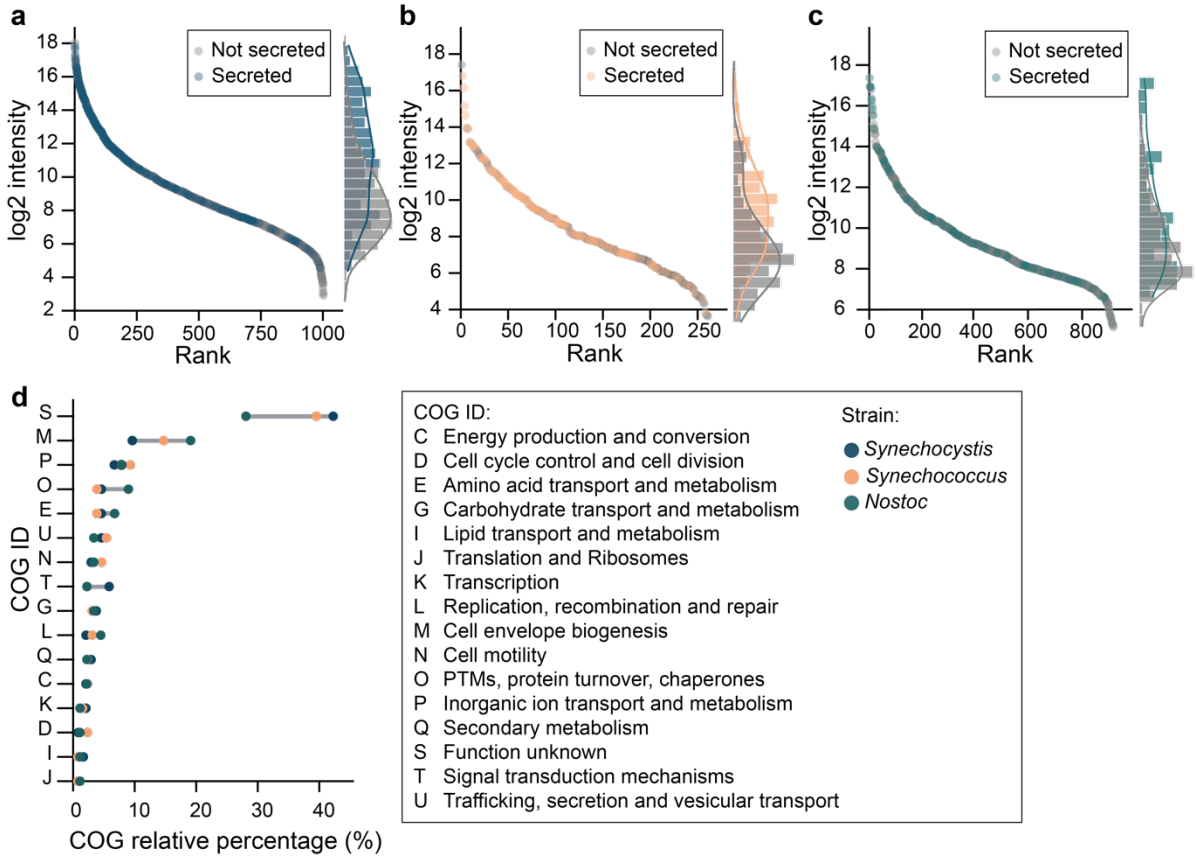

Fig. 4. The secretome functional profile is conserved across *Synechocystis*, *Synechococcus*, and *Nostoc*. **a–c** Protein intensities ordered by rank with indication of secreted (colored dots) and non-secreted (grey dots) proteins in (**a**) *Synechocystis*, (**b**) *Synechococcus* and (**c**) *Nostoc*. Dots represent means of biological replicates. On the secondary y-axis histogram and density plots representing the frequency distribution of secreted and non-secreted proteins are shown. **d** Clusters of Orthologous Genes (COG) analysis of secreted proteins in each species. COG relative percentage represents the number of times a COG ID appears in relation to the total numbers of COG IDs.

Following the criteria defined above, we detected 156, 117 and 81 secreted proteins in *Synechocystis*, *Synechococcus*, and *Nostoc* (corresponding to 62%, 46% and 14% of all potentially secreted proteins in the respective proteomes) (Supplementary Data 4, 5, 6). To determine whether the proteins we predicted as secreted were present at higher intensities in the exoproteome, all proteins identified were plotted by rank and intensity. In support of our experimental approach, secreted proteins clustered at the top of the curve (Fig. 4a, b, c) and accounted for 56%, 60% and 48% of the total protein intensity in *Synechocystis*, *Synechococcus*, and *Nostoc*, respectively (Supplementary Fig. 3). The identification of only 14% of the potential *Nostoc* secretome, in comparison to 62% in *Synechocystis*, is curious due to the fact that *Nostoc* encodes for a much larger number of secretion systems than *Synechocystis* [12]. This is in line with previous work showing that symbiotic cyanobacteria, such as *Nostoc*, exhibit high levels of proteomic plasticity and

only express a small fraction of their proteome at any given point in time [30,63,64]. This is also in agreement with the coverage of the respective endoproteomes where 65% and 20% of all proteins in *Synechocystis* and *Nostoc*, respectively, have been identified by MS-based proteomics [65,66].

We then proceeded to investigate which proteins were identified in the respective secretomes to confirm whether the patterns observed had support in the literature. First, we observed that approximately 78% of all proteins predicted as secreted contain a Sec signal peptide. This is in line with what is known in cyanobacteria where the majority of proteins are translocated via Sec [12]. Next, PilA1, the major component of the retractable fiber of the T4P system, was the most abundant secreted protein in *Synechococcus* (UniProtKB: A0A4P8WZ44) and fourth most abundant in *Synechocystis* (UniProtKB: P73704). This was expected as cell surface appendages are amongst the most prominent proteins typically found in cyanobacterial secretomes [67]. Despite *Nostoc* possessing a functional T4P system [68], PilA1 was not identified in the secretome. This is in agreement with previous work showing that, in *Nostoc*, pili adhere to the filament surface and may, additionally, be encased in a motility-associated polysaccharide layer [57]. Other examples of cell surface appendages found include PilX1 (*Synechocystis*, UniProtKB: P73704), where a potential role in flocculation has been proposed [69], and CccS (*Synechocystis*, UniProtKB: P74672), involved in the construction of cell surface components [70]. The cyanobacterial exoproteome has also been implicated in microelement acquisition and detoxification [71]. Our data also strongly support this role. The second most abundant secreted proteins in *Synechocystis*, *Synechococcus* and *Nostoc* are futA2 (UniProtKB: Q55835), a periplasmic iron-binding protein [72], PhoX (UniProtKB: A0A4P8X4T0), a secreted alkaline phosphatase [73], and a phosphonate ABC transporter (UniProtKB: B2J991), respectively. Porins harboring SLH (S-layer homology) and OprB (carbohydrate-selective porin OprB) protein domains, which have also been connected to OM transport functions, were also identified in all three species. Finally, it is also noteworthy that we identified several large (100 – 490 kDa) secreted proteins in the *Synechocystis* exoproteome (Supplementary Table 1). These proteins were generally characterized by the absence of signal peptides, the presence of Ca^{2+} binding motifs, pI values below pH 5 and an amino acid composition rich in glycine residues but with only a few, if any, cysteines. Taken together, this suggests that these proteins are T1SS substrates [74].

Next, to explore the functional aspects of the individual secretomes, we annotated the subset of secreted proteins with their respective Clusters of Orthologous Genes (COG) categories. We found that in *Synechocystis*, *Synechococcus*, and *Nostoc*, 20.6%, 0% and 5.6% of all secreted proteins lacked a match to the COG database and a further 22.8%, 39.5% and 22.5% were annotated as proteins of unknown function. Remarkably, amongst the proteins with known COG functions, all three secretomes showed similar functional profiles (Fig. 4d). Analyzing the COG distribution, the most highly represented COG functions in the three secretomes were cell wall/membrane/envelope biogenesis (9–19%), posttranslational modification, protein turnover, chaperones (4–9%) and inorganic ion transport and metabolism (3–9%) (Fig. 4d). Amongst all the categories, signal transduction mechanisms showed the largest deviation, with this category being represented almost four times higher in *Synechocystis* than in *Synechococcus* and *Nostoc*. Altogether, these results suggest that cell envelope maintenance and nutrient acquisition are core functions of the secretome and are conserved across cyanobacteria from different habitats.

The stable Synechocystis secretome is enriched in proteins involved in cell envelope maintenance

Bacterial secretomes are remarkably dynamic and have been shown to respond to different media and growth conditions [31,71,75]. The proportion of the secretome that does not vary across conditions can be considered "stable" and is likely constituted by proteins that perform essential services. Therefore, we analyzed the *Synechocystis* secretome in different conditions to determine whether a stable secretome would emerge.

To this end, we compared the *Synechocystis* secretome in the WT condition as described above (Fig. 3, Supplementary Data 4, Supplementary Fig. 4a) with that of a bloom condition and a Δ*hfq* mutant (Fig. 5a). In the bloom condition *Synechocystis* was cultivated in high-CO_2 to form aggregates embedded in a viscous EPS-rich extracellular matrix (Supplementary Fig. 4b). In this condition we hypothesized that, together with EPS secretion we would see an increase in protein secretion and an expanded secretome. In the Δ*hfq* condition we used a *Synechocystis* mutant that lacks several characteristics that depend on secreted proteins such as motility, natural competence and aggregation (Supplementary Fig. 4c) [20,32,35,76]. In addition, this mutant also exhibited a 73% decrease in the secretion of a nanoluciferase reporter [24]. Therefore, in the Δ*hfq* condition we expected a reduced secretome. Accordingly, we identified 2005 and 397 proteins in the exoproteomes of the bloom condition and the Δ*hfq* condition, respectively (Supplementary Fig. 5, Supplementary Data 7, 8). Of these, 212 and 127 were considered secreted (Supplementary Fig. 5, Supplementary Data 7, 8). In comparison to the WT condition, this represents an increase of 56 proteins in the bloom condition and a decrease of 29 proteins in the Δ*hfq* condition.

Looking at the individual conditions, when forming bloom-like aggregates, *Synechocystis* exhibited an expanded secretome, covering nearly 85% of potentially secreted proteins, with 55 proteins unique to this condition (Fig. 5b, Supplementary Data 7). Previous work suggests that EPS-rich extracellular matrices can sequester small molecules and nutrients, thus acting as an extracellular pool of resources available for mobilization by the cyanobacterial secretome [77,78]. We thus reasoned that the expanded secretome could be linked to the mobilization of extracellular resources. To investigate this hypothesis, we analyzed which COG categories were enriched in comparison to the proportion of the secretome shared by all three conditions. In agreement with our hypothesis, we observed a combined 21% increase in the inorganic ion and amino acid transport categories (Fig. 5c). These findings support the idea that cyanobacteria are not only major contributors to the extracellular matrix but can also utilize it as a nutrient resource.

In the Δ*hfq* condition, *Synechocystis* exhibited a reduced secretome with only 127 secreted proteins compared to 156 in the WT condition. Furthermore, 30 proteins, identified in the WT condition, were absent from the secretome of the Δ*hfq* mutant (Fig. 5b). We, therefore, wondered whether the proteins lacking from the secretome of the Δ*hfq* mutant were entirely absent from the cellular proteome or simply not secreted, i.e., present and accumulating in the cell. To test this, EXCRETE was applied to examine the endoproteome of the the Δ*hfq* mutant and compare it to WT endoproteome. 1905 and 2176 proteins were identified in the WT and Δ*hfq* mutant endoproteomes, respectively (Supplementary Data 9, 10). Amongst the endoproteins in the Δ*hfq* mutant, approximately half of the 30 proteins absent from its secretome could be detected with no significant difference in abundance compared to the WT (Supplementary Fig. 6). The lack of secretion, therefore, did not result in intracellular accumulation. Finally, we observed that out of 56 differentially regulated proteins between the Δ*hfq* and the WT conditions, 54 were upregulated in the Δ*hfq* condition (Fig. 5d). Looking at the COG categories of these 54 proteins, functions in cell envelope assembly and maintenance were most prominent (19%) (Supp. Table 2). It has previously been suggested that the Δ*hfq* mutant exhibits a compromised cell envelope [24]. Therefore, an upregulation of proteins involved in cell envelope maintenance may constitute a response to cell envelope stress.

Finally, we sought to identify the function of the stable *Synechocystis* secretome. A total of 126 proteins were identified across all three conditions (Fig. 5b). Interestingly, the secretome in the Δ*hfq* condition almost completely overlapped with the stable secretome, which suggests the loss of *hfq* reduced the secretome to essential functions. Among the stable 126 secreted proteins, the most represented COG category with known function was cell envelope assembly and maintenance (11%) (Fig. 5e). Furthermore, this was the only category enriched in the stable secretome with an 8% increase in comparison to the variable secretome (Fig. 5e). This supports our conclusion that the core function of the cyanobacterial secretome is to assemble and maintain critical external barriers that protect their photoautotrophic lifestyle.

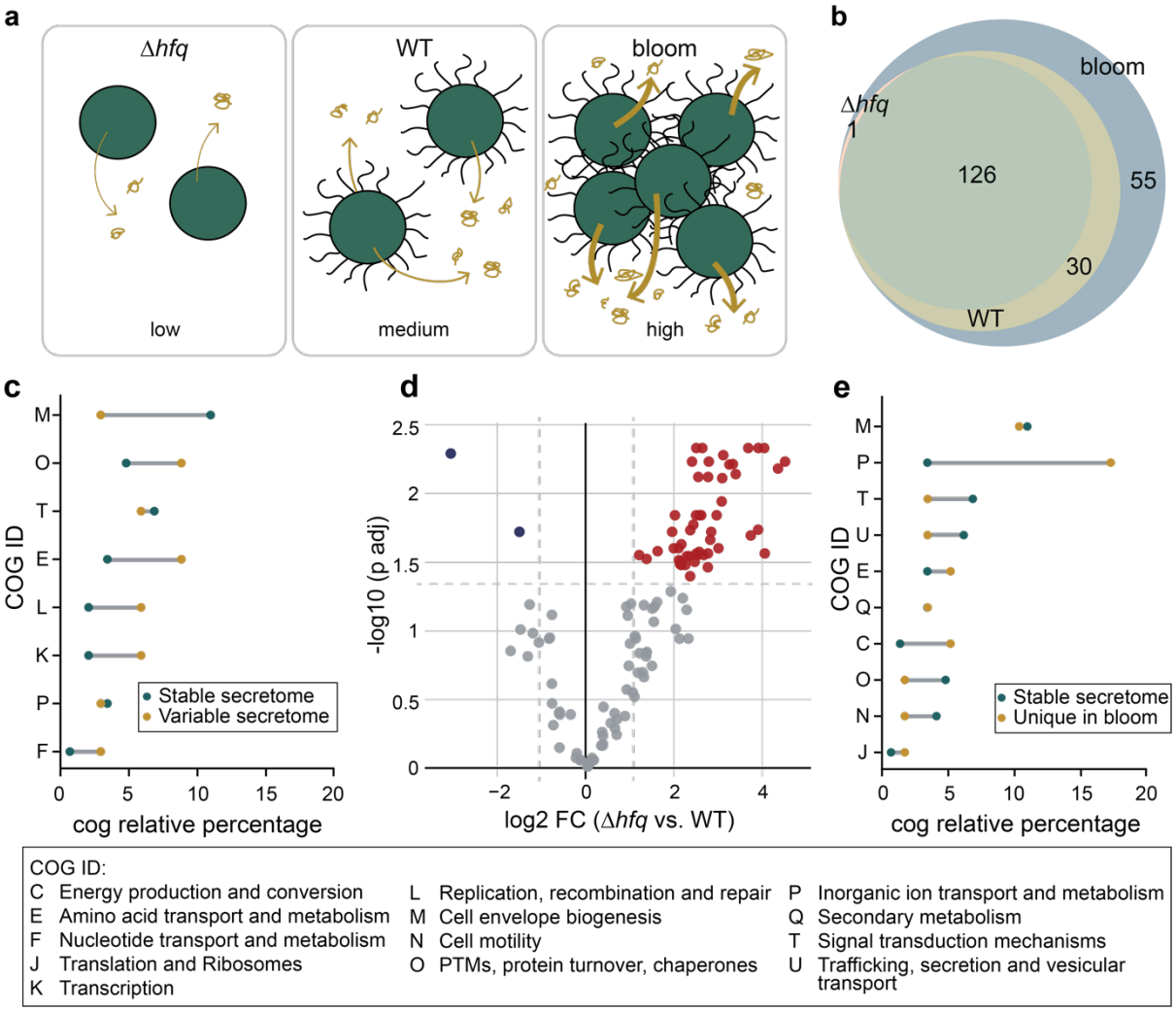

Fig. 5. Secretome comparison between wild-type (WT) *Synechocystis*, a Δ*hfq* mutant and a bloom-like culture **a** Schematic overview of *Synechocystis* secretome comparison resulting in low (Δ*hfq*), medium (WT) and high (bloom-like culture) levels of protein secretion. **b** Venn diagram representing the overlap of protein identifications. **c** Comparison of Clusters of Orthologous Genes (COG) category distribution between all secreted proteins in the stable secretome and those unique to bloom conditions. **d** Volcano plot illustrating differential protein expression between proteins identified in the secretomes of *Synechocystis* in the WT and Δ*hfq* conditions. Each dot on the plot represents an individual protein. A Student's two-sample unpaired t-test with permutation-based multiple test correction was used for differential proteins analysis. The fold-change and significance thresholds are indicated by vertical and horizontal dashed lines, respectively. Color coding indicates significant changes (red for upregulated, blue for downregulated, and grey for non-significant changes). **e** Comparison of COG category distribution between all secreted proteins

in the stable secretome and those in the variable secretome. COG relative percentage represents the number of times a COG ID appears in relation to the total numbers of COG IDs.

Discussion

Cyanobacteria are pioneer organisms that play a key role in shaping global habitats. Therefore, description of the cyanobacterial exoproteome is crucial to understand the mechanisms through which cyanobacteria shape their environment and associated microbiome. In this study we developed EXCRETE, a workflow based on solid-phase enhanced protein aggregation protocols that enabled deep investigation of the cyanobacterial exoproteome. Current methods to interrogate the exoproteome typically start with sample concentration by ultrafiltration or protein precipitation and are often followed by additional clean-up steps. Here, protein losses can occur due to adsorption to the ultrafiltration column, difficulties in washing and resolubilizing the protein pellet following precipitation, and sample degradation due to long processing times and excessive handling. EXCRETE overcomes these issues by concentrating proteins on paramagnetic beads that can be extensively washed with little to no loss of material. This is done in a single container, does not require specialized equipment and samples can be prepared in less than 1 h (Fig. 1a, b). The superior performance of EXCRETE was demonstrated through benchmarking against an ultrafiltration workflow. With EXCRETE we identified 88% more protein groups resulting in 327 unique protein identifications (Fig. 1d, e). Interestingly, these unique identifications included many low abundance proteins, thus increasing the dynamic range (Supplementary Fig. 1b). One aspect where we did not see an improvement was in protein digestion efficiency (Supplementary Fig. 1a). However, this is likely due to the use of a relatively high 1:20 enzyme-to-protein ratio. Previous studies have shown that protein digestion on paramagnetic beads is robust down to a 1:500-1:1000 enzyme-to-protein ratio [34], therefore this could be an area of improvement for future iterations of EXCRETE. We also demonstrated that EXCRETE could be adapted to a 96-well microplate format (Fig. 2). Miniaturization of the workflow will allow processing of hundreds of samples per day, opening the door to large-scale exoproteomic studies. Solid-phase enhanced protein aggregation workflows have already been automatized [55]. Therefore, we expect that our workflow could easily be transferred to a liquid handling robot for high-throughput exoproteome analysis.

Application of EXCRETE to three different cyanobacteria enabled a remarkable increase of up to eight times more exoprotein identifications than previous studies that have used TCA precipitation or ultrafiltration (Fig. 3, Supplementary Fig. 2). However, the identification of such a large number of exoproteins was somewhat unexpected given that cyanobacteria are not typically considered proficient protein secreters [11]. Cells are surrounded by a thick, multilayered envelope with a considerably lower permeability than typical Gram-negative bacteria and porins that only allow the diffusion of inorganic solutes[10]. A COG analysis of secreted proteins pointed towards a specialization in cell envelope management and nutrient acquisition, which, interestingly, was conserved across all species and conditions tested (Figs. 4, 5). These results are not at odds with the traditional view of cyanobacteria. With the ability to synthesize organic compounds themselves, cyanobacteria do not need secretory proteins to hydrolyze organic matter but rather to maintain their protective exterior and facilitate the acquisition of nutrients essential to sustain their photoautotrophic lifestyle [9]. Surprisingly, despite their apparent importance, little is known regarding proteins involved in cell envelope assembly and maintenance in cyanobacteria. Enzymes such as N-acetylmuramoyl-L-alanine amidases, lytic transglycosylases, and peptidyl-prolyl cis-trans isomerases were identified in all secretomes in this work. A recent study in *Anabaena* sp. PCC 7120 also identified several extracellular amidases with

potential roles in peptidoglycan recycling and the formation of nanopores [48]. This suggests that enzymes involved in cell envelope assembly and maintenance are ubiquitous in cyanobacteria and are a promising direction of future studies.

In this study we have greatly expanded the number of identified secreted proteins. However, it remains unclear how hundreds of secreted proteins are transported across a tightly regulated cell envelope. For example, our data support the presence and activity of the one-step T1SS, however putative T1SS substrates only accounted for a small fraction of total secreted proteins (Supplementary Table 1). For two-step secretion in cyanobacteria, it is generally accepted that translocation across the cytoplasmic membrane is mediated by the Sec and Tat pathways [12,62]. Less is known regarding translocation across the outer membrane (OM). In bacteria, the type II secretion system (T2SS) mediates the secretion of a wide variety of proteins. However, cyanobacteria only possess the homologous T4P system which is likely dedicated to the secretion of pilins [24]. Given that the most abundant OM proteins are only permeable to inorganic ions [10], the main protein secretion channels across the OM remain unknown. Protein secretion can also occur via "non-classical secretion". Proteins secreted by these mechanisms lack recognizable signal peptides or secretion motifs and include, for example, cytoplasmic proteins which can exert a second, "moonlighting", function outside the cell [60] and proteins secreted via extracellular vesicles [25,79]. Although non-classical secretion remains poorly characterized in cyanobacteria, a recent study has suggested that extracellular vesicles can be responsible for up to 10% of total protein secretion [25]. Taken together, these results suggest that many open questions remain regarding protein secretion in cyanobacteria.

In summary, our study provides a novel workflow that enables deep exploration of microbial exoproteomes. Application of this approach highlights cyanobacteria as unexpectedly proficient protein secreters and suggests the cyanobacterial secretome provides support to their successful photoautotrophic lifestyle. More broadly, our approach should be applicable to a wide range of bacteria and algae and could be readily adapted to field studies, thus opening new avenues in microbial exoproteomics.

Acknowledgements

The authors thank Elke Dittman for helpful discussions regarding *Nostoc* cultivation. We also thank Tanveer Batth for helpful discussions regarding method development and Felix Schneidmadel for mass spectrometry assistance. The authors acknowledge financial support from the Humboldt Foundation (DAR), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), SFB 1127 ChemBioSys, project number 239748522 (DAR, GP, JAZZ) and the Free state of Thuringia and the European Union via the "Innovationszentrum für Thüringer Medizintechnik-Lösungen" (ThIMEDOP; #2018 IZN 002).

Author Contributions

DAR: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Visualization, Writing – Original draft, Writing – Review & editing, Resources, Funding acquisition. **DO**: Methodology, Investigation, Formal analysis, Data curation, Visualization, Writing – Original draft, Writing – Review & editing. **GP**: Writing – Review & editing, Resources, Funding acquisition. **FM**: Conceptualization, Methodology, Formal analysis, Writing – Review & editing, Resources, Funding acquisition. **JAZZ**: Conceptualization, Methodology, Formal analysis, Visualization, Writing – Original draft, Writing – Review &

editing, Resources, Funding acquisition. All authors have read and approved the final version of the manuscript. DAR and DO contributed equally.

Competing interests

The authors declare no competing interests.

Materials & Correspondence

Correspondence and requests for material should be addressed to David A. Russo.

References

1. Phelan, V. V., Liu, W.-T., Pogliano, K. & Dorrestein, P. C. Microbial metabolic exchange—the chemotype-to-phenotype link. *Nat Chem Biol* **8**, 26–35 (2012).
2. Costa, T. R. D. *et al.* Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol* **13**, 343–359 (2015).
3. Longhurst, A., Sathyendranath, S., Platt, T. & Caverhill, C. An estimate of global primary production in the ocean from satellite radiometer data. *J Plankton Res* **17**, 1245–1271 (1995).
4. Flombaum, P. *et al.* Present and future global distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus. *PNAS* **110**, 9824–9829 (2013).
5. Partensky, F., Hess, W. R. & Garczarek, L. Marine Cyanobacteria. in *The Marine Microbiome* (eds. Stal, L. J. & Cretoiu, M. S.) 103–157 (Springer International Publishing, Cham, 2022). doi:10.1007/978-3-030-90383-1_3.
6. Kahn, A., Oliveira, P., Cuau, M. & Leão, P. N. Incorporation, fate, and turnover of free fatty acids in cyanobacteria. *FEMS Microbiol Rev* **47**, fuad015 (2023).
7. De Philippis, R. & Vincenzini, M. Exocellular polysaccharides from cyanobacteria and their possible applications. *FEMS Microbiol Rev* **22**, 151–175 (1998).
8. Prieto-Barajas, C. M., Valencia-Cantero, E. & Santoyo, G. Microbial mat ecosystems: Structure types, functional diversity, and biotechnological application. *Electron J Biotechnol* **31**, 48–56 (2018).
9. Hoiczyk, E. & Hansel, A. Cyanobacterial Cell Walls: News from an Unusual Prokaryotic Envelope. *J Bacteriol* **182**, 1191–1199 (2000).
10. Kowata, H., Tochigi, S., Takahashi, H. & Kojima, S. Outer Membrane Permeability of Cyanobacterium *Synechocystis* sp. Strain PCC 6803: Studies of Passive Diffusion of Small Organic Nutrients Reveal the Absence of Classical Porins and Intrinsically Low Permeability. *J Bacteriol* **199**, e00371-17 (2017).
11. Zhao, Z., Baltar, F. & Herndl, G. J. Linking extracellular enzymes to phylogeny indicates a predominantly particle-associated lifestyle of deep-sea prokaryotes. *Sci Adv* **6**, eaaz4354 (2020).
12. Russo, D. A. & Zedler, J. A. Z. Genomic insights into cyanobacterial protein translocation systems. *Biol Chem* **402**, 39–54 (2021).
13. Moslavac, S. *et al.* A TolC-Like Protein Is Required for Heterocyst Development in Anabaena sp. Strain PCC 7120. *J Bacteriol* **189**, 7887–7895 (2007).
14. Hahn, A., Stevanovic, M., Mirus, O. & Schleiff, E. The TolC-like protein HgdD of the cyanobacterium Anabaena sp. PCC 7120 is involved in secondary metabolite export and antibiotic resistance. *J. Biol. Chem.* **287**, 41126–41138 (2012).
15. Oliveira, P., Pinto, F., Pacheco, C. C., Mota, R. & Tamagnini, P. HesF, an exoprotein required for filament adhesion and aggregation in *Anabaena* sp. PCC 7120. *Environ Microbiol* **17**, 1631–1648 (2015).

16. Oliveira, P. *et al.* The versatile TolC-like Slr1270 in the cyanobacterium *Synechocystis* sp. PCC 6803. *Environ Microbiol* **18**, 486–502 (2016).
17. Bhaya, D., Bianco, N. R., Bryant, D. & Grossman, A. Type IV pilus biogenesis and motility in the cyanobacterium *Synechocystis* sp. PCC 6803. *Mol Microbiol* **37**, 941–951 (2000).
18. Yoshihara, S. *et al.* Mutational Analysis of Genes Involved in Pilus Structure, Motility and Transformation Competency in the Unicellular Motile Cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol* **42**, 63–73 (2001).
19. Schuergers, N. *et al.* Cyanobacteria use micro-optics to sense light direction. *eLife* **5**, e12620 (2016).
20. Conradi, F. D. *et al.* Factors Controlling Floc Formation and Structure in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803. *J Bacteriol* **201**, e00344-19 (2019).
21. Aguilo-Ferretjans, M. del M. *et al.* Pili allow dominant marine cyanobacteria to avoid sinking and evade predation. *Nat Commun* **12**, 1857 (2021).
22. Zedler, J. A. Z., Michel, M., Pohnert, G. & Russo, D. A. Cell surface composition, released polysaccharides, and ionic strength mediate fast sedimentation in the cyanobacterium *Synechococcus elongatus* PCC 7942. *Environ Microbiol* **25**, 1955–1966 (2023).
23. Russo, D. A. *et al.* Expression and secretion of a lytic polysaccharide monooxygenase by a fast-growing cyanobacterium. *Biotechnol Biofuels* **12**, 74 (2019).
24. Russo, D. A. *et al.* Development of a Highly Sensitive Luciferase-Based Reporter System To Study Two-Step Protein Secretion in Cyanobacteria. *J Bacteriol* **204**, e00504-21 (2022).
25. Matinha-Cardoso, J. *et al.* Novel protein carrier system based on cyanobacterial nano-sized extracellular vesicles for application in fish. *Microb Biotechnol* **15**, 2191–2207 (2022).
26. Frenkel, A. *et al.* Cell specialization in cyanobacterial biofilm development revealed by expression of a cell-surface and extracellular matrix protein. *npj Biofilms Microbiomes* **9**, 10 (2023).
27. Armengaud, J., Christie-Oleza, J. A., Clair, G., Malard, V. & Duport, C. Exoproteomics: exploring the world around biological systems. *Expert Rev of Proteomics* **9**, 561–575 (2012).
28. Oliveira, P. *et al.* The *Anabaena* sp. PCC 7120 Exoproteome: Taking a Peek outside the Box. *Life* **5**, 130–163 (2015).
29. Gao, L. *et al.* Profiling and Compositional Analysis of the Exoproteome of *Synechocystis* Sp. PCC 6803. *J Metabolomics Syst Biol* **1**, 8 (2014).
30. Warshan, D. *et al.* Feathermoss and epiphytic *Nostoc* cooperate differently: expanding the spectrum of plant–cyanobacteria symbiosis. *ISME J* **11**, 2821–2833 (2017).
31. Flores, C. *et al.* The alternative sigma factor SigF is a key player in the control of secretion mechanisms in *Synechocystis* sp. PCC 6803. *Environ Microbiol* **21**, 343–359 (2019).
32. Yegorov, Y. *et al.* A Cyanobacterial Component Required for Pilus Biogenesis Affects the Exoproteome. *mBio* **12**, e03674-20 (2021).
33. Hughes, C. S. *et al.* Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol Syst Biol* **10**, 757 (2014).
34. Batth, T. S. *et al.* Protein Aggregation Capture on Microparticles Enables Multipurpose Proteomics Sample Preparation. *Mol Cell Proteomics* **18**, 1027–1035 (2019).
35. Dienst, D. *et al.* The cyanobacterial homologue of the RNA chaperone Hfq is essential for motility of *Synechocystis* sp. PCC 6803. *Microbiology* **154**, 3134–3143 (2008).

36. Stanier, R. Y., Kunisawa, R., Mandel, M. & Cohen-Bazire, G. Purification and properties of unicellular blue-green algae (order Chroococcales). *Bacteriol Rev* **35**, 171–205 (1971).
37. Stevens Jr., S. E., Patterson, C. O. P. & Myers, J. The Production of Hydrogen Peroxide by Blue-Green Algae: A Survey. *J Phycol* **9**, 427–430 (1973).
38. Bähr, L., Wüstenberg, A. & Ehwald, R. Two-tier vessel for photoautotrophic high-density cultures. *J Appl Phycol* **28**, 783–793 (2016).
39. Cao, J., Russo, D. A., Xie, T., Groß, G. A. & Zedler, J. A. Z. A droplet-based microfluidic platform enables high-throughput combinatorial optimization of cyanobacterial cultivation. *Sci Rep* **12**, 15536 (2022).
40. Rappsilber, J., Ishihama, Y. & Mann, M. Stop and Go Extraction Tips for Matrix-Assisted Laser Desorption/Ionization, Nanoelectrospray, and LC/MS Sample Pretreatment in Proteomics. *Anal. Chem.* **75**, 663–670 (2003).
41. Harney, D. J. *et al.* Proteomic Analysis of Human Plasma during Intermittent Fasting. *J Proteome Res* **18**, 2228–2240 (2019).
42. Meier, F. *et al.* diaPASEF: parallel accumulation–serial fragmentation combined with data-independent acquisition. *Nat Methods* **17**, 1229–1236 (2020).
43. Bruderer, R. *et al.* Extending the Limits of Quantitative Proteome Profiling with Data-Independent Acquisition and Application to Acetaminophen-Treated Three-Dimensional Liver Microtissues. *Mol Cell Proteomics* **14**, 1400–1410 (2015).
44. Huerta-Cepas, J. *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**, D309–D314 (2019).
45. Yu, N. Y. *et al.* PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**, 1608–1615 (2010).
46. Teufel, F. *et al.* SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat Biotechnol* **40**, 1023–1025 (2022).
47. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* **51**, D523–D531 (2023).
48. Sarasa-Buisan, C. *et al.* FurC (PerR) contributes to the regulation of peptidoglycan remodeling and intercellular molecular transfer in the cyanobacterium *Anabaena* sp. strain PCC 7120. *mBio* **0**, e03231-23 (2024).
49. Knecht, S., Eberl, H. C. & Bantscheff, M. Interval-Based Secretomics Unravels Acute-Phase Response in Hepatocyte Model Systems. *Mol Cell Proteomics* **21**, 100241 (2022).
50. Moggridge, S., Sorensen, P. H., Morin, G. B. & Hughes, C. S. Extending the Compatibility of the SP3 Paramagnetic Bead Processing Approach for Proteomics. *J Proteome Res* **17**, 1730–1740 (2018).
51. Johnston, H. E. *et al.* Solvent Precipitation SP3 (SP4) Enhances Recovery for Proteomics Sample Preparation without Magnetic Beads. *Anal Chem* **94**, 10320–10328 (2022).
52. Parnasa, R. *et al.* Small secreted proteins enable biofilm development in the cyanobacterium *Synechococcus elongatus*. *Sci Rep* **6**, (2016).
53. Włodarczyk, A., Selão, T. T., Norling, B. & Nixon, P. J. Newly discovered *Synechococcus* sp. PCC 11901 is a robust cyanobacterial strain for high biomass production. *Commun Biol* **3**, 1–14 (2020).
54. Sielaff, M. *et al.* Evaluation of FASP, SP3, and iST Protocols for Proteomic Sample Preparation in the Low Microgram Range. *J Proteome Res* **16**, 4060–4072 (2017).
55. Müller, T. *et al.* Automated sample preparation with SP3 for low-input clinical proteomics. *Mol Syst Biol* **16**, e9111 (2020).

56. Meeks, J. C. & Elhai, J. Regulation of Cellular Differentiation in Filamentous Cyanobacteria in Free-Living and Plant-Associated Symbiotic Growth States. *Microbiol Mol Biol Rev* **66**, 94–121 (2002).
57. Risser, D. D. Hormogonium Development and Motility in Filamentous Cyanobacteria. *Appl Environ Microbiol* **89**, e00392-23 (2023).
58. Schuergers, N., Mullineaux, C. W. & Wilde, A. Cyanobacteria in motion. *Curr Opin Plant Biol* **37**, 109–115 (2017).
59. Wang, G. *et al.* How are the Non-classically Secreted Bacterial Proteins Released into the Extracellular Milieu? *Curr Microbiol* **67**, 688–695 (2013).
60. Ebner, P. & Götz, F. Bacterial Excretion of Cytoplasmic Proteins (ECP): Occurrence, Mechanism, and Function. *Trends Microbiol* **27**, 176–187 (2019).
61. Tsirigotaki, A., De Geyter, J., Šoštaric´, N., Economou, A. & Karamanou, S. Protein export through the bacterial Sec pathway. *Nat Rev Microbiol* **15**, 21–36 (2017).
62. Frain, K. M., Gangl, D., Jones, A., Zedler, J. A. Z. & Robinson, C. Protein translocation and thylakoid biogenesis in cyanobacteria. *Biochim Biophys Acta, Bioenerg* **1857**, 266–273 (2016).
63. Vries, S. de & Vries, J. de. Evolutionary genomic insights into cyanobacterial symbioses in plants. *Quant Plant Biol* **3**, e16 (2022).
64. Baunach, M., Guljamow, A., Miguel-Gordo, M. & Dittmann, E. Harnessing the potential: advances in cyanobacterial natural product research and biotechnology. *Nat. Prod. Rep.* (2023) doi:10.1039/D3NP00045A.
65. Battchikova, N., Muth-Pawlak, D. & Aro, E.-M. Proteomics of cyanobacteria: current horizons. *Curr Opin Biotechnol* **54**, 65–71 (2018).
66. Álvarez, C., Brenes-Álvarez, M., Molina-Heredia, F. P. & Mariscal, V. Quantitative Proteomics at Early Stages of the Symbiotic Interaction Between *Oryza sativa* and *Nostoc punctiforme* Reveals Novel Proteins Involved in the Symbiotic Crosstalk. *Plant Cell Physiol* **63**, 1433–1445 (2022).
67. Schuergers, N. & Wilde, A. Appendages of the Cyanobacterial Cell. *Life* **5**, 700–715 (2015).
68. Khayatan, B., Meeks, J. C. & Risser, D. D. Evidence that a modified type IV pilus-like system powers gliding motility and polysaccharide secretion in filamentous cyanobacteria. *Mol Microbiol* **98**, 1021–1036 (2015).
69. Oeser, S. *et al.* Minor pilins are involved in motility and natural competence of the cyanobacterium *Synechocystis* sp. PCC 6803. *Mol Microbiol* **116**, 1–23 (2021).
70. Yoshimura, H. *et al.* CccS and CccP are Involved in Construction of Cell Surface Components in the Cyanobacterium *Synechocystis* sp. strain PCC 6803. *Plant Cell Physiol* **51**, 1163–1172 (2010).
71. Giner-Lamia, J. *et al.* Extracellular Proteins: Novel Key Components of Metal Resistance in Cyanobacteria? *Front Microbiol* **7**, 878 (2016).
72. Badarau, A. *et al.* FutA2 Is a Ferric Binding Protein from *Synechocystis* PCC 6803. *J Biol Chem* **283**, 12520–12527 (2008).
73. Sebastian, M. & Ammerman, J. W. The alkaline phosphatase PhoX is more widely distributed in marine bacteria than the classical PhoA. *ISME J* **3**, 563–572 (2009).
74. Holland, I. B. *et al.* Type I Protein Secretion—Deceptively Simple yet with a Wide Range of Mechanistic Variability across the Family. *EcoSal Plus* **7**, (2016).
75. Tsolis, K. C. *et al.* Secretome Dynamics in a Gram-Positive Bacterial Model. *Mol Cell Proteomics* **18**, 423–436 (2019).
76. Schuergers, N. *et al.* Binding of the RNA chaperone Hfq to the type IV pilus base is crucial for its function in *Synechocystis* sp. PCC 6803. *Mol Microbiol* **92**, 840–852 (2014).

77. Stuart, R. K. *et al.* Light Regimes Shape Utilization of Extracellular Organic C and N in a Cyanobacterial Biofilm. *mBio* **7**, e00650-00616 (2016).
78. Kamennaya, N. A. *et al.* High pCO2-induced exopolysaccharide-rich ballasted aggregates of planktonic cyanobacteria could explain Paleoproterozoic carbon burial. *Nat Commun* **9**, 2116 (2018).
79. Lima, S. *et al.* Extracellular vesicles as an alternative copper-secretion mechanism in bacteria. *J Hazard Mater* **431**, 128594 (2022).