Characterizing metabolic stress-induced phenotypes of *Synechocystis* PCC6803 with Raman spectroscopy

Imen Tanniche[1], Eva Collakova[2], Cynthia Denbow[2] and Ryan S. Senger[1,3]

[1] Department of Biological Systems Engineering, Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, United States of America
[2] School of Plant & Environmental Sciences, Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, United States of America
[3] Department of Chemical Engineering, Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, United States of America

ABSTRACT

Background. During their long evolution, *Synechocystis* sp. PCC6803 developed a remarkable capacity to acclimate to diverse environmental conditions. In this study, Raman spectroscopy and Raman chemometrics tools (Rametrix^{TM}) were employed to investigate the phenotypic changes in response to external stressors and correlate specific Raman bands with their corresponding biomolecules determined with widely used analytical methods.

Methods. *Synechocystis* cells were grown in the presence of (i) acetate (7.5–30 mM), (ii) NaCl (50–150 mM) and (iii) limiting levels of $MgSO_4$ (0–62.5 mM) in BG-11 media. Principal component analysis (PCA) and discriminant analysis of PCs (DAPC) were performed with the Rametrix^{TM} LITE Toolbox for MATLAB®. Next, validation of these models was realized via Rametrix^{TM} PRO Toolbox where prediction of accuracy, sensitivity, and specificity for an unknown Raman spectrum was calculated. These analyses were coupled with statistical tests (ANOVA and pairwise comparison) to determine statistically significant changes in the phenotypic responses. Finally, amino acid and fatty acid levels were measured with well-established analytical methods. The obtained data were correlated with previously established Raman bands assigned to these biomolecules.

Results. Distinguishable clusters representative of phenotypic responses were observed based on the external stimuli (i.e., acetate, NaCl, $MgSO_4$, and controls grown on BG-11 medium) or its concentration when analyzing separately. For all these cases, Rametrix^{TM} PRO was able to predict efficiently the corresponding concentration in the culture media for an unknown Raman spectra with accuracy, sensitivity and specificity exceeding random chance. Finally, correlations ($R > 0.7$) were observed for all amino acids and fatty acids between well-established analytical methods and Raman bands.

Corresponding author
Ryan S. Senger, senger@vt.edu

INTRODUCTION

Microbial cells undergo phenotypic variations in response to changing environmental conditions. These changes directly affect the chemical composition of the cell, allowing it to maintain integrity, and the study of these changes is referred to as "microbial phenotyping" (*Tummala et al., 2003*; *Sellick & Reece, 2005*; *Alsaker, Paredes & Papoutsakis, 2010*; *Nicolaou, Gaida & Papoutsakis, 2010*). For example, exposure to high salt concentrations (e.g., NaCl, $MgSO_4$) results in ionic and osmotic imbalances (*Markovitz & Sylvan, 1962*; *Lusk, Williams & Kennedy, 1968*; *Reed et al., 1985*; *Mikkat et al., 2000*; *Karandashova & Elanskaya, 2005*; *Marin et al., 2006*). This leads to the accumulation of carboxylic acids (e.g. acetate), which can be toxic and require a reorganization of the cell membrane for survival (*Papoutsakis et al., 1987*; *Russell, 1992*; *Lasko et al., 1997*; *Roe et al., 2002*). Thus, cells have the ability to respond to such chemical stresses, adapt for survival, and even acquire tolerance (*Nicolaou, Gaida & Papoutsakis, 2010*). These characteristics are necessary in biotechnology, where microbes are engineered to produce products that are toxic in high concentrations (e.g., biofuels), operate in stressful environments (e.g., high salt, low pH), or perform bioremediation. Here, Raman spectroscopy was further developed as a microbial phenotyping methodology, with the focus on analyzing responses of cyanobacteria to salt (NaCl and $MgSO_4$) and acid (acetate) stressors.

In this context, the cyanobacterium *Synechocystis* sp. PCC6803 (hereafter *Synechocystis*) is of interest to biotechnology because of its ability to grow on low-cost resources (i.e., CO_2 and sunlight) and produce high-value chemicals (*Yu et al., 2013*). This species demonstrates a highly versatile carbon metabolism, growing in autotrophic as well as under heterotrophic and mixotrophic conditions (*Vermaas, 1996*). It is also a model organism for many cyanobacteria, which are, in general, known to tolerate high salt concentrations (*Reed et al., 1985*). Acclimation of *Synechocystis* to high salt concentrations begins with adjustment of osmotic equilibrium with the outside medium (*Marin et al., 2004*; *Marin et al., 2006*; *Karandashova & Elanskaya, 2005*), followed by the accumulation of the osmoprotectant glycosylglycerol (*Reed et al., 1986*; *Marin et al., 2004*; *Marin et al., 2006*; *Hagemann, 2013*). This osmoprotectant enhances the internal osmotic potential and protects proteins and membranes (*Borges et al., 2002*; *Hincha & Hagemann, 2004*). In addition, it has been shown that long-term salt stress induces accumulation of specific proteins (*Hagemann, Techel & Rensing, 1991*; *Hagemann & Erdmann, 1994*; *Fulda et al., 2000*; *Huang et al., 2006*) and alterations in fatty acids composition (*Huflejt et al., 1990*; *Khomutov et al., 1990*; *Allakhverdiev et al., 1999*; *Singh, Sinha & Häder, 2002*).

Synechocystis has the ability to grow mixotrophically in the presence of glucose and acetate (*Wu, Shen & Wu, 2002*); however, utilization of acetate as the sole carbon source is limited as it lacks major enzymes required for uptake (*Knoop et al., 2013*; *Varman et al., 2013*; *Thiel et al., 2017*). Acetate could serve as a supplementary precursor for the synthesis of certain biomass building blocks such as amino acids (*Varman et al., 2013*; *Thiel et al., 2017*) and lipids (*Thiel et al., 2017*). Mechanisms of acetate toxicity have been studied mainly in Gram-negative bacteria and include alterations in cell membranes, active secretion, and activation of stress responses (*Isken & De Bont, 1998*; *Sardessai & Bhosle,*

2002). Magnesium, on the other hand, is an important cofactor for activities of many enzymes, while sulfur is a major macronutrient (0.6% of biomass dry weight) (*Schwarz & Forchhammer, 2005*; *Kim et al., 2011*). It is incorporated in proteins, lipids, and other vital compounds, and it is needed for the thioredoxin-mediated regulation of enzymes involved in photosynthesis and respiratory activities (*Zhang et al., 2008*; *Kiyota, Ikeuchi & Hirai, 2012*).

Commonly, biological studies of physiological responses to chemical stressors have taken place using genomic methods (*Fulda et al., 2000*; *Marin et al., 2004*; *Karandashova & Elanskaya, 2005*; *Huang et al., 2006*; *Zhang et al., 2008*) to study gene expression and biochemical approaches (*Marin et al., 2006*; *Zavřel, Očenášová & Červený, 2017*) to characterize biomolecules. These require extensive sample preparation and long analysis times. Near real-time methods of analysis are needed alongside standardized procedures to provide rapid information to further understand microbial metabolic responses to environmental stress or for phenotype screening purposes. With such methods, cultures could be monitored in near real-time in biotechnological applications so that the appearance of unhealthy or unproductive phenotypes could be remedied quickly.

Raman spectroscopy has been proven a powerful analytical technique for the analysis of biological materials. It uses monochromatic light to provide a qualitative measurement of the biochemical composition of the biological specimen (*Butler et al., 2016*) and gives distinctive signals from macromolecules (i.e., proteins, lipids, carbohydrates, and nucleic acids) (*Das & Agrawal, 2011*). Therefore, a characteristic fingerprint is assigned to each biological sample and allows a better understanding of its chemical composition. Raman spectroscopy allows near real-time and noninvasive acquisition of phenotype data from living systems (*Das & Agrawal, 2011*; *Zu et al., 2014*; *Butler et al., 2016*; *Freedman et al., 2016*; *Zu, Athamneh & Senger, 2016*; *García-Timermans et al., 2019*). In addition, the analysis is label free, requires minimal or no sample preparation, and there is no spectral interference from water (*Zu et al., 2014*; *Freedman et al., 2016*; *Zu, Athamneh & Senger, 2016*). The application of Raman spectroscopy is an expanding field and it includes fermentation monitoring (*Sivakesava, Irudayaraj & Demirci, 2001*; *Ewanick et al., 2013*; *Olson et al., 2016*; *Zu et al., 2016*), detection and identification of microorganisms (*Nelson, Manoharan & Sperry, 1992*; *Kirschner et al., 2001*; *Pahlow et al., 2015*), monitoring the kinetics of germination of individual *Clostridium difficile* spores (*Wang et al., 2015*) and detection of its toxins (*Koya et al., 2018*). It has been also demonstrated that Raman spectroscopy can be used in near-real time phenotyping of *Escherichia coli* exposed to alcohol (*Zu et al., 2014*; *Zu, Athamneh & Senger, 2016*) and antibiotics (*Athamneh et al., 2014*), single cell phenotyping (*Wu et al., 2011*; *Serrano et al., 2014*; *Sun et al., 2015*; *García-Timermans et al., 2019*), and characterizing phenotypic differences among *E. coli* enriched for 1-butanol tolerance (*Freedman et al., 2016*; *Zu et al., 2014*). Thus, Raman spectroscopy was selected as an effective method for probing changes in *Synechocystis* phenotypes in near real-time when exposed to external stimuli.

In this research, Raman spectroscopy was used to perform microbial phenotyping of *Synechocystis* cultures growing in different concentrations of (i) acetic acid (7.5–30 mM), (ii) NaCl (50–150 mM), and (iii) $MgSO_4$ (0–62.5 mM) in BG-11 media. This was done to test

the hypothesis that Raman spectroscopy and spectral processing with Raman chemometrics (Rametrix[TM]) and statistical analyses could distinguish among phenotypes and quantify changes in biomolecular composition. Routine analytical methodology was used to measure the levels of several biomolecules (e.g., fatty acids and amino acids) so correlations could be made with individual Raman band intensities. This technique was shown previously (*Zu et al., 2014*) to correlate Raman data with fatty acid and amino acid analytical data for *Escherichia coli*. Here, this approach was expanded to include *Synechocystis* and the stress-induced phenotypes studied. *Synechocystis* was chosen for this study because it is a well-studied model organism for many cyanobacteria. It has well-developed genomic tools, relevance to biotechnology, and stress-response and chemical composition data available in the literature.

MATERIAL AND METHODS

Bacterial strain and culture conditions

Synechocystis sp. PCC6803 (ATCC$^{®}$27184^{TM}) were used to create axenic kanamycin resistant mutants for the comparative growth experiments (*Heidorn et al., 2011*; Tanniche, Collakova & Senger, 2019, unpublished data). This strain was used with kanamycin (15 μg/mL) to help prevent contamination. Cells were cultivated in the BG-11 medium (*Rippka et al., 1978*) (adjusted to pH 7.0) supplemented with 5 mM glucose. Three different stress conditions were analyzed: (i) NaCl (50, 100, and 150 mM), (ii) acetate (7.5, 15, and 30 mM) and (iii) MgSO$_4$ (0, 15.6, 31.2, and 62.5 mM; representing 0%, 25%, 50% and 100% of BG-11). Cultures grown in BG-11 were used as controls, and cells harvested in mid-exponential phase were used to initiate stress experiments. Each culture type condition experiment was performed in biological triplicate. Cultures were grown in 50 mL volume in 125 mL flasks, incubated at 25 °C under continuous light (20 μE), agitated at 140 rpm, and ambient CO$_2$. Cells were harvested at mid-exponential phase and washed twice with ice-cold purified water. 2 μL of cells were kept for Raman scans and the remaining pellet was freeze-dried and prepared for analyses by UPLC and GC-FID.

Raman spectroscopy

To prepare samples for Raman analysis, 2 μL of washed cells were dried at room temperature on aluminum foil (three samples per biological replicate). Dried cells were analyzed using a PeakSeeker PRO-785 Raman microscope (Agiltron; Woburn, MA) with 10X objective. Measurements were carried out using the following settings: laser excitation of 785 nm (30 mW) for 5 seconds with spectral resolution of 13 cm^{-1}. Twenty individual scans were taken per sample by focusing on different areas of the dried cells. Raman data were collected using RSIQ^{TM} software.

Computational methods

Acquired Raman spectra were processed and analyzed using the Rametrix^{TM} LITE Toolbox (*Fisher et al., 2018*) and Rametrix^{TM} PRO Toolbox (*Senger & Robertson, 2020*) for MATLAB$^{®}$. MATLAB$^{®}$ R1018A with the Statistics and Machine Learning Toolbox was used for all calculations. Data analysis with Rametrix^{TM} LITE Toolbox (*Fisher et al.,*

2018) consisted of averaging the 20 spectra replicates and then baselining with the Goldindec algorithm (*Liu et al., 2015*) (baseline polynomial order = 3, estimated peak ratio = 0.5, and smoothing window size = 5) and vector normalizing over the biological range (400 cm^{-1}–1,800 cm^{-1}). PCA was performed, and outlier spectra were identified and excluded from analysis. Next, DAPC was applied, and several DAPC models were constructed by varying the number of principal components (PCs) used to build each one. DAPC models enabled the separation of Raman spectra according to a preset experimental factor. Here, the culture medium composition and environmental stimuli (acetate, NaCl, and $MgSO_4$) concentrations were used. The Rametrix^{TM} PRO Toolbox (*Senger & Robertson, 2020*) was used to apply leave-one-out analysis to DAPC clustering. This analysis determines the ability of DAPC models to predict correctly the classification when presented with an unknown Raman spectrum. The performance of DAPC models are quantitated by prediction accuracy (percent of unknown spectra classified correctly), sensitivity (true positive percentage), and specificity (true negative percentage).

Analysis of variance (ANOVA), pairwise comparisons using Tukey's honest significant difference (HSD), and regression statistical analyses required the representation of the Rametrix^{TM} LITE outputs (spectra intensity values, PCA, and DAPC) as a single numerical value. For example, the hundreds of intensity values at each Raman shift (400–1,800 cm^{-1}) had to be reduced to a single numerical value for each spectrum. This is also true as data-rich spectra were reduced from hundreds of intensity values to tens of PCs by PCA and canonical values by DAPC. To calculate this single numerical value for statistical comparisons between spectra, the distance formula was applied to determine the similarity between two spectra, where one was a reference. When using PCA data, the Total Principal Component Distance (TPD) was calculated according to Eq. (1), where the first five PCs (representing more than 95% of the dataset variance) were used.

$$TPD = \sum_{i=1}^{5} \sqrt{\left(P_{x,i} - P_{reference,i}\right)^2}. \tag{1}$$

Here, $P_{x,i}$ is the *i*th PC of spectrum *x* and $P_{reference,i}$ represents the *i*th PC of the reference. The calculation was repeated for all *x* spectra in the analysis. Different spectra were chosen as the reference in each of the four sub-studies presented in this research. These are identified for each in Appendix S1 in the Raw Distances and Mean Distances sheets. More details about TPD and its value in statistical calculations have been published (*Senger et al., 2019*). Similarly, the distance calculation was applied to each spectrum to calculate the Total Spectral Distance (TSD), based on normalized intensity values across all Raman shifts, and the Total Canonical Distance (TCD), based on the first five canonicals in the DAPC model. Reduction of spectral data into a single value allowed statistical analysis including analysis of variance (ANOVA), pairwise comparisons using Tukey's honest significant difference (HSD), and regression. These analyses were performed to assess whether significant differences exist among the separated groups of spectra (ANOVA), to determine which group differences are statistically significant (pairwise comparisons), and establish a linear fit among data.

Biomass measurements

Fatty acid content was determined by gas chromatography-flame ionization detection (GC-FID) after direct hydrolysis and conversion to fatty acid methyl esters (FAME). Lyophilized cells were incubated with methanolic HCl and 10 µg heptadecanoic acid (C17:0) as internal standard at 75 °C for 2 h to form FAME. GC-FID was done on an Agilent 7890A series GC equipped with FID. For fatty acid identification purposes, the same GC was coupled to a 5,975C series single quadrupole mass spectrometer (MS) and fatty acids were identified by comparing the spectra with the corresponding spectra in the NIST library (Agilent Technologies, Santa Clara, CA).

Protein hydrolysis was carried out in a custom-made teflon hydrolysis chamber (see photo in Appendix S2) under vacuum at 110 °C in the presence of 6N HCl. 50 mM norvaline was used as an internal standard. Amino acid levels were determined after derivatization with AccQ-Tag™ reagent using a Waters Acquity™H-class ultra-performance liquid chromatography (UPLC) system equipped with a fluorescent detector (Waters Corporation, Mildford, MA).

Public access

The Rametrix™ LITE Toolbox is shared through GitHub with license agreement (https://github.com/SengerLab/RametrixLITEToolbox), and the Rametrix™ PRO Toolbox is also accessible through GitHub with license agreement (https://github.com/SengerLab/RametrixPROToolbox). Raw Raman spectra (in *.SPC file format) and biological measurement raw data are shared in the Supplemental Information 3.

RESULTS

Raman spectroscopy of *Synechocystis* grown in the presence of acetate and salts

Raman spectroscopy was used to identify phenotypic changes in *Synechocystis* induced by the presence of: (i) acetate, (ii) NaCl, and (iii) $MgSO_4$. Cells were scanned with Raman microscopy and the resulting spectra were analyzed using the Rametrix™ LITE (*Fisher et al., 2018*) and PRO (*Senger & Robertson, 2020*) Toolboxes. The obtained spectra were averaged, truncated (400–1,800 cm^{-1}), baselined, and vector normalized prior to further processing using PCA and DAPC in the Rametrix™ LITE. Rametrix™ PRO was used to determine the ability to predict a particular metabolic phenotype given a Raman spectrum of cells grown under unknown growth conditions. Next, statistical analyses (i.e., ANOVA and pairwise comparisons) were performed on TCD and TPD results. TPD and TSD produced similar results, allowing TSD to be omitted in the analysis. Finally, specific Raman bands assigned to certain metabolites (i.e., amino acids and fatty acids) were compared to standardized analytical methods (i.e., UPLC, GC-FID). The remainder of the results section is divided according to analysis type: (i) Rametrix™ analyses and (ii) correlation of individual Raman bands with well-established analytical approaches.
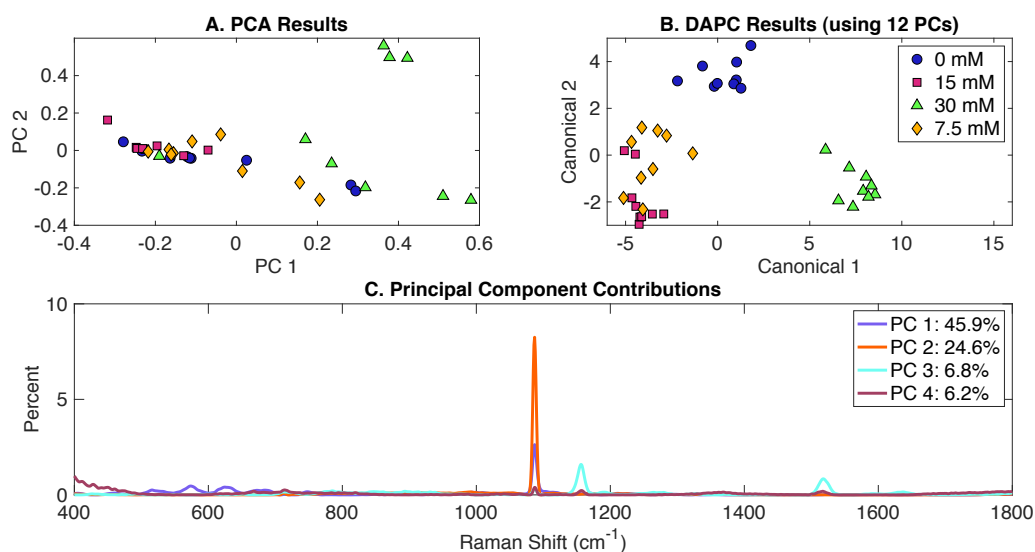
Figure 1 Rametrix^TM LITE analysis of acetate-induced phenotypes. (A) PCA, (B) DAPC, and (C) Raman shifts contributions between groups in PCA. 0 mM represents native BG-11 medium.

Rametrix^TM analyses
Acetate-induced phenotypes

Raman spectra of cultures growing in three concentrations of acetate (7.5 mM, 15 mM and 30 mM) and the control (BG-11 without acetate) were analyzed by Raman spectroscopy and the Rametrix^TM LITE and Rametrix^TM PRO Toolboxes. PCA, DAPC, and PC contributions results are shown in Fig. 1. PCA of Raman spectra (Fig. 1A) showed some separation of clusters in the two first PCs, which contained more than 68% of the dataset variance. This clustering is more noticeable with the DAPC model (Fig. 1B) constructed with 12 PCs (representing more than 99% of the dataset variance). Although four marked clusters were revealed, based on the concentration of acetate in the BG-11 medium, only three major groups were observed. These clusters suggest that phenotypic differences could be generated as a result of acetate addition to BG-11 medium and that cells grown in the presence of 7.5 mM and 15 mM could have similar phenotypic responses. Next, Raman shift contributions for both PCA and DAPC were studied to point to molecular differences between clusters of Raman spectra. In this particular case, it provided information about the molecular differences between cultures grown in the presence and absence of acetate. It does not provide information about accumulation or depletion of individual biomolecules (e.g., carotenoids), but it does suggest the associated band intensities change significantly. Raman shift contributions between groups in PCA are shown in Fig. 1C and the contributions associated with DAPC model are represented in Fig. S1. Raman band assignments were selected based on published libraries (*De Gelder et al., 2007*; *Movasaghi, Rehman & Rehman, 2007*; *Zhu et al., 2011*). The full list of assignments is given in Appendix S1, and highlights are given in Table 1.

Table 1	Highlighted Raman shift contributions to PCA and DAPC models.	
Study	**Model**	**Biomolecules and Raman bands (cm^{-1})**
Acetate-induced	PCA	Glycerol (630 cm^{-1}), amino acid (573 cm^{-1}) protein (1,156 cm^{-1}), and carotenoids (1,156, 1,518 cm^{-1})
Acetate-induced	DAPC	Proteins (524, 618, 1,005, 1,155 cm^{-1}), carotenoids (1,155, 1,528 cm^{-1}), amino acids (532, 638, 715, 920, 988, 1,005, 1,196 1,327, 1,495, 1,550, 1,558, 1,615 cm^{-1}), phosphatidylinositol (415 cm^{-1}), polysaccharides (477 cm^{-1}), DNA/RNA (746, 1,510 cm^{-1}), and lipids (1,086 cm^{-1})
NaCl-induced	PCA	Glycerol (630 cm^{-1}); proteins, amides and lipids (1,260 cm^{-1}); carotenoids (1,157, 1,520 cm^{-1}); and amino acids (515, 573, 669, 684, 1,188 cm^{-1})
NaCl-induced	DAPC	Sterols (430 cm^{-1}), phosphatidylinositol (594 cm^{-1}), DNA/RNA (746 cm^{-1}), DNA (1,090, 1,329 cm^{-1}), amino acids (730, 854, 1,040, 1,495 cm^{-1}), carotenoids (1,157, 1,517 cm^{-1}), nucleic acids (650, 1,200, 1,329 cm^{-1}), and lipids (1,255 cm^{-1})
MgSO$_4$-induced	PCA	Sterols (429 cm^{-1}), glycerol (630 cm^{-1}), carotenoids and carotene (1,157, 1,520 cm^{-1}), amino acids (573, 684, 1,188 cm^{-1}), lipids (1,380 cm^{-1}), and phospholipids (1,330 cm^{-1})
MgSO$_4$-induced	DAPC	Carotenoids (1,157, 1,520 cm^{-1}) and amino acids (850, 992 cm^{-1}) in addition to phosphatidylinositol (412 cm^{-1}) and saccharides (850, 868, 1,456 cm^{-1})

Next, TCD, TPD and TSD (both raw and mean values) were calculated using PCA and DAPC data and used in ANOVA and pairwise comparisons. Full results are given in the Appendix S1. In general, as the phenotype deviates from the reference (represented here by phenotypes of cells grown on BG-11 medium), the TCD (or the TPD) mean value should become larger. Similar phenotypes should have comparable values. In the acetate-induced phenotypes, TCD and TPD mean values increased proportionally to the increase of acetate concentration. Based on TCD data (using 12 PCs in DAPC model) ANOVA results indicated that differences between the groups are significant ($p < 0.001$). With TPD data, ANOVA returned a significant p-value ($p < 0.001$). This result indicates that even though no clear clustering was observed in PCA (Fig. 1A), there are statistically significant differences between the groups. Pairwise comparisons identified all phenotypes (groups) to be significantly different based on the TCD data ($p < 0.001$). However, using TPD data, 7.5 mM and 15 mM induced phenotypes were found not significantly different ($p = 0.997$). The small difference in mean TPD and the near clustering observed in Fig. 1B are consistent with this result. Next, regression (Appendix S1) was computed between acetate concentrations (0 mM, 7.5 mM, 15 mM, 30 mM) and the TCD and TPD mean values, respectively. Results show a linear relationship among the data with a coefficient of determination (R^2) of 0.97 and 0.84 for TPD and TCD, respectively.

RametrixTM PRO was used to apply leave-one-out analysis to the DAPC clustering in Fig. 1B. The purpose of this analysis is to determine the ability of the DAPC model to predict precisely the classification (i.e., 0 mM, 7.5 mM, 15 mM, 30 mM of acetate) when

Table 2   Highlighted Rametrix™ PRO results.

Study	Classification predicted	Sensitivity	Specificity
All Studies	Random chance[**]	25%	75%
Acetate-induced	0 mM	100%	89%
Acetate-induced	7.5 mM	67%	81%
Acetate-induced	15 mM	56%	100%
Acetate-induced	30 mM	56%	96%
NaCl-induced	0 mM	100%	100%
NaCl-induced	50 mM	33%	96%
NaCl-induced	100 mM	44%	100%
NaCl-induced	150 mM	44%	89%
MgSO$_4$-induced	62.5 mM	100%	100%
MgSO$_4$-induced	31.2 mM	89%	96%
MgSO$_4$-induced	15.6 mM	100%	89%
MgSO$_4$-induced	0 mM	56%	78%
All phenotypes	BG-11 control	100%	100%
All phenotypes	Acetate-induced	56%	100%
All phenotypes	NaCl-induced	71%	95%
All phenotypes	MgSO$_4$-induced	93%	100%

Notes.

[*]All DAPC models were built with 12 PCs representing 99% of the dataset variance.

[**]The random chance sensitivity and specificity values were calculated statistically, not by Rametrix™ PRO. They were the same for all studies here.

presented with a Raman spectrum of *Synechocystis* cells grown in an unknown condition. Rametrix™ PRO results are also given in Appendix S1, and more details of Rametrix™ PRO calculations are given in *Senger & Robertson (2020)*. Here, four classification options, 0 mM, 7.5 mM, 15 mM and 30 mM, were assessed. With four different classifications, the random chance specificity was calculated as 25% and the random chance specificity as 75% for this dataset. Therefore, model performance sensitivity and specificity must be higher than these values if the model is truly able to predict *Synechocystis* phenotypes based on Raman spectra. Highlighted Rametrix™ PRO results are given in Table 2, and all sensitivity and specificity values exceeded random chance values, with specificity and sensitivity values ranging between 56–100% and 81–100%, respectively.

NaCl-induced metabolic phenotypes

In order to determine the effect of NaCl addition to BG-11 medium on *Synechocystis* metabolic phenotypes, cells were grown in BG-11 at three salt concentrations (50 mM, 100 mM, and 150 mM) and compared to a control (cells grown in BG-11 alone with no NaCl added). PCA, DAPC and PC contribution results generated by the Rametrix™ LITE are represented in Fig. 2. PCA of Raman spectra (Fig. 2A), again, showed no apparent clustering in the first two PCs, which comprised about 69% of the dataset variance. However, the application of DAPC (Fig. 2B) demonstrated four clusters representative of the three different concentrations of NaCl and the control. These results were obtained using 12 PCs (representing over 99% of the dataset variance). Based on the presence of
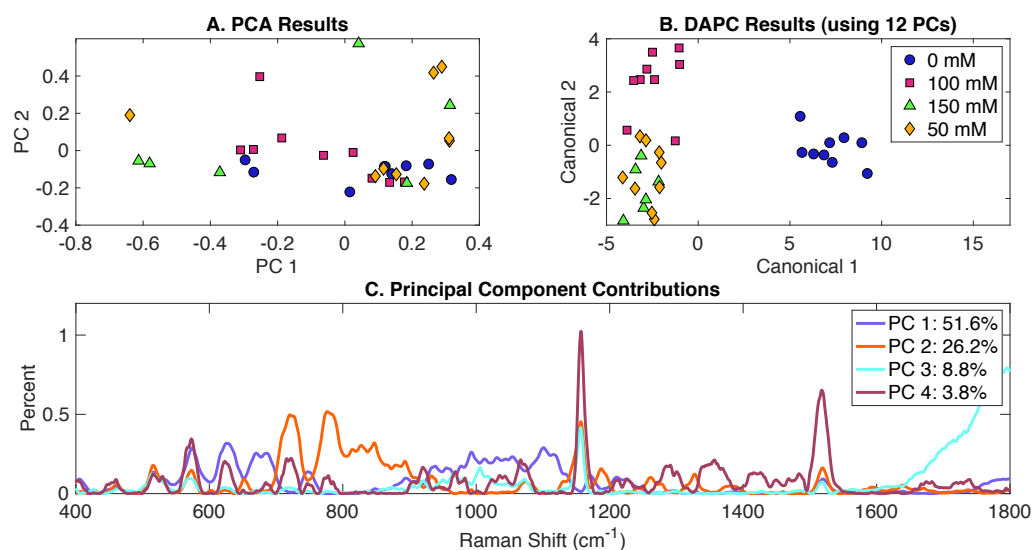
Figure 2 **Rametrix™ LITE analysis of NaCl-induced phenotypes.** Rametrix™ LITE analysis of NaCl-induced phenotypes. (A) PCA, (B) DAPC, and (C) Raman shifts contributions between groups in PCA. 0 mM represents native BG-11 medium.

NaCl, two major groups were noticed (Fig. 2B), suggesting significant cellular phenotype differences could occur in response to as little as 50 mM NaCl in BG-11 medium. Next, Raman shift contributions to the PCA and DAPC results were examined. Raman shift contributions between the groups in PCA are shown in Fig. 2C, and those related to DAPC contributions are represented in Fig. S2. Specific lists of Raman band contributions and molecular assignments (for both PCA and DAPC) are given in Appendix S1, and highlights are given in Table 1.

DAPC model and PCA were used to compute TCD, TPD, TSD, and their mean values. Multi-way ANOVA and pairwise comparisons were used to assess these parameters, and full results are given in Appendix S1. As mentioned previously, the TCD and TPD mean values increase with the increased difference from the reference phenotype. Using the TPD mean values, as the concentration of NaCl increased, the distance also increased. However, a small difference was observed between the 50 mM and the 100 mM TPD means, suggesting comparable phenotypes. The TCD mean values and the clustering in the Fig. 2B confirm this result. ANOVA revealed that molecular phenotypes induced at different NaCl concentrations were significantly different ($p < 0.001$) for TCD and TPD data. Pairwise comparison allowed the identification of significantly different phenotypes. Based on the TPD data, the phenotypic responses associated with 100 mM NaCl were not significantly different from the BG-11 medium ($p = 0.165$) and 50 mM ($p = 0.091$). Pairwise comparisons also showed that the phenotypes observed at 150 mM NaCl were not statistically different from the phenotypes at 50 mM ($p = 0.04$). On the other hand, based on the TCD data, the phenotypes resulting from 50 mM, 100 mM, and 150 mM NaCl treatments were all statistically different from the BG-11 medium control (all p-values < 0.001), but they were not different from one another (all p-values >0.027). Finally,

Regression (Appendix S1) was applied between NaCl concentrations and the TCD and TPD mean values. A coefficient of determination (R^2) of 0.64 and 0.51 were obtained for TPD and TCD, respectively.

Rametrix^{TM} PRO was used to apply leave-one-out analysis to DAPC in Fig. 2B. Similar to the previous analysis, four classification options were included: 0 mM, 50 mM, 100 mM, and 150 mM NaCl. Each of these conditions was selected as the positive condition to generate accuracy, sensitivity, and specificity results. As mentioned previously, sensitivity and specificity values should exceed those of the random chance values to confirm the model has value in identifying *Synechocystis* phenotypes. Highlighted results for a DAPC model built with 12 PCs (representing more than 99% of dataset variability) are given in Table 2. Results for additional DAPC models are given in Appendix S1. The best-performing model (Table 2) returned sensitivity and specificity values that exceeded the random chance values. In addition, 0 mM classification showed the highest sensitivity and specificity rates (100%). This suggests that NaCl-induced phenotypes may occur below 50 mM and remain somewhat similar through 150 mM. The BG-11 control phenotypes were significantly distinguishable from the ones observed for the other three conditions, as shown in Fig. 2B.

MgSO$_4$-induced phenotypes

Synechocystis cells growing under different levels of MgSO$_4$ present in BG-11 medium were analyzed using Raman spectroscopy. PCA, DAPC and PCA contributions generated by the Rametrix^{TM} LITE Toolbox are presented in Fig. 3. Again, no clustering was observed with PCA of Raman spectra (Fig. 3A) in the first two PCs (representing over 75% of the dataset variance). Application of DAPC using 12 PCs (representative of more than 99% of the dataset variance) revealed four distinguishable clusters (Fig. 3B), based on the concentration of MgSO$_4$ in BG-11 medium. These clusters formed two separate groups representing cells grown under limiting concentrations of MgSO$_4$ and BG-11 medium, confirming that limiting MgSO$_4$ in the medium is linked with altering cellular phenotypes. Analysis of Raman shift contributions to the PCA and DAPC groups was also performed. Full results are given in Appendix S1, and highlights are given in Table 1. DAPC contributions are also identified in Fig. S3.

TSD, TCD and TPD values were calculated based on spectra, DAPC, and PCA results. These values were analyzed further by ANOVA and pairwise comparisons. All results are available in the Appendix S1. In general, TCD and TPD increased with the decrease of MgSO$_4$ concentration, with the 15.6 mM classification mean values being closer to the reference when using the TPD means, compared to TCD mean values. Next, ANOVA and pairwise comparison tests were performed. It was found that differences between the groups (in both DAPC and PCA) are statistically significant ($p < 0.001$). Pairwise comparisons allowed the identification of significantly different phenotypes. With both TPD and TCD data, all phenotypes were different from those observed when cells were grown in the original BG-11 medium containing 62.5 mM MgSO$_4$. These results confirm that even a 50% reduction in MgSO$_4$ from the original BG-11 medium gives rise to various phenotypes. Results also show that phenotypes associated with the reduced MgSO$_4$ content (31 mM) are not significantly different from 0 mM ($p = 0.096$) and phenotypes with
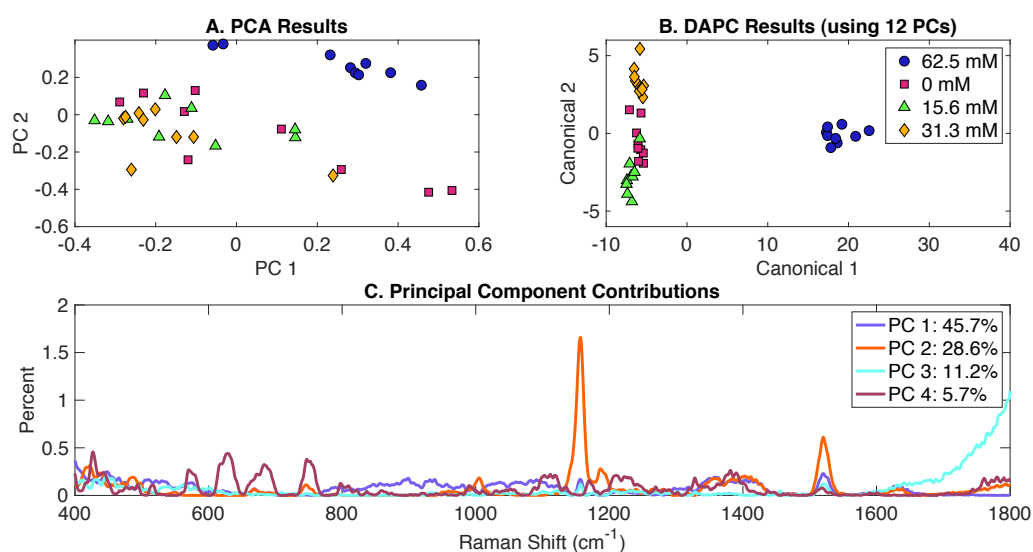
Figure 3 **Rametrix**[TM] **LITE analysis of MgSO₄-induced phenotypes.** Rametrix[TM] LITE analysis of MgSO₄-induced phenotypes. (A) PCA, (B) DAPC, and (C) Raman shifts contributions between groups in PCA. 62.5 mM represents native BG-11 medium.

even more reduced content (15.6 mM) ($p = 0.032$), when using TPD data. Similar results were found when using TCD values. Regression analysis (Appendix S1) between MgSO₄ concentrations and both TCD and TPD mean values was performed. Results show linear relationships between the data for both cases with a correlation coefficient (R) of 0.81 and 0.76 for TPD and TCD, respectively.

For Rametrix[TM] PRO leave-one-out analysis on the DAPC model (Fig. 3B), four classification options were included for the four tested concentrations of MgSO₄. Each one of these conditions was treated as a positive condition to generate accuracy, sensitivity, and specificity results. Again, results for several DAPC models are given in Appendix S1, and results for the DAPC model generated with 12 PCs (the best performer) are given in Table 2. Results confirmed the BG-11 cluster (62.5 mM MgSO₄) (100% sensitivity and specificity) is significantly different from the other clusters. Again, all sensitivity and specificity values exceeded random chance, meaning the DAPC model was capable of applying these classifications for an unknown Raman spectrum of *Synechocystis* with accuracy greater than random chance.

Classification of all phenotypes

All Raman scans performed in this study were compiled in a single dataset and included the growth conditions in the presence of the following environmental stimuli: (i) acetate, (ii) NaCl, (iii) MgSO₄ and (iv) BG-11 medium (set as the control). Data analysis included the generation of PCA and DAPC models and validation with the leave-one-out analysis using Rametrix[TM] LITE and Rametrix[TM] PRO Toolboxes, respectively. PCA results are given in Fig. 4A and show some clustering, based on the two first two PCs which covered over 59% of the dataset variance. A DAPC model was constructed based on 15 PCs (representing
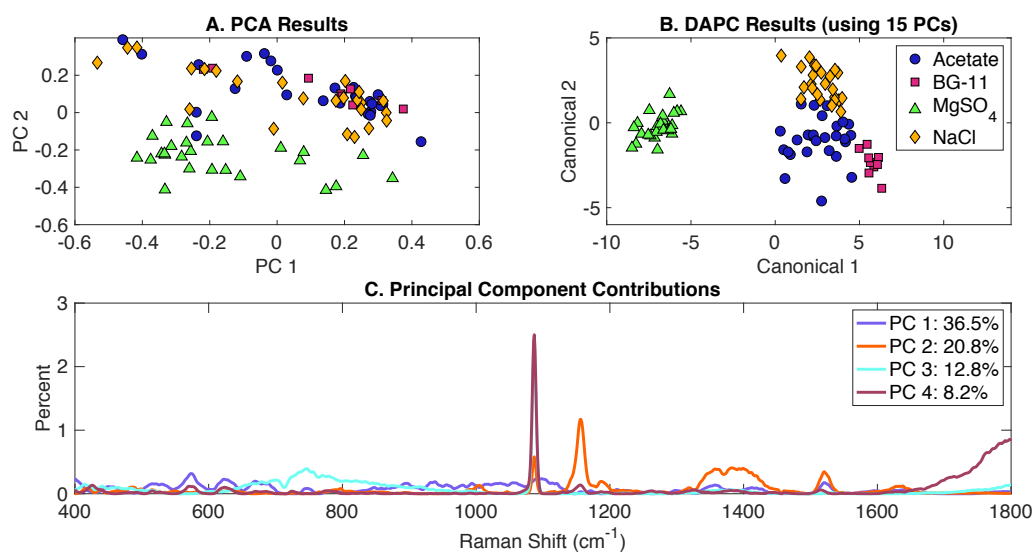
Figure 4 Rametrix^TM LITE analysis of all phenotypes. Rametrix^TM LITE analysis of all phenotypes. (A) PCA, (B) DAPC, and (C) Raman shifts contributions between groups in PCA.

over 99% of the dataset variance) and indicates four distinct clusters based on the presence or absence of environmental stimuli in the BG-11 medium (Fig. 4B). These clusters could be the result of potential differences in the cellular phenotypes obtained with every growth condition. Most importantly, a major segregation was observed for cells grown under $MgSO_4$, clearly separating from the other conditions. Raman shift contributions to PCA (Fig. 4C) and DAPC groups (Fig. S4) are listed in the Appendix S1.

The analysis of TCD and TPD, using the BG-11 medium as a reference, allowed the classification of phenotypes. Both TCD and TPD mean values indicated the distance from the control (indicating more different phenotypes): (i) acetate-, (ii) NaCl-, and (iii) limited $MgSO_4$-induced phenotypes. It appears that $MgSO_4$ limitation generated the most significant changes in the cellular phenotypic responses (also shown in Fig. 4B). In addition, multi-way ANOVA indicated statistically significant differences between the groups representing each phenotype ($p < 0.001$). Using TPD data, pairwise comparisons indicated that all groups are significantly different from each other, only acetate (over all concentrations together) was not statistically different from BG-11 ($p = 0.021$). Based on TCD data (using 15 PCs), all phenotypes are significantly different ($p < 0.001$).

Next, Rametrix^TM PRO was applied to determine if the phenotypes generated by the DAPC model could be predicted from Raman spectra (Appendix S1). Results are highlighted in Table 2. Again, all results exceeded the random chance values, indicating ability by the DAPC model to identify environmental stimuli exposure with accuracy much better than random chance.

Correlation of individual Raman bands with well-established analytical approaches

Another method of comparing phenotypes is analysis of Raman bands assigned to functional groups of biomolecules. These signals are then correlated with results from accepted analytical methods such as GC-FID and UPLC for analyzing the levels of fatty acids and amino acids, respectively. The analysis of individual Raman band intensities was conducted based on the sets of Raman bands identified in previous research (*De Gelder et al., 2007*; *Movasaghi, Rehman & Rehman, 2007*; *Zhu et al., 2011*) and all band assignment are available in Appendix S2.

Acetate study

The levels of amino acids (protein-derived and free amino acids) were determined by both Raman spectroscopy and UPLC analysis (Fig. S5). The general trend is an increase in amino acids levels for the three tested concentration of acetate (7.5 mM, 15 mM, 30 mM) compared to the control with an over-production of amino acids at 15 mM acetate and comparable levels at 7.5 mM and 30 mM acetate.

Raman assigned bands for the different amino acids (*De Gelder et al., 2007*; *Zhu et al., 2011*) were correlated with the UPLC data (Fig. S6). For most amino acids, correlation coefficients (R) varying between 0.61 and 0.97 were obtained. Ala, Leu, Lys, Met, and Ser had correlations with values of R ranging from 0.43 to 0.59. Since the acidic protein hydrolysis method used to release amino acid prevents from determining Cys and Trp levels (*Olcott & Fraenkel-Conrat, 1947*; *Hugli & Moore, 1972*; *Yamada, Moriya & Tsugita, 1991*), these amino acids were predicted using Raman band assignment alone (Fig. S7). The same trend was observed when Cys and Trp levels were higher at 15 mM of acetate.

Similarly, fatty acid content was evaluated and then compared to Raman bands (*De Gelder et al., 2007*; *Czamara et al., 2015*). Overall, there was an increase in fatty acid levels, mainly for C16:0, C18:2 and C18:3, where, at 15 mM acetate, levels of these fatty acids were the highest (Fig. S8). However, minor changes were observed for C16:1 and C18:1. Next, Raman signals corresponding to those fatty acids were compared with the GC-FID results (Fig. S9). Correlations were observed for C16:0 ($R = 0.92$) and C18:2 ($R = 0.82$), and correlation coefficients of 0.67 and 0.68 were found for C18:1 and C18:3, respectively. C16:1 had a correlation 0.46.

NaCl study

Cells growing in different concentrations of NaCl were examined for their amino acid and fatty acid levels with UPLC and GC-FID, respectively. Raman signals corresponding to these biomolecules were used for comparison. Amino acid analysis revealed an increase in their levels proportionally to the amount of NaCl supplied to the media with a distinguishable over production of amino acids in the cells grown in the presence of 150 mM NaCl (Fig. S10). Comparison with the associated Raman bands showed correlation between the UPLC and Raman microscopy data (Fig. S11), where correlation (R) values were between 0.7–0.99, and slightly less correlation was observed for Val ($R = 0.64$). As stated previously, Cys and Trp are only predicted using Raman spectroscopy since they degrade during
acidic hydrolysis (*Olcott & Fraenkel-Conrat, 1947*; *Hugli & Moore, 1972*; *Yamada, Moriya & Tsugita, 1991*). Again, based on the Raman assignments, we were able to resolve the patterns of these two amino acids, showing an increase in their levels mainly at 150 mM of NaCl (Fig. S12).

Finally, GC-FID analysis of the levels of free and lipid-derived fatty acids (Fig. S13) revealed an increase in the levels of C16:0, C18:2, and C18:3 in cells subjected to 50 mM of NaCl. In addition, no changes were observed for C16:1 and C18:1 levels. These results were compared with specific Raman bands assigned to each fatty acid. With the exception of C18:2, correlations (R) between 0.7 and 0.95 were obtained. A lower correlation was observed for the C18:2 between the Raman and GC-FID data ($R = 0.28$). This value improved to 0.7 when the 0 mM NaCl data point was removed from the analysis.

MgSO$_4$ study

The correlation of the individual Raman bands with their associated biomolecules could not be achieved for this study. MgSO$_4$ is the major source of both elements in the BG-11 medium and sufficient amounts of the biomass could not be obtained for cells deprived of this salt, which prevented UPLC and GC-FID analyses.

DISCUSSION

Newly developed methods of analysis by Raman spectroscopy using Rametrix^{TM} enabled the assessment of phenotypic variations among *Synechocystis* cells growing in the presence of external substrates or stressors (acetate, NaCl, and MgSO$_4$). Two different approaches of analysis were undertaken: (i) overall metabolic phenotype changes using Rametrix^{TM} chemometric methods and (ii) correlation of specific Raman bands with their corresponding biomolecules and comparisons with data obtained by well-established analytical methods (i.e., UPLC and GC-FID). While more advanced instrumentation and single-cell methodologies exist (*Wu et al., 2011*; *Serrano et al., 2014*), the purpose of this research was to develop phenotyping capabilities with a relatively inexpensive and portable Raman microscope (with $10\times$ objective imaging) that could have broad appeal to microbiology labs. Several Raman devices are now capable of field research, and Rametrix^{TM} technologies may be able to be implemented there as well.

Synechocystis can grow in the presence of acetate in heterotrophic conditions (*Wu, Shen & Wu, 2002*). However, this microorganism cannot utilize acetate efficiently as it does not express necessary transporters (*Knoop et al., 2013*; *Varman et al., 2013*; *Thiel et al., 2017*). Phenotype changes due to different concentrations of acetate were observed through Rametrix^{TM} LITE clustering of Raman spectral data and validated through statistical analyses (ANOVA and pairwise comparisons). Rametrix^{TM} PRO displayed higher values of sensitivity and specificity (100% and 89%, respectively) over random chance values, supporting that our model can predict effectively the level of acetate in the culture media. Previous studies revealed that acetate is used for the biosynthesis of amino acids (*Varman et al., 2013*; *Thiel et al., 2017*) and lipids (*Thiel et al., 2017*) which is in agreement with our results (both with Raman spectroscopy and standard analytical methods). An over-accumulation of amino acids and fatty acids was observed mainly for acetate concentration

of 15 mM. However, addition of 30 mM acetate into the BG-11 medium did not have a stronger effect compared to the 15 mM acetate. It is obvious that the amount of acetate supplied is limited by diffusion through the cell membrane (*Knoop et al., 2013*; *Varman et al., 2013*; *Thiel et al., 2017*) and that 15 mM might be the maximum concentration that could induce changes in the levels of the tested biomolecules. This result does not exclude the changes in the phenotypic responses observed with the clustering in Rametrix™ LITE and suggest that other changes might happen at 30 mM acetate. According to DAPC contributions, these changes include altered polysaccharide composition, and the levels and types of nucleotides and carotenoids.

NaCl-induced phenotypes were clearly different among the cells grown in the presence of different concentrations of this salt. ANOVA confirmed these results and pairwise comparisons revealed which phenotypes are significantly different from each other. In addition, Rametrix™ PRO was able to predict with accuracy, sensitivity (100%) and specificity (100%) greater than random chance the concentration of NaCl based on an unknown Raman spectrum of *Synechocystis* cells. Data from both Raman and the chromatography-based methods showed high correlations and were in an agreement with previously published data. In fact, salt stress triggers specific cellular responses, including synthesis of stress-related proteins (*Hagemann, Techel & Rensing, 1991*; *Hagemann & Erdmann, 1994*; *Fulda et al., 2000*; *Huang et al., 2006*) and alterations in fatty acid composition (*Huflejt et al., 1990*; *Khomutov et al., 1990*; *Allakhverdiev et al., 1999*; *Singh, Sinha & Häder, 2002*). In particular, 150 mM NaCl induces an over-production of amino acids (exceeding a 15-fold increase compared to the control) and significant changes in fatty acid composition.

Clustering of Raman spectral data via Rametrix™ LITE allowed the differentiation of metabolic phenotypes among *Synechocystis* cells in the presence of $MgSO_4$-limiting concentrations. These results were validated by ANOVA and pairwise comparison. In addition, the DAPC model could predict the $MgSO_4$ levels in the culture media with 100% specificity and sensitivity. The importance of Raman spectroscopy is well-demonstrated also as it provided valuable information with a limited biomass sample when the traditional analytical approaches could not be implemented. In addition, Raman shift contributions to the PCA and DAPC allowed the identification of specific biomolecules that contributed to the phenotypic changes observed with Rametrix™ LITE. These biomolecules included carotenoids, amino acids, lipids, phospholipids, and saccharides. Specific responses observed in photosynthetic organisms during acclimation to Mg and S deprivation include glycogen accumulation (*Schwarz & Forchhammer, 2005*), changes in amino acids composition (*Kiyota, Ikeuchi & Hirai, 2012*) as well as lipids (*Sato et al., 2017*) and degradation of photosynthetic components (i.e., chlorophyll a, PBS and carotenoids) (*Jensen & Rachin, 1984*; *Richaud et al., 2001*; *Schmidt, Erdle & Köst, 2014*). Interestingly, most of these biomolecules were identified with Raman shift contributions, suggesting that our findings are in an agreement with previously established results.

CONCLUSIONS

Raman spectroscopy and analysis with the Rametrix^{TM} Toolboxes and statistical tests enabled the identification of phenotypic changes in cells growing in the presence of different concentrations of external stressors or nutrients. In general, correlations were observed between Raman band intensities and measurements obtained with well-established analytical methods. Most importantly, Raman spectroscopy enabled the identification of the major biomolecules contributing to the phenotypic changes in samples with limited biomass, demonstrating its usefulness.

Funding

This research was funded by the National Science Foundation award NSF1243988. Eva Collakova and Ryan Senger are also supported by HATCH funds. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
National Science Foundation: NSF1243988.
HATCH funds.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Imen Tanniche conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Eva Collakova conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Cynthia Denbow performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Ryan S. Senger conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The Rametrix^{TM} LITE Toolbox is available at GitHub under the MIT license (https://github.com/SengerLab/RametrixLITEToolbox), and the Rametrix^{TM} PRO Toolbox is also accessible through GitHub under the MIT license (https://github.com/SengerLab/RametrixPROToolbox).

Raw Raman scans and biological measurement raw data are available in the Supplementary File.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.8535#supplemental-information.

REFERENCES

Allakhverdiev SI, Nishiyama Y, Suzuki I, Tasaka Y, Murata N. 1999. Genetic engineering of the unsaturation of fatty acids in membrane lipids alters the tolerance of *Synechocystis* to salt stress. *Proceedings of the National Academy of Sciences of the United States of America* **96**:5862–5867 DOI 10.1073/pnas.96.10.5862.

Alsaker KV, Paredes C, Papoutsakis ET. 2010. Metabolite stress and tolerance in the production of biofuels and chemicals: gene-expression-based systems analysis of butanol, butyrate, and acetate stresses in the anaerobe *Clostridium acetobutylicum*. *Biotechnology and Bioengineering* **105**:1131–1147 DOI 10.1002/bit.22628.

Athamneh AI, Alajlouni RA, Wallace RS, Seleem MN, Senger RS. 2014. Phenotypic profiling of antibiotic response signatures in *Escherichia coli* using Raman spectroscopy. *Antimicrobial Agents and Chemotherapy* **58**:1302–1314 DOI 10.1128/aac.02098-13.

Borges N, Ramos A, Raven ND, Sharp RJ, Santos H. 2002. Comparative study of the thermostabilizing properties of mannosylglycerate and other compatible solutes on model enzymes. *Extremophiles* **6**:209–216 DOI 10.1007/s007920100236.

Butler HJ, Ashton L, Bird B, Cinque G, Curtis K, Dorney J, Esmonde-White K, Fullwood NJ, Gardner B, Martin-Hirsch PL, Walsh MJ, McAinsh MR, Stone N, Martin FL. 2016. Using Raman spectroscopy to characterize biological materials. *Nature Protocols* **11**:664–687 DOI 10.1038/nprot.2016.036.

Czamara K, Majzner K, Pacia MZ, Kochan K, Kaczor A, Baranska M. 2015. Raman spectroscopy of lipids: a review. *Journal of Raman Spectroscopy* **46**:4–20 DOI 10.1002/jrs.4607.

Das RS, Agrawal YK. 2011. Raman spectroscopy: recent advancements, techniques and applications. *Vibrational Spectroscopy* **57**:163–176 DOI 10.1016/j.vibspec.2011.08.003.

De Gelder J, De Gussem K, Vandenabeele P, Moens L. 2007. Reference database of Raman spectra of biological molecules. *Journal of Raman Spectroscopy* **38**:1133–1147 DOI 10.1002/jrs.1734.

Ewanick SM, Thompson WJ, Marquardt BJ, Bura R. 2013. Real-time understanding of lignocellulosic bioethanol fermentation by Raman spectroscopy. *Biotechnology for Biofuels* **6**:28 DOI 10.1186/1754-6834-6-28.

Fisher AK, Carswell WF, Athamneh AIM, Sullivan MC, Robertson JL, Bevan DR, Senger RS. 2018. The RametrixTM LITE toolbox v1.0 for MATLAB®. *Journal of Raman Spectroscopy* **49**:885–896 DOI 10.1002/jrs.5348.

Freedman BG, Zu TNK, Wallace RS, Senger RS. 2016. Raman spectroscopy detects phenotypic differences among *Escherichia coli* enriched for 1-butanol tolerance using a metagenomic DNA library. *Biotechnology Journal* **11**:877–889 DOI 10.1002/biot.201500144.

Fulda S, Huang F, Nilsson F, Hagemann M, Norling B. 2000. Proteomics of *Synechocystis* sp. strain PCC 6803. *European Journal of Biochemistry* **267**:5900–5907 DOI 10.1046/j.1432-1327.2000.01642.x.

García-Timermans C, Rubbens P, Heyse J, Props R, Kerckhof F-M, Skirtach A, Waegeman W, Boon N. 2019. Characterizing phenotypic heterogeneity in isogenic bacterial populations using flow cytometry and Raman spectroscopy. *bioRxiv* DOI 10.1101/545681.

Hagemann M. 2013. Chapter two—genomics of salt acclimation: synthesis of compatible solutes among cyanobacteria. In: Chauvat F, Cassier-Chauvat C, eds. *Advances in botanical research*. Vol. 65. Academic Press, 27–55 DOI 10.1016/B978-0-12-394313-2.00002-0.

Hagemann M, Erdmann N. 1994. Activation and pathway of glucosylglycerol synthesis in the cyanobacterium *Synechocystis* sp. PCC 6803. *Microbiology* **140**:1427–1431 DOI 10.1099/00221287-140-6-1427.

Hagemann M, Techel D, Rensing L. 1991. Comparison of salt- and heat-induced alterations of protein synthesis in the cyanobacterium *Synechocystis* sp. PCC 6803. *Archives of Microbiology* **155**:587–592 DOI 10.1007/BF00245354.

Heidorn T, Camsund D, Huang H-H, Lindberg P, Oliveira P, Stensjö K, Lindblad P. 2011. Synthetic biology in cyanobacteria engineering and analyzing novel functions. *Methods in Enzymology* **497**:539–579 DOI 10.1016/B978-0-12-385075-1.00024-X.

Hincha DK, Hagemann M. 2004. Stabilization of model membranes during drying by compatible solutes involved in the stress tolerance of plants and microorganisms. *Biochemical Journal* **383**:277–283 DOI 10.1042/BJ20040746.

Huang F, Fulda S, Hagemann M, Norling B. 2006. Proteomic screening of salt-stress-induced changes in plasma membranes of *Synechocystis* sp. strain PCC 6803. *Proteomics* **6**:910–920 DOI 10.1002/pmic.200500114.

Huflejt ME, Tremolieres A, Pineau B, Lang JK, Hatheway J, Packer L. 1990. Changes in membrane lipid composition during saline growth of the fresh water cyanobacterium *Synechococcus* 6311. *Plant Physiology* **94**:1512–1521 DOI 10.1104/pp.94.4.1512.

Hugli TE, Moore S. 1972. Determination of the tryptophan content of proteins by ion exchange chromatography of alkaline hydrolysates. *Journal of Biological Chemistry* **247**:2828–2834.

Isken S, De Bont JA. 1998. Bacteria tolerant to organic solvents. *Extremophiles* **2**:229–238 DOI 10.1007/s007920050065.

Jensen TE, Rachin JW. 1984. Effect of varying sulphur deficiency on structural components of a cyanobacterium *Synechococcus leopoliensis*: a morphometric study. *Cytobios (AGRIS)* **41**:35–46.

Karandashova IV, Elanskaya IV. 2005. Genetic control and mechanisms of salt and hyperosmotic stress resistance in cyanobacteria. *Russian Journal of General Chemistry* **41**:1311–1321 DOI 10.1007/s11177-006-0001-z.

Khomutov G, Fry IV, Huflejt ME, Packer L. 1990. Membrane lipid composition, fluidity, and surface charge changes in response to growth of the fresh water cyanobacterium *Synechococcus* 6311 under high salinity. *Archives of Biochemistry and Biophysics* **277**:263–267 DOI 10.1016/0003-9861(90)90577-L.

Kim HW, Vannela R, Zhou C, Rittmann BE. 2011. Nutrient acquisition and limitation for the photoautotrophic growth of *Synechocystis* sp. PCC6803 as a renewable biomass source. *Biotechnology and Bioengineering* **108**:277–285 DOI 10.1002/bit.22928.

Kirschner C, Maquelin K, Pina P, Thi NANgo, Choo-Smith LP, Sockalingum GD, Sandt C, Ami D, Orsini F, Doglia SM, Allouch P, Mainfait M, Puppels GJ, Naumann D. 2001. Classification and identification of enterococci: a comparative phenotypic, genotypic, and vibrational spectroscopic study. *Journal of Clinical Microbiology* **39**:1763–1770 DOI 10.1128/JCM.39.5.1763-1770.2001.

Kiyota H, Ikeuchi M, Hirai MY. 2012. Response of amino acid metabolism to sulfur starvation in *Synechocystis* sp. PCC 6803. In: De Kok LJ, Tabe L, Tausz M, Hawkesford MJ, Hoefgen R, McManus MT, Norton RM, Rennenberg H, Saito K, Schnug E, eds. *Sulfur metabolism in plants. Proc Int plant sulfur work.* Springer Netherlands, 53–59.

Knoop H, Gründel M, Zilliges Y, Lehmann R, Hoffmann S, Lockau W, Steuer R. 2013. Flux balance analysis of cyanobacterial metabolism: the metabolic network of *Synechocystis* sp. PCC 6803. *PLOS Computational Biology* **9**:e1003081 DOI 10.1371/journal.pcbi.1003081.

Koya SK, Brusatori M, Martin JV, Yurgelevic S, Huang C, Liberati DM, Auner GW, Diebel LN. 2018. Rapid detection of Clostridium difficile toxins in serum by Raman spectroscopy. *Journal of Surgical Research* **232**:195–201 DOI 10.1016/j.jss.2018.06.031.

Lasko DR, Schwerdel C, Bailey JE, Sauer U. 1997. Acetate-specific stress response in acetate-resistant bacteria: an analysis of protein patterns. *Biotechnology Progress* **13**:519–523 DOI 10.1021/bp970075f.

Liu J, Sun J, Huang X, Li G, Liu B. 2015. Goldindec: a novel algorithm for Raman spectrum baseline correction. *Applied Spectroscopy* **69**:834–842 DOI 10.1366/14-07798.

Lusk JE, Williams RJ, Kennedy EP. 1968. Magnesium and the growth of *Escherichia coli*. *Journal of Biological Chemistry* **243**:2618–2624.

Marin K, Kanesaki Y, Los DA, Murata N, Suzuki I, Hagemann M. 2004. Gene expression profiling reflects physiological processes in salt acclimation of *Synechocystis* sp. strain PCC 6803. *Plant Physiology* **136**:3290–3300 DOI 10.1104/pp.104.045047.

Marin K, Stirnberg M, Eisenhut M, Kramer R, Hagemann M. 2006. Osmotic stress in *Synechocystis* sp. PCC 6803: low tolerance towards nonionic osmotic stress results from lacking activation of glucosylglycerol accumulation. *Microbiology* **152**:2023–2030 DOI 10.1099/mic.0.28771-0.

Markovitz A, Sylvan S. 1962. Effect of sodium sulfate and magnesium sulfate on heteropolysaccharide synthesis in Gram-negative soil bacteria. *Journal of Bacteriology* **83**:483–489 DOI 10.1128/JB.83.3.483-489.1962.

Mikkat S, Galinski EA, Berg G, Minkwitz A, Schoor A. 2000. Salt adaptation in Pseudomonads: characterization of glucosylglycerol-synthesizing isolates from brackish coastal waters and the rhizosphere. *Systematic and Applied Microbiology* **23**:31–40 DOI 10.1016/S0723-2020(00)80043-0.

Movasaghi Z, Rehman S, Rehman IU. 2007. Raman spectroscopy of biological tissues. *Applied Spectroscopy Reviews* **42**:493–541 DOI 10.1080/05704920701551530.

Nelson WH, Manoharan R, Sperry JF. 1992. UV resonance Raman studies of bacteria. *Applied Spectroscopy Reviews* **27**:67–124 DOI 10.1080/05704929208018270.

Nicolaou SA, Gaida SM, Papoutsakis ET. 2010. A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: from biofuels and chemicals, to biocatalysis and bioremediation. *Metabolic Engineering* **12**:307–331 DOI 10.1016/j.ymben.2010.03.004.

Olcott HS, Fraenkel-Conrat H. 1947. Formation and loss of cysteine during acid hydrolysis of proteins. Role of tryptophan. *Journal of Biological Chemistry* **171**:583–594.

Olson ML, Johnson J, Carswell WF, Reyes LH, Senger RS, Kao KC. 2016. Characterization of an evolved carotenoids hyper-producer of *Saccharomyces cerevisiae* through bioreactor parameter optimization and Raman spectroscopy. *Journal of Industrial Microbiology and Biotechnology* **43**:1355–1363 DOI 10.1007/s10295-016-1808-9.

Pahlow S, Meisel S, Cialla-May D, Weber K, Rösch P, Popp J. 2015. Isolation and identification of bacteria by means of Raman spectroscopy. *Advanced Drug Delivery Reviews* **89**:105–120 DOI 10.1016/j.addr.2015.04.006.

Papoutsakis ET, Bussineau CM, Chu IM, Diwan AR, Huesemann M. 1987. Transport of substrates and metabolites and their effect on cell metabolism (in butyric-acid and methylotrophic fermentations). *Annals of the New York Academy of Sciences* **506**:24–50 DOI 10.1111/j.1749-6632.1987.tb23808.x.

Reed RH, Borowitzka LJ, Mackay MA, Chudek JA, Foster R, Warr SRC, Moore DJ, Stewart WDP. 1986. Organic solute accumulation in osmotically stressed cyanobacteria. *FEMS Microbiology Letters* **39**:51–56 DOI 10.1016/0378-1097(86)90060-1.

Reed RH, Warr SRC, Richardson DL, Moore DJ, Stewart WDP. 1985. Multiphasic osmotic adjustment in a euryhaline cyanobacterium. *FEMS Microbiology Letters* **28**:225–229 DOI 10.1111/j.1574-6968.1985.tb00796.x.

Richaud C, Zabulon G, Joder A, Thomas J-C. 2001. Nitrogen or sulfur starvation differentially affects phycobilisome degradation and expression of the nblA gene in *Synechocystis* strain PCC 6803. *Journal of Bacteriology* **183**:2989–2994 DOI 10.1128/JB.183.10.2989-2994.2001.

Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY. 1978. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Journal of General Microbiology* **111**:1–61.

Roe AJ, O'Byrne C, McLaggan D, Booth IR. 2002. Inhibition of *Escherichia coli* growth by acetic acid: a problem with methionine biosynthesis and homocysteine toxicity. *Microbiology* **148**:2215–2222 DOI 10.1099/00221287-148-7-2215.

Russell JB. 1992. Another explanation for the toxicity of fermentation acids at low pH: anion accumulation versus uncoupling. *Journal of Applied Bacteriology* **73**:363–370 DOI 10.1111/j.1365-2672.1992.tb04990.x.

Sardessai Y, Bhosle S. 2002. Tolerance of bacteria to organic solvents. *Research in Microbiology* **153**:263–268 DOI 10.1016/S0923-2508(02)01319-0.

Sato N, Kamimura R, Kaneta K, Yoshikawa M, Tsuzuki M. 2017. Species-specific roles of sulfolipid metabolism in acclimation of photosynthetic microbes to sulfur-starvation stress. *PLOS ONE* **12**:e0186154 DOI 10.1371/journal.pone.0186154.

Schmidt A, Erdle I, Köst H-P. 2014. Changes of C-phycocyanin in *Synechococcus* 6301 in relation to growth on various sulfur compounds materials and methods. *Zeitschrift für Naturforschung C* **37**:870–876 DOI 10.1515/znc-1982-1004.

Schwarz R, Forchhammer K. 2005. Acclimation of unicellular cyanobacteria to macronutrient deficiency: emergence of a complex network of cellular responses. *Microbiology* **151**:2503–2514 DOI 10.1099/mic.0.27883-0.

Sellick CA, Reece RJ. 2005. Eukaryotic transcription factors as direct nutrient sensors. *Trends in Biochemical Sciences* **30**:405–412 DOI 10.1016/j.tibs.2005.05.007.

Senger RS, Kuvaru V, Sullivan M, Gouldin A, Lundgren S, Merrifield K, Steen C, Baker E, Vu T, Agnor B, Martinez G, Coogan H, Carswell W, Karageorge L, Dev D, Du P, Sklar A, Orlando G, Pirkle J, Robertson JL. 2019. Spectral characteristics of urine specimens from healthy human volunteers analyzed using Raman chemometric urinalysis (Rametrix). *PLOS ONE* **14(9)**:e0222115 DOI 10.1371/journal.pone.0222115.

Senger RS, Robertson JL. 2020. The Rametrix[TM] PRO Toolbox v1.0 for MATLAB®. *PeerJ* **8**:e8179 DOI 10.7717/peerj.8179.

Serrano P, Wagner D, Bottger U, De Vera JP, Lasch P, Hermelink A. 2014. Single-cell analysis of the methanogenic archaeon *Methanosarcina soligelidi* from Siberian permafrost by means of confocal Raman microspectroscopy for astrobiological research. *Planetary and Space Science* **98**:191–197.

Singh S, Sinha RP, Häder D-P. 2002. Role of lipids and fatty acids in stress tolerance in cyanobacteria. *Acta Protozoologica* **41**:297–308.

Sivakesava S, Irudayaraj J, Demirci A. 2001. Monitoring a bioprocess for ethanol production using FT-MIR and FT-Raman spectroscopy. *Journal of Industrial Microbiology and Biotechnology* **26**:185–190 DOI 10.1038/sj.jim.7000124.

Sun S, Wang X, Gao X, Ren L, Su X, Bu D, Ning K. 2015. Condensing Raman spectrum for single-cell phenotype analysis. *BMC Bioinformatics* **16(Suppl 18)**:S15 DOI 10.1186/1471-2105-16-S18-S15.

Thiel K, Vuorio E, Aro E-M, Kallio PT. 2017. The effect of enhanced acetate influx on *Synechocystis* sp. PCC 6803 metabolism. *Microbial Cell Factories* **16**:21 DOI 10.1186/s12934-017-0640-x.

Tummala SB, Junne SG, Paredes CJ, Papoutsakis ET. 2003. Transcriptional analysis of product-concentration driven changes in cellular programs of recombinant

Clostridium acetobutylicum strains. *Biotechnology and Bioengineering* **84**:842–854 DOI 10.1002/bit.10851.

Varman AM, Yu Y, You L, Tang YJ. 2013. Photoautotrophic production of D-lactic acid in an engineered cyanobacterium. *Microbial Cell Factories* **12**:117 DOI 10.1186/1475-2859-12-117.

Vermaas W. 1996. Molecular genetics of the cyanobacterium *Synechocystis* sp. PCC 6803: principles and possible biotechnology applications. *Journal of Applied Phycology* **8**:263–273 DOI 10.1007/BF02178569.

Wang S, Shen A, Setlow P, Li Y. 2015. Characterization of the dynamic germination of individual *Clostridium difficile* spores using Raman spectroscopy and differential interference contrast microscopy. *Journal of Bacteriology* **197**:2361 DOI 10.1128/JB.00200-15.

Wu GF, Shen ZY, Wu QY. 2002. Modification of carbon partitioning to enhance PHB production in *Synechocystis* sp. PCC6803. *Enzyme and Microbial Technology* **30**:710–715 DOI 10.1016/S0141-0229(02)00044-3.

Wu H, Volponi JV, Oliver AE, Parikh AN, Simmons BA, Singh S. 2011. In vivo lipidomics using single-cell Raman spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* **108**:3809–3814 DOI 10.1073/pnas.1009043108.

Yamada H, Moriya H, Tsugita A. 1991. Development of an acid hydrolysis method with high recoveries of tryptophan and cysteine for microquantities of protein. *Analytical Biochemistry* **198**:1–5 DOI 10.1016/0003-2697(91)90496-G.

Yu Y, You L, Liu D, Hollinshead W, Tang YJ, Zhang F. 2013. Development of *Synechocystis* sp. PCC 6803 as a phototrophic cell factory. *Marine Drugs* **11**:2894–2916 DOI 10.3390/md11082894.

Zavřel T, Očenášová P, Červený J. 2017. Phenotypic characterization of *Synechocystis* sp. PCC 6803 substrains reveals differences in sensitivity to abiotic stress. *PLOS ONE* **12**:e0189130 DOI 10.1371/journal.pone.0189130.

Zhang Z, Pendse ND, Phillips KN, Cotner JB, Khodursky A. 2008. Gene expression patterns of sulfur starvation in *Synechocystis* sp. PCC 6803. *BMC Genomics* **9**:344 DOI 10.1186/1471-2164-9-344.

Zhu G, Zhu X, Fan Q, Wan X. 2011. Raman spectra of amino acids and their aqueous solutions. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **78**:1187–1195 DOI 10.1016/j.saa.2010.12.079.

Zu NT, Athamneh IA, Senger SR. 2016. Characterizing the phenotypic responses of *Escherichia coli* to multiple 4-carbon alcohols with Raman spectroscopy. *Fermentation* **2**(**1**):3 DOI 10.3390/fermentation2010003.

Zu TNK, Athamneh AIM, Wallace RS, Collakova E, Senger RS. 2014. Near-real-time analysis of the phenotypic responses of *Escherichia coli* to 1-butanol exposure using Raman spectroscopy. *Journal of Bacteriology* **96**:3983–3991 DOI 10.1128/JB.01590-14.

Zu TNK, Liu S, Germane KL, Servinsky MD, Gerlach ES, Mackie DM, Sund CJ. 2016. Predictive modeling in *Clostridium acetobutylicum* fermentations employing Raman spectroscopy and multivariate data analysis for real-time culture monitoring. In: *Proc. SPIE 9863, Smart biomedical and physiological sensor technology XIII, 98630I (13 May 2016)*. DOI 10.1117/12.2228545.