**Supporting Information** 

## Influence of Carbon Sources on Quantification of Deuterium Incorporation in Heterotrophic Bacteria: A Raman-Stable Isotope Labeling Approach

Georgette Azemtsop Matanfack<sup>1,2,3</sup>, Martin Taubert<sup>4</sup>, Shuxia Guo<sup>1,2,3</sup>, Rola Houhou<sup>1,2,3</sup>, Thomas Bocklitz<sup>1,2,3</sup>, Kirsten Küsel<sup>4,5</sup>, Petra Rösch<sup>\*1,2,3</sup> and Jürgen Popp<sup>1,2,3</sup>

<sup>1</sup>Institute of Physical Chemistry and Abbe Center of Photonics (IPC), Friedrich-Schiller-University Jena, Helmholtzweg 4, 07743 Jena, Germany
<sup>2</sup>Leibniz Institute of Photonic Technology (Leibniz-IPHT), Albert-Einstein-Straße 9, 07745 Jena, Germany
<sup>3</sup>Research Campus Infectognostics e.v. Jena, Germany
<sup>4</sup>Aquatic Geomicrobiology, Institute of Biodiversity, Friedrich Schiller University Jena, Dornburger Str. 159, 07743 Jena, Germany
<sup>5</sup>German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5E, 04103 Leipzig, Germany

## **Corresponding Author**

\*E-mail: petra.roesch@uni-jena.de. Phone: +49-36419-48381. Fax: +49-36419-48302.

Figure S1. Mean spectra of *Bacillus-1* cultured under different carbon sources in  $H_2O$  or heavy water (D<sub>2</sub>O).

**Figure S2.** Mean spectra of *Bacillus-2* cultured under different carbon sources in  $H_2O$  or heavy water (D<sub>2</sub>O).

Figure S3. Mean spectra of *Microbacterium* cultured under different carbon sources in  $H_2O$  or heavy water (D<sub>2</sub>O).

**Figure S4.** Mean spectra of *Arthrobacter* cultured under different carbon sources in H<sub>2</sub>O or heavy water (D<sub>2</sub>O).

**Figure S5.** Mean spectra of *Paenibacillus* cultured under different carbon sources in H<sub>2</sub>O or heavy water (D<sub>2</sub>O).

**Figure S6.** Mean spectra of *Pedobacter cells* cultured in heavy water ( $D_2O$ ), after correction to separate overtone and combination bands of carotenoids from the C-D stretching vibration band.

**Figure S7.** Relative deuterium amount of all data combined for all the six carbon source conditions grown in H<sub>2</sub>O and D<sub>2</sub>O.

**Figure S8.** C-D ratios of all true positive spectra from the PCA-LDA-"full-spectrum model", built on the whole spectral range, for the differentiation between labeled and non-labeled cells.

**Figure S9.** C-D ratios of all true positive spectra from the PCA-LDA-"C-D region model", built on the C-D region  $(2040 - 2300 \text{ cm}^{-1})$ , to distinguish deuterated and non-deuterated cells.

**Figure S10.** C-D ratios and some prominent examples of the misclassified single cells spectra from the PCA-LDA-"C-D region model", built on the C-D region ( $2040 - 2300 \text{ cm}^{-1}$ ), for the differentiation of deuterium labeled and non-labeled cells.

**Figure S11.** PCA-LDA loading vector and the corresponding score plot resulting from the "full-spectrum model."

Table S1. Overview of heterotrophic bacterial growth in different carbon sources.

**Table S2.** LDA classification model for the differentiation between deuterium labeled and nonlabeled bacteria based on the whole Raman spectral range.

**Table S3.** LDA classification model for the differentiation between deuterium labeled and nonlabeled bacteria based on the restricted C-D stretching vibrations region (2040-2300 cm<sup>-1</sup>).



**Figure S1.** Mean spectra of *Bacillus-1* grown in H<sub>2</sub>O (**A**) and heavy water (**B**). Each spectrum represents a carbon source. Growth conditions in minimal medium with single carbon sources: casaminoacids (ca), peptone (pe), SCFA (fa), sugars (su), as well as complex medium (cm). *Bacillus-1* could not grow in aromatics. **C:** Relative amount of deuterium incorporation into cells under various carbon source conditions. Dots represent single cell spectra; each boxplot shows median, upper and lower quartile. The black line represents the mean value of the relative deuterium amount of all six carbon source conditions, whereas the number in each boxplot represents the mean of each plotted group. Significant differences to respective systems were determined by Wilcoxon test (\*p<0.05; \*\*p<0.01; \*\*\*p<0.001; \*\*\*\*p<0.0001; N.S., not significant).



**Figure S2.** Mean spectra of *Bacillus-2* grown in H<sub>2</sub>O (**A**) and heavy water (**B**). Each spectrum represents a carbon source. Growth conditions in minimal medium with single carbon sources: aromatics (ar), casaminoacids (ca), peptone (pe), SCFA (fa), sugars (su), as well as complex medium (cm). **C:** Relative amount of deuterium incorporation into cells under various carbon source conditions. Dots represent single cell spectra; each boxplot shows median, upper and lower quartile. The black line represents the mean value of the relative deuterium amount of all six carbon source conditions, whereas the number in each boxplot represents the mean of each plotted group. Significant differences to respective systems were determined by Wilcoxon test (\*p<0.05; \*\*p<0.01; \*\*\*p<0.001; N.S., not significant).



**Figure S3.** Mean spectra of *Microbacterium* grown H<sub>2</sub>O (**A**) and heavy water (**B**). Each spectrum represents a carbon source. Growth conditions in minimal medium with single carbon sources: aromatics (ar), casaminoacids (ca), peptone (pe), SCFA (fa), sugars (su), as well as complex medium (cm). **C:** Relative amount of deuterium incorporation into cells under various carbon source conditions. Dots represent single cell spectra; each boxplot shows median, upper and lower quartile. The black line represents the mean value of the relative deuterium amount of all six carbon source conditions, whereas the number in each boxplot represents the mean of each plotted group. Significant differences to respective systems were determined by Wilcoxon test (\*p<0.05; \*\*p<0.01; \*\*\*p<0.001; N.S., not significant).



**Figure S4.** Mean spectra of *Arthrobacter* grown in H<sub>2</sub>O (**A**) and heavy water (**B**). Each spectrum represents a carbon source. Growth conditions in minimal medium with single carbon sources: aromatics (ar), casaminoacids (ca), peptone (pe), SCFA (fa), sugars (su), as well as complex medium (cm). **C:** Relative amount of deuterium incorporation into cells under various carbon source conditions. Dots represent single cell spectra; each boxplot shows median, upper and lower quartile. The black line represents the mean value of the relative deuterium amount of all six carbon source conditions, whereas the number in each boxplot represents the mean of each plotted group. Significant differences to respective systems were determined by Wilcoxon test (\*p<0.05; \*\*p<0.01; \*\*\*p<0.001; N.S., not significant).



**Figure S5.** Mean spectra of *Paenibacillus* grown in H<sub>2</sub>O (**A**) and heavy water (**B**). Each spectrum represents a carbon source. Growth conditions in minimal medium with sugars (su) as single carbon source and complex medium (cm). *Paenibacillus* was able to grow only in sugars and complex medium. **C**: Relative amount of deuterium incorporation into cells under various carbon source conditions. Dots represent single cell spectra; each boxplot shows median, upper and lower quartile. The black line represents the mean value of the relative deuterium amount of all six carbon source conditions, whereas the number in each boxplot represents the mean of each plotted group. Significant differences to respective systems were determined by Wilcoxon test (\*p<0.05; \*\*p<0.01; \*\*\*p<0.001; N.S., not significant).



**Figure S6.** Mean spectra of *Pedobacter* grown in heavy water ( $D_2O$ ) after correction to separate overtone and combination bands of carotenoid from the C-D stretching vibration band. Each spectrum represents a carbon source. Growth conditions in minimal medium with single carbon sources: casaminoacids (ca), peptone (pe), SCFA (fa), sugars (su), as well as complex medium (cm). *Pedobacter* can not grow in aromatics.



**Figure S7.** Relative deuterium amount of all data combined for all the six carbon source conditions grown in  $D_2O$  and  $H_2O$ . Dots represent single cell spectra; each boxplot shows median, upper and lower quartile. Significant differences to respective systems were determined by Students't-test (\*\*\*\*p<0.0001), owing to the large samples size in each plotted group.



**Figure S8.** Box-plots showing of the C-D ratios of all the 2054 true positive ( $D_2O$ ) spectra from PCA-LDA-"full-spectrum model", built to distinguish deuterium labeled and non-labeled cells. The carbon sources correspond to aromatics (ar), casaminoacids (ca), peptone (pe), SCFA (fa), sugars (su), as well as complex medium (cm). And the bacterial genera correspond to G1: *Pseudomonas*, G2: *Microbacterium*, G3: *Phenylobacterium*, G4: *Bacillus-1*, G5: *Bacillus-2*, G6: *Paenibacillus*, G7: *Pedobacter* and G8: *Arthrobacter*.



**Figure S9.** C-D ratios of all the 2070 true positive ( $D_2O$ ) spectra from the PCA-LDA-"C-D region model", built on the restricted C-D region (2040 – 2300 cm<sup>-1</sup>), for the differentiation between labeled and non-labeled cells. The carbon sources correspond to aromatics (ar), casaminoacids (ca), peptone (pe), SCFA (fa), sugars (su), as well as complex medium (cm). And the bacterial genera correspond to G1: *Pseudomonas*, G2: *Microbacterium*, G3: *Phenylobacterium*, G4: *Bacillus-1*, G5: *Bacillus-2*, G6: *Paenibacillus*, G7: *Pedobacter* and G8: *Arthrobacter*.



**Figure S10.** Mean spectra (**A**) and C-D ratios (**B**) of the 56 single cells, which were misclassified as non-labeled ( $H_2O$ ) in the PCA-LDA-"C-D region model", built to differentiate deuterium labeled and non labeled cells. **A** (**i**): Mean spectrum of forty-seven misclassified spectrum from *Pedobacter* grown in short chain fatty acids (*Pedobacter\_fa*). **A** (**ii**): Mean spectrum of nine misclassified spectra from *Arthrobacter* grown in aromatics (*Arthrobacter\_ar*) **B**: Box plots showing the C-D ratios of the 56 missclassified spectra per isolate. The black line represents the mean of all the plotted groups, and the number in each boxplot represents the mean of each plotted group.



**Figure S11.** PCA-LDA loading vector (**A**) and the corresponding score plot (**B**) calculated from the "full-spectrum model" for the differentiation between deuterium labeled (yellow) and non-labeled (green) cells. Negative peaks and scores characterize the deuterium labeled cells, while positive peaks and scores represent the non-labeled cells.

**Table S1.** Overview of heterotrophic bacterial growth in different carbon sources. +: good cell growth, -:slow cell growth, /: no cell reproduction. No differences were observed in H<sub>2</sub>O and D<sub>2</sub>O cultures.

Carbon sources Genera	Aromatics	Casaminoacids	Peptone	SCFA	Sugars	Complex Medium
Pseudomonas	+	+	+	+	+	+
Phenylobacterium	+	+	+	+	+	+
Pedobacter	1	+	+		+	+
Bacillus isolate 1	/	+	+	+	+	+
Bacillus isolate 2	+	+	+	+	+	+
Microbacterium	+	+	+	+	+	+
Arthrobacter	+	+	+	+	+	+
Paenibacillus	/	/	/	/	+	+

**Table S2.** Confusion matrix of the PCA-LDA-"full-spectrum model" for the differentiation between labeled ( $D_2O$ ) and non-labeled ( $H_2O$ ) bacterial cells, with an overall accuracy of 98.3%.

"Full-spectrum model"	True class labels		
Predicted class labels	D <sub>2</sub> O	H <sub>2</sub> O	
D <sub>2</sub> O	2054	0	
H <sub>2</sub> O	72	2081	

**Table S3.** Confusion matrix of the PCA-LDA-"C-D region model", based on the restricted C-D stretching vibration region (2040-2300 cm<sup>-1</sup>) for the differentiation between labeled ( $D_2O$ ) and non-labeled ( $H_2O$ ) bacterial cells, with an overall accuracy of 98.7 %.

"C-D region model"	True class labels		
Predicted class labels	D <sub>2</sub> O	H <sub>2</sub> O	
D <sub>2</sub> O	2070	0	
H <sub>2</sub> O	56	2081	