Supporting Information

Single-Cell and Time-Resolved Profiling of Intracellular *Salmonella* Metabolism in Primary Human Cells

Jiabao Xu[1], Lorena Preciado-Llanes[2,3,Γ], Anna Aulicino[2,3,Γ], Christoph Martin Decker[4], Maren Depke[4], Manuela Gesell Salazar[4], Frank Schmidt[4,5], Alison Simmons[2,3] and Wei E. Huang[1]*

1. Begbroke Science Park, Department of Engineering Science, University of Oxford, Woodstock Road, Oxford, OX5 1PF, United Kingdom.

2. MRC Human Immunology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford OX3 9DS, UK

3. Translational Gastroenterology Unit, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK

4. Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Felix-Hausdorff-Str. 8, 17475 Greifswald, Germany

5. Proteomics Core, Weill Cornel Medicine-Qatar, Education City, PO 24144 Doha, Qatar

[Γ]L. P. and A. A. have contributed equally to this study.

*Corresponding author: Wei E. Huang

wei.huang@eng.ox.ac.uk

Telephone: +44 (0)1865 283786, Fax: +44 (0)1865 3749

Table of Contents

Supporting Methods

Bacterial strains and culture conditions. The ability of *Salmonella* to incorporate D_2O was tested by growing bacteria overnight in LB Lennox broth (Sigma, UK) supplemented with 35% D_2O (v/v), at 37 °C under continuous shaking at 200 rpm. The degree of D_2O incorporation by STM-LT2 was compared with EC-DH5α, used as reference.

For the infection experiments, STM-LT2, STM-D23580 and ST-Ty2 were grown to logarithmic growth phase in LB Lennox broth (Sigma) supplemented with 10% sucrose (Sigma) and 50% D_2O. Aliquots were kept frozen at –80 °C until use, whilst bacterial viability was monitored periodically. For each experiment, an aliquot of bacteria was thawed, washed and resuspended in RPMI 1640 (Sigma) to obtain a multiplicity of infection (MOI) of 30:1 (bacteria:THP-1), 15:1 (bacteria:macrophage) or 10:1 (bacteria:dendritic cell). As a control for bacterial extracellular growth, RPMI was inoculated with each strain at 37 °C under continuous shaking at 200 rpm. At specific time points, corresponding to the kinetic of the infection, bacteria were collected, washed and processed for the single-cell Raman spectra (SCRS) measurements and analysis, as described below. The number of microorganisms was assessed at each time point post infection (p.i.) by plating 10-fold dilutions of the bacterial suspension, in quadruplicate, on LB Lennox agar (Sigma). The number of bacteria was determined as Colony Forming Units (CFU).

Propagation of THP-1 cells and infection. THP-1 cells were obtained from the American Type Culture Collection (ATCC, TIB-202). The cells were maintained in vented T75 flasks (Cellstar) at 37 °C, 5% CO_2 and split every 2-3 days when confluent (2 x10^6 cells/mL). One day before the infection, THP-1 cells were harvested, washed, resuspended at a density of 1 x10^6 cells/mL/well in 24-well plates (Costar) in complete RPMI medium containing 10% Fetal Bovine Serum (FBS, Sigma), 2 mM L-Glutamine, 100 U/mL of penicillin and 100 μg/mL streptomycin (Sigma), and supplemented with 25 ng/mL of phorbol 12-myristate 13-acetate (PMA, Sigma) to induce differentiation into macrophage-like cells, while incubating at 37 °C, 5% CO_2 for 24h.

Generation of monocyte-derived dendritic cells and macrophages and infection. Leukocyte Reduction System cones were obtained from the UK National Blood Centre with informed consent following local ethical guidelines. Blood from healthy donors was diluted in PBS and layered on a standard density gradient (Lymphoprep^{TM}). Peripheral Blood Mononuclear Cells (PBMC) were collected from the interface and washed three times in PBS at 4 °C. Monocytes were obtained by magnetic isolation using human $CD14^+$ MicroBeads (Miltenyi Biotec, Germany) according to manufacturer's protocol. Monocyte-derived dendritic cells (MoDCs) and macrophages (Mf) were obtained by incubation of freshly isolated monocytes in complete

medium in the presence of 40 ng/mL of recombinant human (rh) Granulocyte Macrophage-Colony-Stimulating Factor (GM-CSF) plus 40 ng/mL rh Interleukin-4 (IL-4) or 100 ng/mL of rh Macrophage-Colony-Stimulating Factor (M-CSF) (Peprotech), respectively, at 37 °C in humidified atmosphere at 5% CO_2.

Cells differentiating into MoDCs were plated in tissue culture dishes (Falcon) at a density of 1 $x10^6$ cells/mL, while cells differentiating into Mf were plated in 24 well plate (Costar) at a density of 5 $x10^5$ cells/mL. After 5 days of culture, MoDCs were harvested, washed in PBS, resuspended in complete medium without antibiotics and seeded in polypropylene tubes (Falcon) at a density of 1 $x10^6$ cells/tube.

After 6 days of culture, Mf were washed with PBS and 500 µL of complete medium without antibiotics was added to each well.

Single-cell Raman spectra (SCRS) measurements and analysis. At specific time point p.i., bacteria grown in RPMI and cell lysates of THP-1, MoDCs and Mf were washed with PBS, fixed in 4% paraformaldehyde/PBS (PFA, EMS) (v/v) at RT for 15 min and washed twice with Milli-Q water. Samples were then diluted and spotted onto a Raman slide. Individual cells were observed under a 100×/0.8 microscope objective. SCRS were acquired using an HR Evolution confocal Raman microscope (Horiba Jobin-Yvon, France) equipped with a 532 nm neodymium-yttrium aluminum garnet laser. The laser power was attenuated to 4.7 mW by neutral density (ND) filters. Raman scattering was detected by a charge coupled device (CCD) cooled at –68°C. The spectra were acquired in the range of 500 to 3500 cm^{-1} with a 300 grooves/mm diffraction grating. The acquisition time was 10 s per spectrum and at least 30 single bacterial cells were acquired per each condition.

Raman data pre-processing and analysis. Bands associated with Carbohydrates I (382 – 470 cm^{-1}), Carbohydrates II (510 – 600 cm^{-1}), Adenine (705 – 736 cm^{-1}), DNA/RNA (765 – 795 cm^{-1}), Tyrosine (843 – 869 cm^{-1}), Phenylalanine (985 – 1015 cm^{-1}), Carbohydrates III (1020 – 1070 cm^{-1}), Amide III (1208 – 1275 cm^{-1}), Lipid I (1275 – 1357 cm^{-1}), Lipid II (1425 – 1470 cm^{-1}), Amide I (1630 – 1710 cm^{-1}), saturated C-H (2880 – 2950 cm^{-1}) and unsaturated C-H (2990 – 3015 cm^{-1}) were integrated to quantify intracellular biomolecule concentration. The C-D (2070 - 2300 cm^{-1}) and C-H (2800-3030 cm^{-1}) band areas were calculated to determine deuterium content originating from incubation with D_2O.

LC-MS/MS measurements and analysis. At specific time points p.i., MoDCs were lysed by addition of 500 µL of saponin 2% (w/v) in PBS followed by 5 min incubation at 37°C, and bacterial pellets were resuspended in 20 mM ammonium bicarbonate (ABC; Sigma-Aldrich, St.

Louis, MO, USA). Cell solution was then supplemented with 50% trifluoroacetic acid (TFA; Uvasol®, Merck, Darmstadt, Germany) to reach a final concentration of 1% TFA. After 15 min of incubation at 37 °C pH was adjusted to pH 7.0 using 0.5 M ABC. Sequencing grade modified trypsin (0.3 µg; Promega, Madison, WI, USA) was added and incubated for 16 h at 37 °C. Tryptic digestion was stopped by acetic acid (RotiPuran®; Carl Roth GmbH + Co. KG, Karlsruhe, Germany) at a final concentration of 1%. The suspension was centrifuged at 13,000 g for 10 min at 4 °C. Supernatant was collected and purified using ZipTip C18 (Millipore, Schwalbach, Germany). Purified tryptic peptides were dried by lyophilization, dissolved in Buffer A consisting of 2% acetonitrile (J.T. Baker®; part of Fisher Scientific, Waltham, MA, USA), 0.1% acetic acid (Roth) in water (J.T. Baker®), and subjected to LC-MS/MS analysis. Two biological and three technical replicates were performed.

Due to the high pathogenicity of *Salmonella typhi,* we optimized the way to safely kill all bacteria in BSL3 area in order to continue with the peptide preparation outside. Previous studies have shown a strong lytic effect of low pH conditions on gram negative bacteria and kills them rigorous [1,2]. Tests with *Salmonella typhi* incubated with 1% trifluoracetic acid for 15 min didn't show any further bacterial growth and after neutralization, trypsin was efficiently working. Also *Escherichia coli* was treated under identical conditions using a low number of bacteria and was analyzed in mass spectrometry. Using this method, we could identify a sufficient number of proteins and we did not found any significant peptide modifications caused by the use of trifluoracetic acid.

Due to the neutralization with ABC after 1% trifluoracetic acid treatment, the volume already increased and a further add of additional compounds would lead to very high volumes. Additional precipitation steps would be necessary and would yield to a loss of proteins/peptides. In addition, any use of e.g. urea leads to unspecific artificial modification on a various number of peptides [1,2]. This leads to a split of the primary peptide intensity into more variances with lower intensities and further to less identifications, because only a limited number of bacteria and proteins/peptides were available. We therefore tried to limit the chemical treatments to a minimum.

Raw data files from the LC-MS/MS runs were analyzed using MaxQuant version 1.5.3.8. Analysis was performed in strain-specific batches (10 files each for STM-D23580 and ST-Ty2, 9 files for STM-LT2). In all MaxQuant analyses, parameters were set to specific Trypsin/P digestion with maximal 2 missed cleavages. Oxidation of methionine, oxidation of cysteine, deamidation of asparagine, and N-terminal acetylation were chosen as variable modifications with at most 5 modifications allowed per peptide. Fixed modifications were not set. The peptide

mass was limited to 9,900 Da. Label-free quantification (LFQ) with an LFQ minimal ratio count of 2 was selected. Protein identification was supported by alignment between the samples of each analysis batch using the “match between runs” feature with a match time window of 0.7 min and an alignment time window of 20 min. The false discovery rate (FDR) was controlled using a target-decoy approach and limited to 1%. Quantification was performed on normalized data and a minimal ratio count of 2, with unique and razor peptides, including unmodified as well as modified peptides. Other parameters were set to default values. Resulting data of protein identifications were filtered prior to analysis of biological content. Contaminants, false-positive (reverse) hits and protein groups containing more than single proteins were removed. The following analysis was done under an R environment. Log2 transformed label-free quantification (LFQ) intensities were calculated and pairwise comparisons of p.i. 0.75 hr, 2 hr, 4 hr and 6 hr (6 pairs in total) were done using empirical Bayes approach in LIMMA package. Proteins with $p < 0.05$ (moderated t-test) at any pairwise comparison were determined as differentially expressed proteins (DEPs). DEPs were then functionally annotated using Biocyc database.

Supporting Tables

Table S1. Details of LC-MS/MS analysis

Reversed Phase Liquid Chromatography (RPLC)	
instrument	Ultimate 3000 RSLC (Dionex, Idstein, Germany)
trap column	75 µm inner diameter, packed with 3 µm C18 particles (Acclaim PepMap100, Thermo Scientific)
analytical column	Accucore 150-C18: 25 cm x 75 µm, 2.6 µm C18 particles, 150 Å pore size (Thermo Fisher Scientific)
buffer system	binary buffer system consisting of 0.1% acetic acid, 2% ACN (buffer A) and 0.1% acetic acid in 100% ACN in (buffer B)
flow rate	300 nl/min
gradient	linear gradient of buffer B from 2% up to 25%
gradient duration	60 min of linear gradient, total gradient time 95 min
gradient time and percentage of Buffer B	0 min – 2%, 2 min – 5%, 10 min – 5%, 70 min – 25%, 75 min – 40%, 77 min – 90%, 82 min – 90%, 85 min – 2%, 95 min – 2%
column oven temperature	40 °C
Mass Spectrometry (MS)	
instrument	Q Exactive plus mass spectrometer (Thermo Scientific)
electrospray	Nanospray Flex Ion Source (Thermo Scientific)
operation mode	data-dependent
MS scan resolution	70,000
MS ion target value	3E6
Scan range	300-1650 m/z
maximum ion injection time for the MS scan	120 ms
MS spectra data type	profile
selection for MS/MS	10 most abundant isotope patterns with charge ≥2 from survey scan
MS/MS resolution	17,500
MS/MS ion target value	2E5
maximum ion injection time for the MS/MS scans	120 ms
MS/MS spectra data type	centroid
isolation window	3 *m/z*
dissociation mode	higher energy collisional dissociation (HCD)
normalized collision energy	27.5%
peptide match	preferred
charge exclusion	1, >6

Table S2. Biological and vibrational assignments of relevant Raman bands in *Salmonella* spectra

Raman wavenumber centred at (cm^{-1})	**Assignments**	**Vibrational mode**	**Reference**
410	Carbohydrates I	skeletal mode of carbohydrates	[3]
545	Carbohydrates II	C-O-C glycosidic ring deformation	[4]
722	Adenine	Ring breathing	[5]
780	DNA/RNA	Cytosine, uracil ring breathing	[6]
853	Tyrosine	Ring breathing	[5]
1003	Phenylalanine	Ring breathing	[7]
1043	Carbohydrates III	C-O stretching	[8]
1240	Amide III of proteins	40% C-N stretch stretch, 30% N-H bend	[9]
1320	Lipid I	C-H deformation	[10]
1450	Lipid II	CH_2 deformation	[10]
1574	Amide II of proteins	60% N-H bend and 40% C-N stretch	[9]
1662	Amide I of proteins	80% C=O stretch	[9]

Table S3. Confusion matrix and performance evaluation of the kNN classification model classifying 116 STM-LT2 spectra, 152 STM-D23580 spectra and 125 ST-Ty2 spectra.

		Actual Class		
		STM-LT2	STM-D23580	ST-Ty2
Predicted class	STM-LT2	**106**	10	0
	STM-D23580	10	**135**	3
	ST-Ty2	0	7	**122**
Performance evaluation	Sensitivity	**91.4%**	**88.8%**	**97.6%**
	Accuracy		**92.4%**	

Supporting Figures

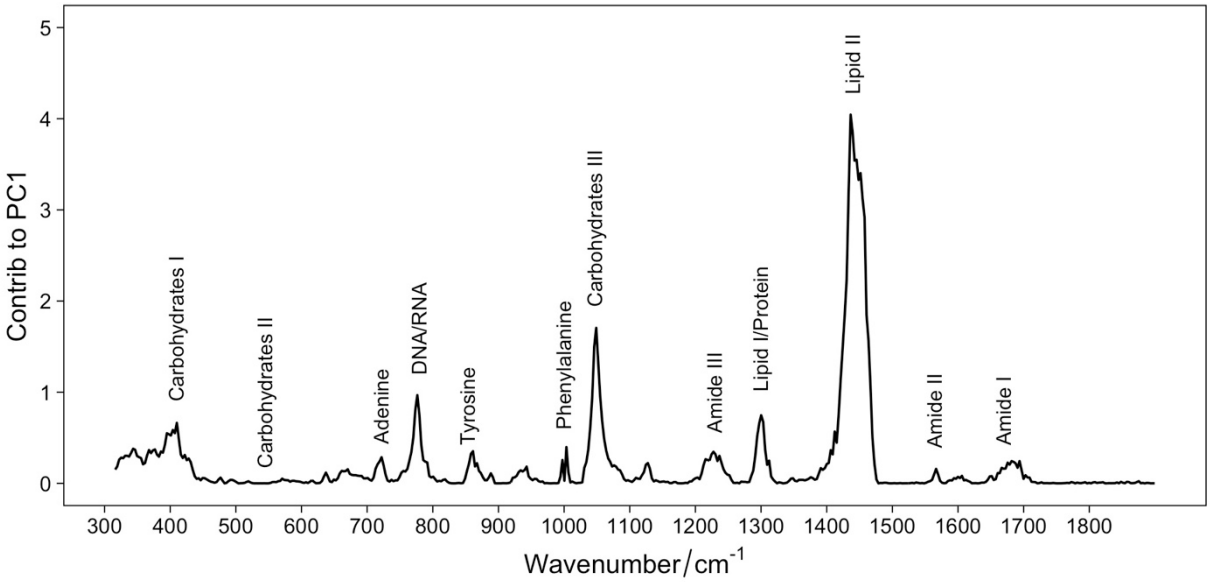

Figure S1. Loading plot of SCRS of STM-LT2, STM-D23580 and ST-Ty2, with Raman wavenumbers against their contributions to PC1 of the PCA. The labelled assignments are according to the wavenumbers in Figure 1A and Table S1. Other unlabelled bands include wavnumbers at around 630 (C-C twisting of aromatic compounds), 671 (ring breathing in DNA bases), 940 (C-C stretching of protein backbone) and 1125 cm^{-1} (C-C stretching of proteins and lipids).

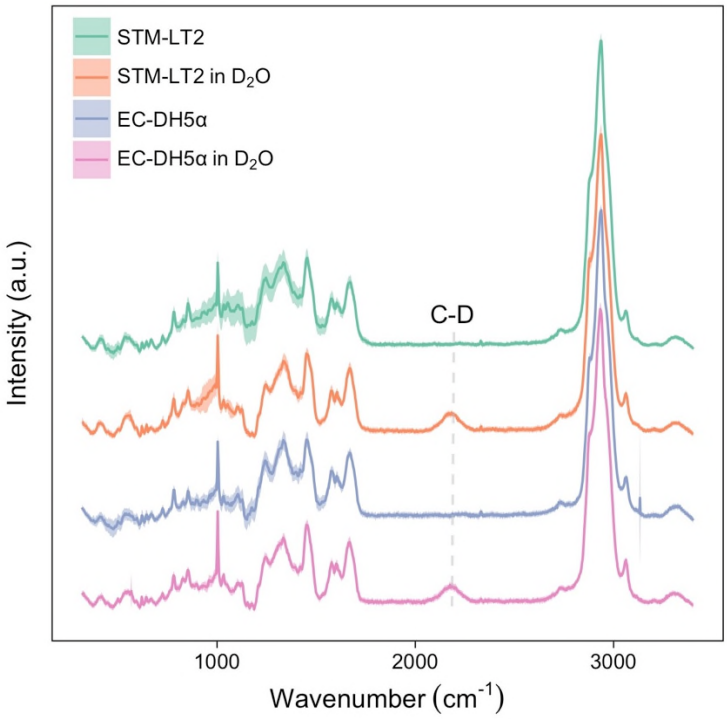

Figure S2. SCRS of STM-LT2 and *E. coli* DH5α (EC- DH5α). Bacterial cultures were grown overnight in LB broth in the presence of 35% D_2O. A broad C-D band indicates incorporation of D during active metabolism. The shaded area represents standard deviation from 30 single-cell measurements.

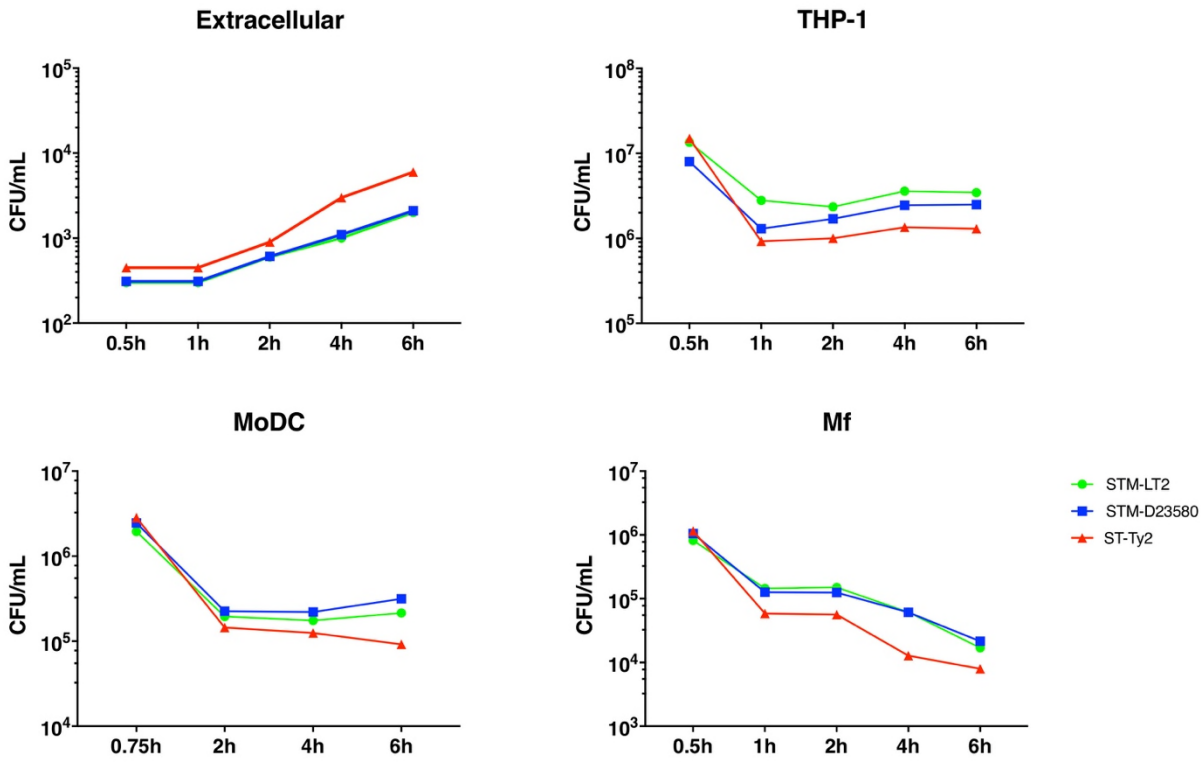

Figure S3. Number of recovered live *Salmonella*, expressed as CFU/mL during (A) extracellular growth of D-labelled bacteria in RPMI and intracellular growth of D-labelled bacteria in (B) THP-1, (C) Monocyte-derived Dendritic cell (MoDC) and (D) macrophages (Mf). An average of two experiments is shown.

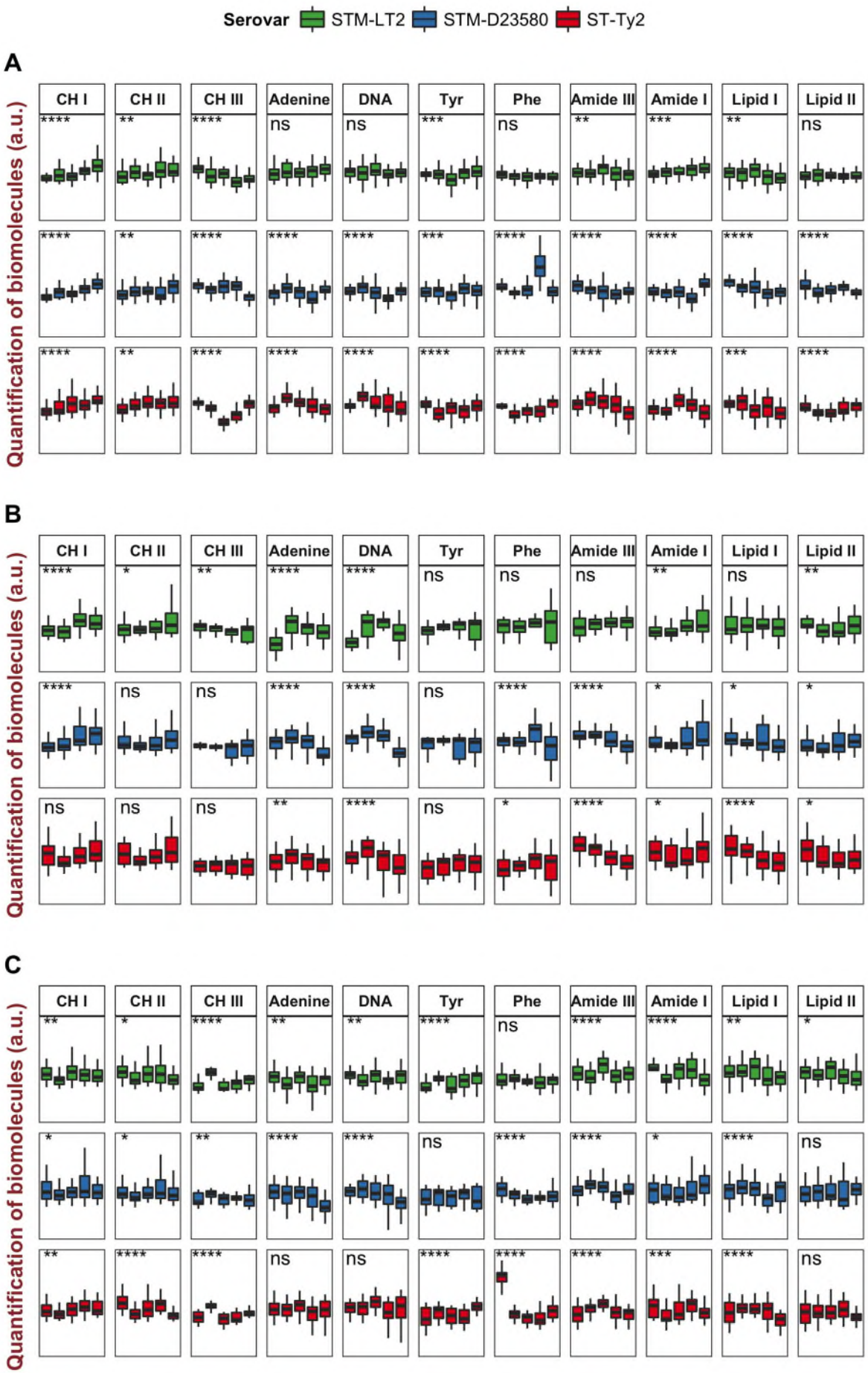
Serovar
STM-LT2
STM-D23580
ST-Ty2
A
B
C
Quantification of biomolecules (a.u.)
CH I
CH II
CH III
Adenine
DNA
Tyr
Phe
Amide III
Amide I
Lipid I
Lipid II

Figure S4. Complete Raman profiling of biomolecules in intracellular *Salmonella* inside (A) THP-1 cells, (B) MoDCs and (C) Mf, based on their single-cell Raman spectra.

Boxplots showing semi-quantification of selected biomolecules by integrating relevant Raman bands in fingerprint region of SCRS of STM-LT2, STM-D23580 and ST-Ty2 at different time point of infection in (A) THP-1 cells, (B) dendritic cells and (C) macrophages. Statistics was done by using one-way ANOVA with a global p-value to indicate general differences among time points (ns: $p > 0.05$; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****; $p < 0.0001$).

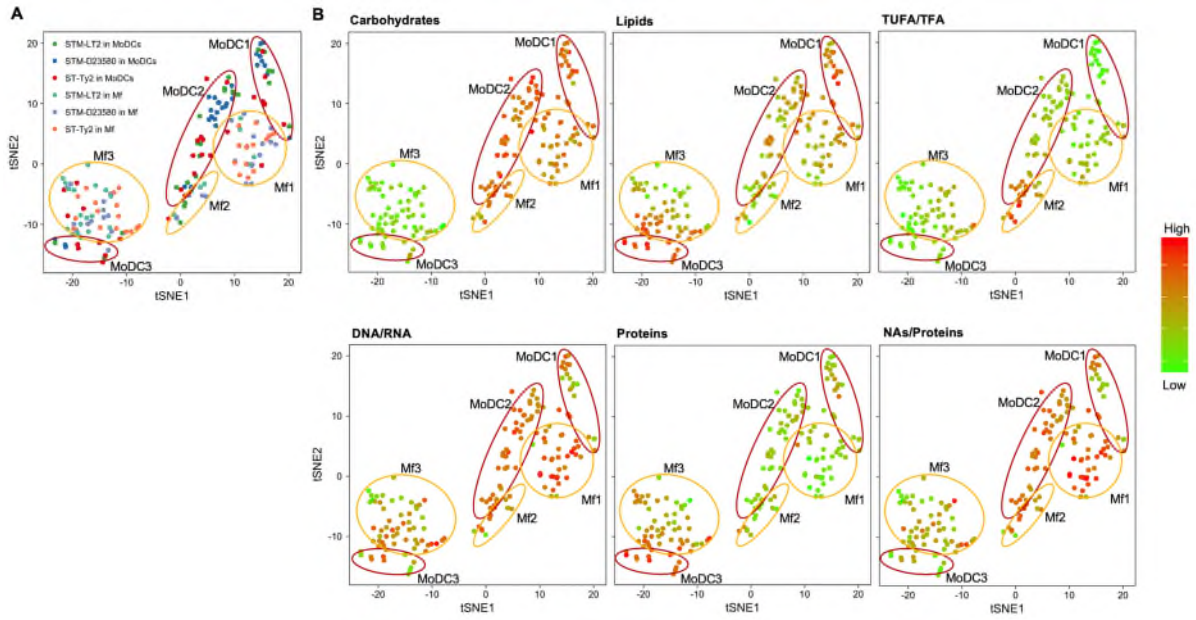

Figure S5. (A) A t-SNE plot of single cells (perplexity = 10) computed based on SCRS of STM-LT2, STM-D23580 and ST-Ty2 at 2 hr p.i. of MoDCs and Mf. Three bacterial subpopulations were clustered in MoDCs (MoDC1, MoDC2 and MoDC3) or in Mf (Mf1, Mf2 and Mf3). (B) t-SNE plots colored based on quantification of different biomolecules. The intensity gradient is mapped from green (low) to red (high).

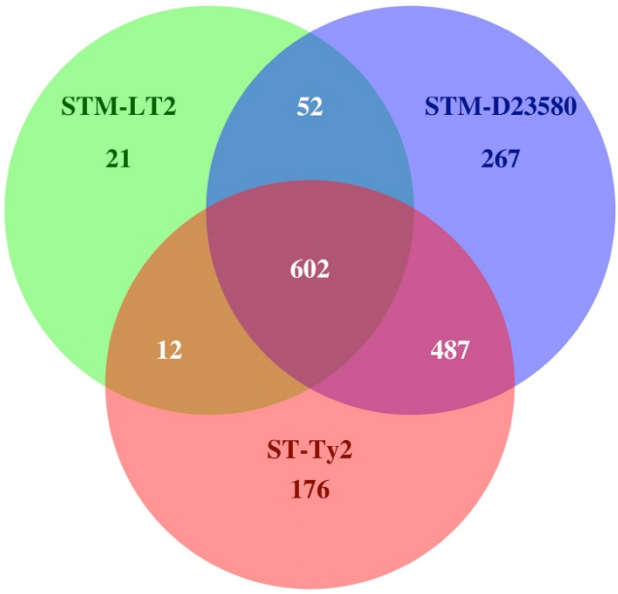

Figure S6. Venn diagram of proteins detected by LC in STM-LT2, STM-D23580 and ST-Ty2 during infection in human dendritic cells.

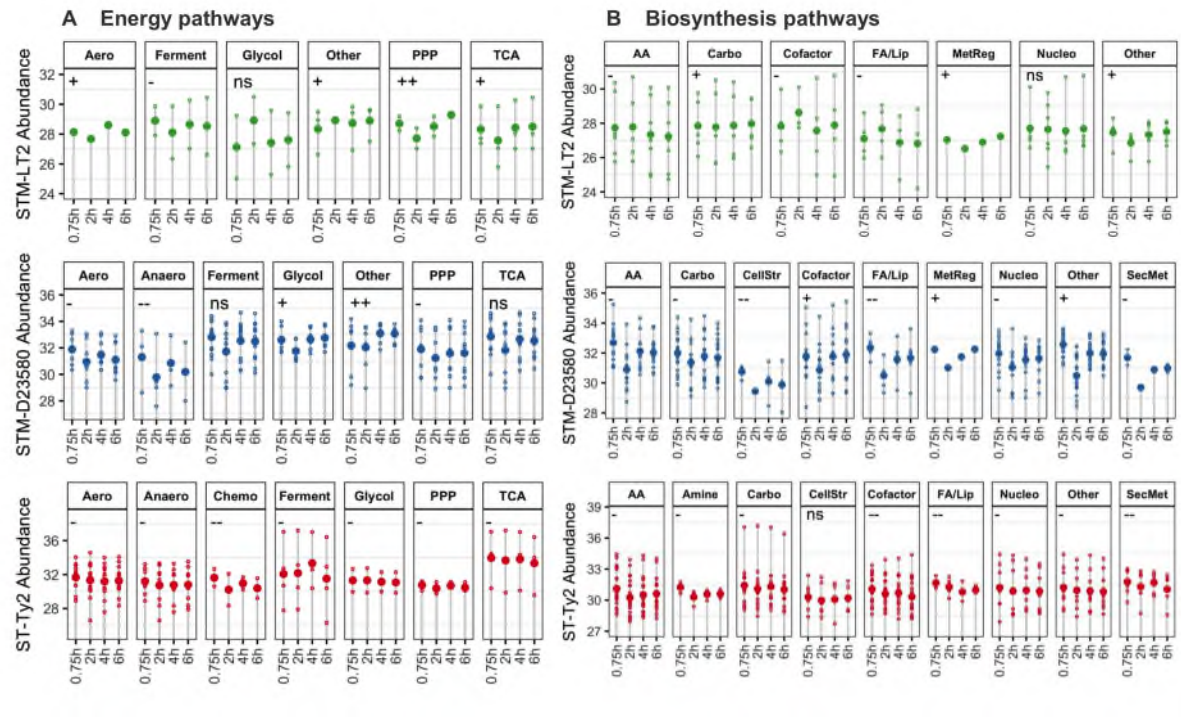

Figure S7. Dotplots of abundance of DEPs as log2 of the LFQ related to (A) energy and (B) biosynthesis pathways in STM-LT2 (green), STM-D23580 (blue) and ST-Ty2 (red). The number of DEPs related to energy pathways is 65 in STM-LT2, 92 in STM-D23580 and 94 in ST-Ty2, respectively. The number of DEPs related to biosynthesis pathways is 44 in STM-LT2, 82 in STM-D23580 and 62 in ST-Ty2, respectively.

The large dot represents the average of all log2 protein expression values belonging to a specific biosynthesis pathway at one time point. Each of the small dots represents a data value for an individual protein within the pathway. Time-wise changes were calculated as linear regression slopes and represented as "++" and "+" being upregulated ($>10^{-1}$ and $1>0^{-2}$), "--" and "-" being downregulated ($<-10^{-1}$ and $<-10^{-2}$), and "ns" being non-significant ($<10^{-3}$) over time.

Abbreviations of the pathways are listed as follows: "AA" stands for "Amino Acid Biosynthesis"; "Amine" stands for "Amines and Polyamines Biosynthesis"; "Carbo" stands for "Carbohydrates Biosynthesis"; "CellStr" stands for "Cell Structures Biosynthesis"; "Cofactor" stands for "Cofactors, Prosthetic Groups, Electron Carriers Biosynthesis"; "FA/Lip" stands for "Fatty Acids and Lipids Biosynthesis"; "MetReg" stands for "Metabolic Regulators Biosynthesis"; "Nucleo" stands for "Nucleosides and Nucleotides Biosynthesis"; "Other" stands for "Other Biosynthesis"; "SecMet" stands "Secondary Metbolites Biosynthesis".

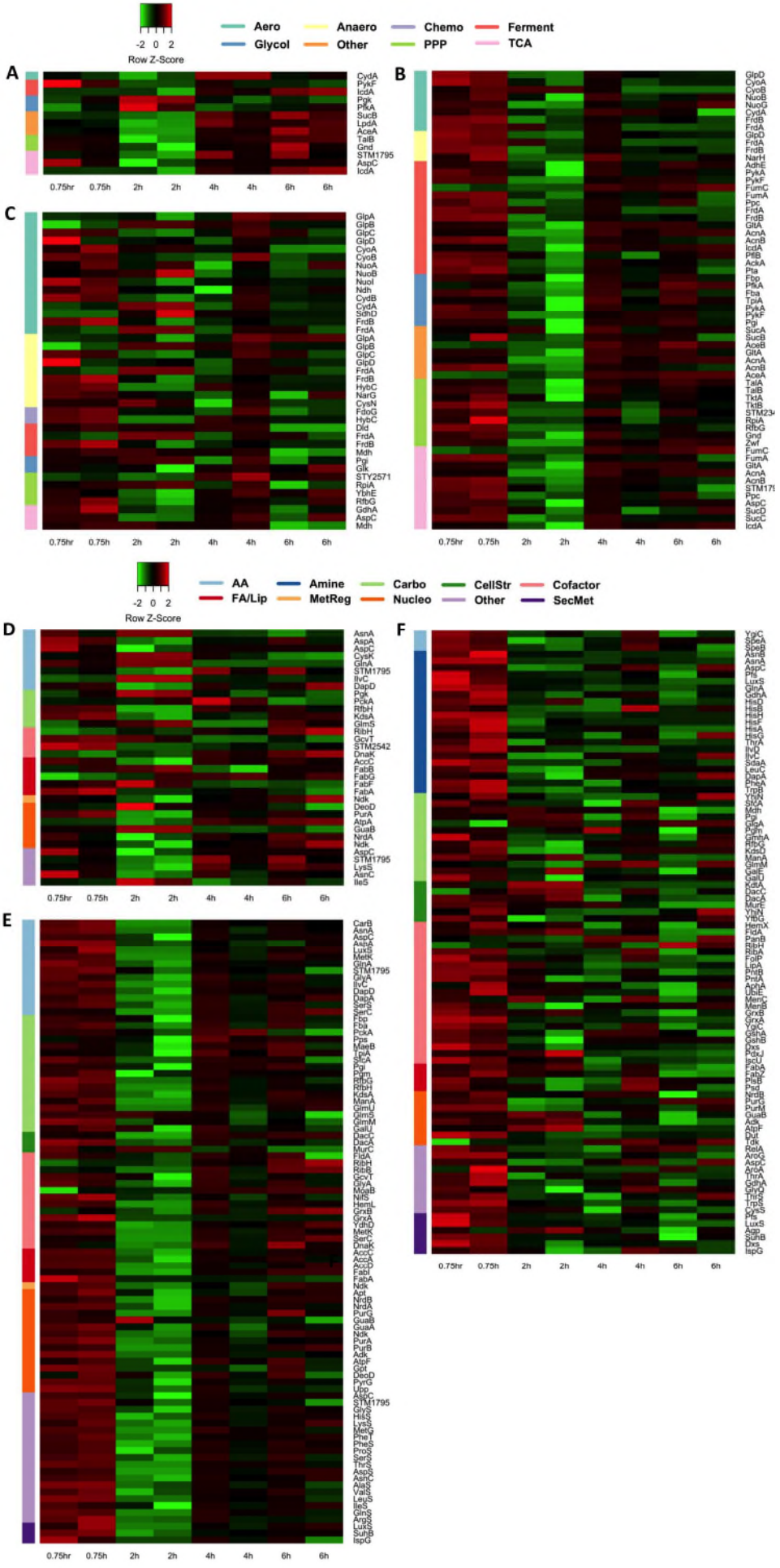
-2 0 2
Row Z-Score
Aero
Glycol
Anaero
Other
Chemo
PPP
Ferment
TCA
A
B
C
D
E
F
AA
FA/Lip
Amine
MetReg
Carbo
Nucleo
CellStr
Other
Cofactor
SecMet
0.75hr 0.75h 2h 2h 4h 4h 6h 6h

Figure S8. Heatmaps of row-normalized abundance of differentially expressed proteins (DEPs) related to energy pathways in (A) STM-LT2, (B) STM-D23580 and (C) ST-Ty2 or biosynthesis pathways in (D) STM-LT2, (E) STM-D23580 and (F) ST-Ty2. Color bars on the left present assignment of protein to one specific energy pathway and individual protein names are listed on the right.

Abbreviations of the pathways are listed as follows: “Aero” stands for “Aerobic Respiration”; “Anaero” stands for “Anaerobic Respiration”; “Chemo” stands for “Chemoautotrophic Energy Metabolism”; “Ferment” stands for “Fermentation”; “Glycol” stands for “Glycolysis”; “Other” stands for “Other Energy Pathways”; “PPP” stands for “Pentose Phosphate Pathway”; “TCA” stands for “The Citric Acid Cycle”; “AA” stands for “Amino Acid Biosynthesis”; “Amine” stands for “Amines and Polyamines Biosynthesis”; “Carbo” stands for “Carbohydrates Biosynthesis”; “CellStr” stands for “Cell Structures Biosynthesis”; “Cofactor” stands for “Cofactors, Prosthetic Groups, Electron Carriers Biosynthesis”; “FA/Lip” stands for “Fatty Acids and Lipids Biosynthesis”; “MetReg” stands for “Metabolic Regulators Biosynthesis”; “Nucleo” stands for “Nucleosides and Nucleotides Biosynthesis”; “Other” stands for “Other Biosynthesis”; “SecMet” stands for “Secondary Metbolites Biosynthesis”.

SI References

(1) Alakomi, H. L.; Skytta, E.; Saarela, M.; Mattila-Sandholm, T.; Latva-Kala, K.; Helander, I. M. *Applied and Environmental Microbiology* **2000**, *66*, 2001-2005.
(2) Kollipara, L.; Zahedi, R. P. *Proteomics* **2013**, *13*, 941-944.
(3) Góral, J.; Zichy, V. *Spectrochimica Acta Part A: Molecular Spectroscopy* **1990**, *46*, 253-275.
(4) Maquelin, K.; Kirschner, C.; Choo-Smith, L. P.; van den Braak, N.; Endtz, H. P.; Naumann, D.; Puppels, G. J. *J Microbiol Methods* **2002**, *51*, 255-271.
(5) De Gelder, J.; De Gussem, K.; Vandenabeele, P.; Moens, L. *Journal of Raman Spectroscopy* **2007**, *38*, 1133-1147.
(6) Ruiz-Chica, A. J.; Medina, M. A.; Sánchez-Jiménez, F.; Ramírez, F. J. *Journal of Raman Spectroscopy* **2004**, *35*, 93-100.
(7) Hernández, B.; Pflüger, F.; Kruglik, S. G.; Ghomi, M. *Journal of Raman Spectroscopy* **2013**, *44*, 827-833.
(8) Xu, J.; Webb, I.; Poole, P.; Huang, W. E. *Anal Chem* **2017**, *89*, 6336-6340.
(9) Rygula, A.; Majzner, K.; Marzec, K. M.; Kaczor, A.; Pilarczyk, M.; Baranska, M. *Journal of Raman Spectroscopy* **2013**, *44*, 1061-1076.
(10) Czamara, K.; Majzner, K.; Pacia, M. Z.; Kochan, K.; Kaczor, A.; Baranska, M. *Journal of Raman Spectroscopy* **2015**, *46*, 4-20.