Infrared spectroscopy coupled with a dispersion model for quantifying the real-time dynamics of kanamycin resistance in artificial microbiota

Journal:	*Analytical Chemistry*
Manuscript ID	ac-2017-01765u.R2
Manuscript Type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Jin, Naifu; Lancaster University, LEC Paraskevaidi, Maria; University of Central Lancashire, School of Pharmacy and Biomedical Sciences Semple, Kirk; Lancaster University, Department of Environmental Science Martin, Francis; University of Central Lancashire, School of Pharmacy and Biomedical Sciences Zhang, Dayi; Lancaster University,

SCHOLARONE™
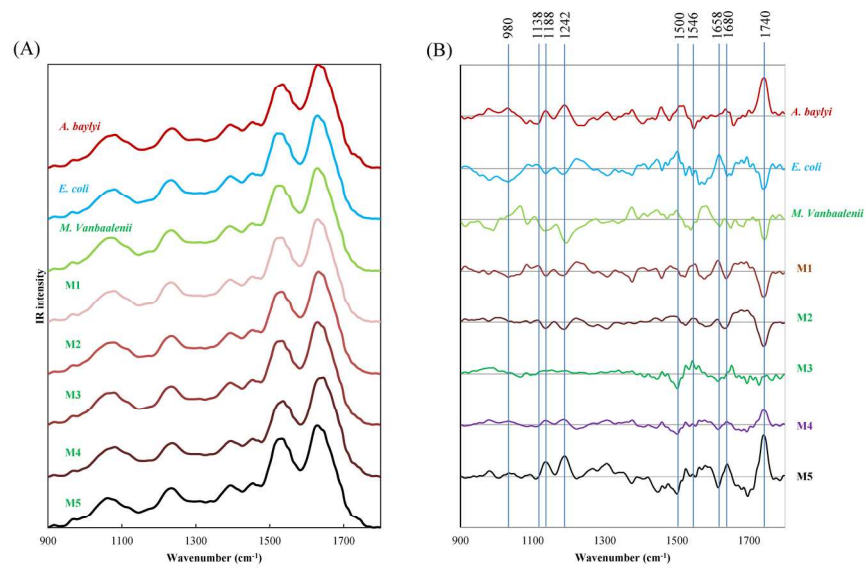Manuscripts
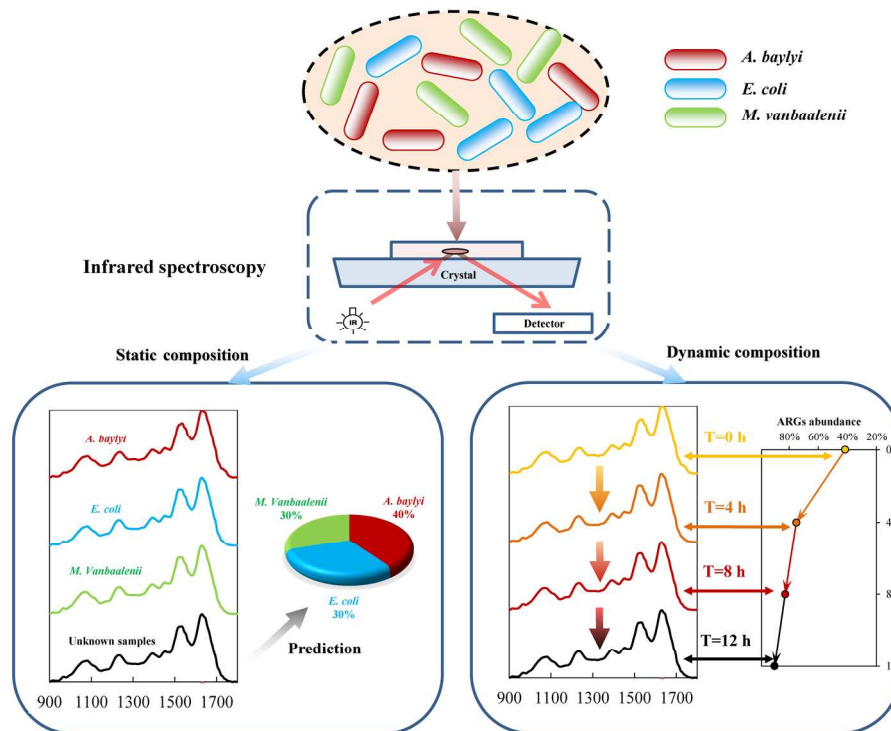
ToC graphic

190x142mm (300 x 300 DPI)

Figure 1

190x142mm (300 x 300 DPI)
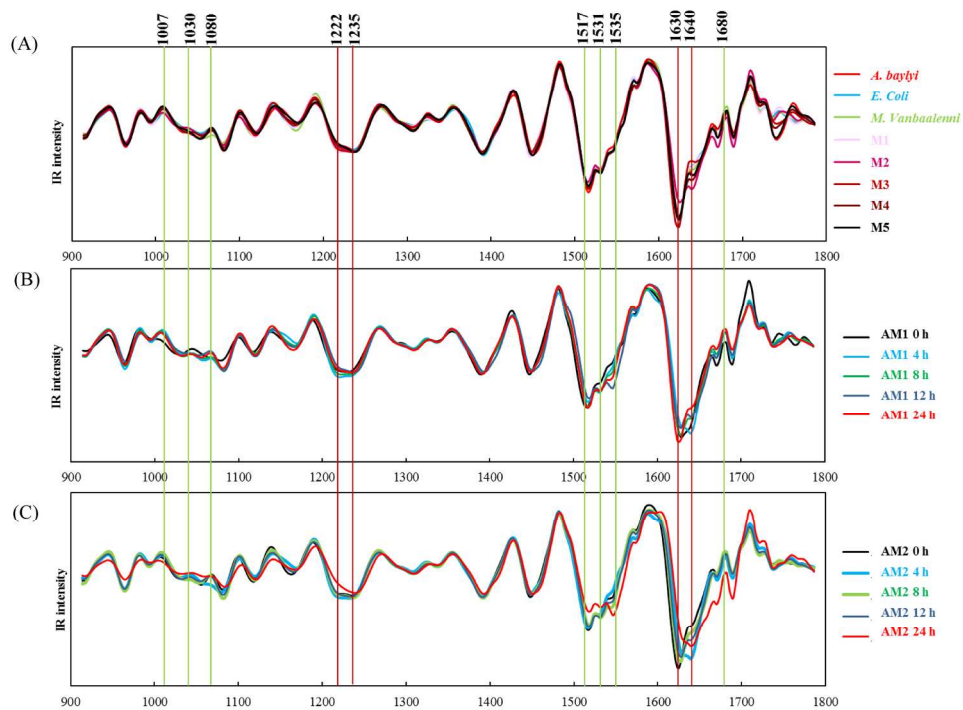
Figure 2

190x142mm (300 x 300 DPI)
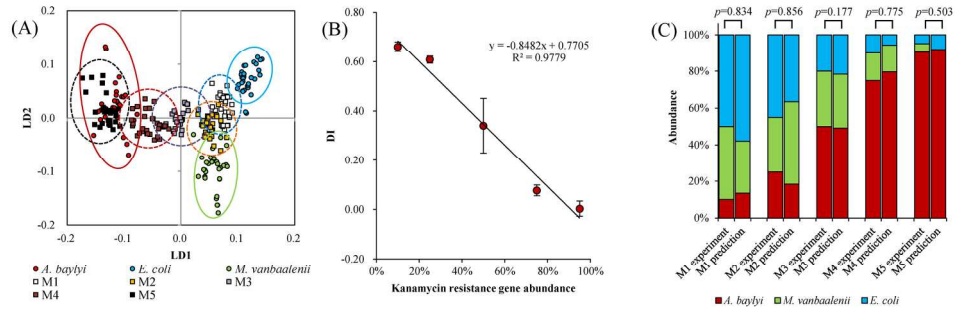
Figure 3

190x142mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

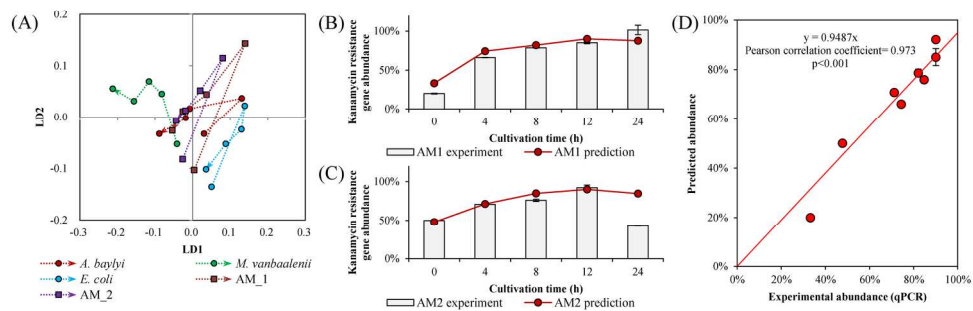

Figure 4

190x142mm (300 x 300 DPI)

1   **Infrared spectroscopy coupled with a dispersion model for quantifying the**

2   **real-time dynamics of kanamycin resistance in artificial microbiota**

3   Naifu Jin[1], Maria Paraskevaidi[2], Kirk T. Semple[1], Francis L. Martin[2,*], Dayi Zhang[1,*]

4   [1]Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, UK;

5   [2]School of Pharmacy and Biomedical Sciences, University of Central Lancashire,

6   Preston PR1 2HE, UK

7

8

9

10

11

12

13   *Corresponding authors:*

14   Dayi Zhang, Lancaster Environment Centre, Lancaster University, Lancaster LA1

15   4YQ, UK; Tel.: +44(0)1524 510288; Fax: +44(0)1524 510082, Email:

16   d.zhang@lancaster.ac.uk

17   Francis L. Martin, School of Pharmacy and Biomedical Sciences, University of

18   Central Lancashire, Preston PR1 2HE, UK; Tel.: +44(0)1772 896482; Email:

19   flmartin@uclan.ac.uk

20

21 **Abstract**

22 Over-usage of antibiotics leads to the widespread induction of antibiotic resistance

23 genes (ARGs). Developing an approach to allow real-time monitoring and fast

24 prediction of ARGs dynamics in clinical or environmental samples has become an

25 urgent matter. Vibrational spectroscopy is potentially an ideal technique towards the

26 characterization of the microbial composition of microbiota as it is non-destructive,

27 high-throughput and label-free. Herein, we employed attenuated total reflection

28 Fourier-transform infrared (ATR-FTIR) spectroscopy and developed a

29 spectrochemical tool to quantify the static and dynamic composition of kanamycin

30 resistance in artificial microbiota to evaluate microbial antibiotic resistance. Second

31 order differentiation was introduced in identifying the spectral biomarkers, and

32 principal component analysis followed by linear discriminant analysis (PCA-LDA)

33 was used for the multivariate analysis of the entire spectral features employed. The

34 calculated results of the mathematical dispersion model coupled with PCA-LDA

35 showed high similarity to the designed microbiota structure, with no significant

36 difference ($P > 0.05$) in the static treatments. Moreover, our model successfully

37 predicted the dynamics of kanamycin resistance within artificial microbiota under

38 kanamycin pressures. This work lends new insights into the potential role of

39 spectrochemical analyses in investigating the existence and trends of antibiotic

40 resistance in microbiota.

41

44

2

Introduction

45

46 Antibiotics have played a vital role in modern medicine contributing to a considerable

47 reduction in childhood mortality and increasing life expectancy[1]. However, the

48 increasing number of fatal infections caused by antibiotic-resistant bacteria is

49 gradually developing into a global threat. The environment has become the primary

50 "sink" for most applied antibiotics and their residues arising from human or animal

51 excretion[1-3]. Since bacteria with antibiotic resistance genes (ARGs) can tolerate

52 antibiotics, selection pressures from contaminated water or soil will boost the

53 abundance of ARGs in the environment and increase the possibility of their spread

54 through microbial species[4,5]. Therefore, real-time monitoring and quantification of

55 ARGs or antibiotic-resistant bacteria is urgently required.

56   Besides measuring the concentration of antibiotics *via* chemical analysis, various

57 biological analytical methods have been used to determine the presence, abundance

58 and diversity of ARGs in the microbiota to capture a "static map" of their existence,

59 *e.g.*, meta-sequence and quantitative polymerase chain reaction (qPCR)[6,7]. However,

60 genetically identical cells from the same population have stochasticity in gene

61 expression, meaning that there is significant variation in their molecular content and

62 phenotype, even under similar environmental influences. Moreover, bacterial

63 resistance to the antibiotics can also be affected and regulated epigenetically[8]. In

64 combination, these factors provide an opportunity for phenotypic and cell-type

65 diversity regardless of genotype[9]. This questions the reliability of determining ARGs

66 abundance by molecular biological approaches in real-world situations, leading to the

67 necessity of developing a phenotypic assay that depicts *in situ* dynamics of ARGs or

68 microbial antibiotic resistance in environmental samples.

69   It is well accepted that genetic and epigenetic factors cannot be studied

70 independently as a complete phenotype emerges from both together[10]. The

71 spectrochemical analysis is an alternative approach to characterize the phenotypic

72 features of organisms and has already demonstrated its ability to investigate clinical

3

73 samples, as well as to describe and identify bacterial species[11,12]. Previous studies

74 indicates that spectroscopic techniques are capable of studying phenotypic features, at

75 either population[13] or single-cell[14] level, such as diagnosing the distinct spectral

76 signatures and metabolomes from isogenic cell lines[15]. However, the current

77 techniques have limited application in characterizing ARGs under antibiotic pressures,

78 mainly due to the lack of appropriate analytical models and well-trained databases.

79 Recently, some studies using spectroscopic techniques have set out to investigate

80 biological response to environmental stress, like nanomaterials[16,17] and antimicrobial

81 reagents[18,19]. The introduction of spectroscopic techniques coupled with a suitable

82 prediction model to characterize microbial composition may bring new insights in

83 detecting the presence or even the dynamics of microbial antibiotic resistance in

84 environmental microbiota in real-time, owing to its non-destructive, high-throughput

85 and label-free character[20,21]. It also allows for *in situ* spectral measurements, helping

86 in understanding the interactions between microbes and their physical environment.

87 Kanamycin is a subclass of aminoglycoside antibiotics, one of the most widely

88 applied antibiotics in health and molecular biology[22]. Because of the well-established

89 mechanisms of kanamycin resistance and characterized sequence[23], it was selected as

90 the model antibiotic in the present study. Herein, we used attenuated total reflection

91 Fourier-transform infrared (ATR-FTIR) spectroscopy, coupled with the multivariate

92 analysis and the dispersion indicator model, to quantify the kanamycin resistance

93 within artificial microbiota and evaluate their phenotypic change associated with

94 kanamycin resistance, from both static and dynamic perspectives. This work raises the

95 potential feasibility of applying spectroscopic techniques to diagnose ARGs

96 phenotypic dynamics in the microbial community *in situ*.

97

Experimental section

Sample preparation

100 The present study included two strains without kanamycin-resistant-gene,

101 *Mycobacterium vanbaalenii* PYR-1 and *Escherichia coli* DH5α, and one

4

102 kanamycin-resistant strain *Acinetobacter baylyi* ADPWH_recA, which has a

103 continuously expressed kanamycin resistance gene *kan*[R] (from Mini-Tn5/Km[24],

104 Genbank accession number: U32991.1) inserted into the *recA* gene in the

105 chromosome of *A. baylyi* ADP1[25]. Before the experiment, they were all cultured in

106 Luria-Bertani (LB) broth medium for 24 h at 30 ± 2°C.

107   The three control groups contained pure *M. vanbaalenii* PYR-1, *E. coli* DH5α

108 and *A. baylyi* ADPWH_recA, respectively. The artificial microbiotas were prepared

109 for both static (M1 to M5) and dynamic (AM1 and AM2) experiments by gently

110 mixing the cells in the compositions listed in Table 1. The optical density at 600 nm

111 (OD_{600}) in each treatment was monitored continuously for 24 h by a multimode plate

112 reader (FLUOstar Omega, Germany) to evaluate bacterial growth. For static tests, the

113 cells were directly collected by centrifugation (4000 rpm for 5 min), washed three

114 times with sterile deionized water to remove the residues of growth media and then

115 suspended in 70% ethanol to fix the bacterial cells. For dynamic tests, all the artificial

116 microbiotas were treated with kanamycin (final concentration 10 mg/L). After

117 exposure for 4, 8, 12 or 24 h, the cells from microbiotas were harvested following the

118 same procedure as above.

119

120 **Table 1.** The compositions of artificial microbiotas (volume ratio, v:v:v).

Treatments	Control			Static test					Dynamic test	
	M. vanbaalenii	*E. coli*	*A. baylyi*	M_1	M_2	M_3	M_4	M_5	AM_1	AM_2
M. vanbaalenii	100%	-	-	40%	30%	30%	15%	5%	40%	25%
E. coli	-	100%	-	50%	45%	20%	10%	5%	40%	25%
A. baylyi	-	-	100%	10%	25%	50%	75%	90%	20%	50%

121

122 *ATR-FTIR spectroscopy*

123 The washed cell pellets (minimal amount >5 μL) were applied onto Low-E slides for

124 the interrogation by ATR-FTIR spectroscopy. A TENSOR 27 FTIR spectrometer

125 (Bruker Optics Ltd., UK) equipped with a Helios ATR attachment (containing a

5

126   diamond internal IRE; incidence angle of the IR beam: 45°) was used. Instrument

127   parameters were set at 32 scans and spatial resolution of 8 cm^{-1}. Before the

128   measurement of a new sample, the crystal was cleaned with deionized water and

129   background readings were retaken. A total of 30 spectra were randomly acquired for

130   each treatment (3 replicates).

131   *Computational analysis*

132   The primary analysis methods employed in this study involved multivariate analysis

133   and the dispersion indicator model. The initial data generated from ATR-FTIR

134   spectroscopy were analyzed within MATLAB R2011a (*TheMathsWorks, Natick, MA,*

135   *USA*) software, coupled with IRootLab toolbox (http://irootlab.googlecode.com)[26].

136   Unless otherwise stated, the acquired spectra were cut to the biochemical-cell

137   fingerprint region (1800-900 cm^{-1}), rubberband baseline corrected and normalized to

138   Amide I (1650 cm^{-1}). Second order differentiation baseline correction and vector

139   normalization was also performed as an alternative mean to process the data (the

140   number of the filter coefficients of the Savitzky-Golay smoothing/differentiation filter

141   was 9). Principal component analysis followed by linear discriminant analysis

142   (PCA-LDA) was subsequently applied to the pre-processed data to reduce the number

143   of spectra to 10 uncorrelated principal components (PCs), which account for >99% of

144   the total variance; LDA is a supervised technique coupled with PCA in order to

145   maximize inter-class and minimize intra-class variance[21]. In addition, cluster vector

146   approach was conducted to visualise the discriminating difference[21,27]. This method

147   takes input from PCA-LDA to create a loadings vector for each category contributing

148   to respective data points. The pseudo-spectra allow identifying which variables (or

149   wavenumber) are responsible for variance in the data set related to the original

150   spectra[21,27]. The detailed information of the dispersion indicator model was described

151   in the Electronic Supplementary Information (ESI).

152   *Biological analysis*

153   The copy numbers of total bacterial 16S rRNA and targeted kanamycin resistance

154   gene (*kan^R*) were determined by quantitative polymerase chain reaction (qPCR). For

155   16S rRNA, the primer pair set was 341F (5'-CCTACGGGNGGCWGCAG-3') and

156   805R (5'-GACTACHVGGGTATCTAATCC-3'), and the primer pair for *kan^R* was

157   KanF (5'-TGTCATACCACTTGTCCGCC-3') and KanR

6

158   (5'-ATCGAGCTGTATGCGGAGTG-3'). The 20 µL qPCR system consisted of 2 µL

159   of each primer, 1 µL DNA template, 5 µL molecular water and 10 µL iTaq™

160   Universal SYBR® Green Supermix (BioRad, USA). The relative abundance of *kanR*

161   in each pure strain was calculated as the ratio of *kanR* copy numbers to 16S rRNA

162   copy numbers (kanR/16S). The microbial kanamycin resistance within the artificial

163   microbiota was calculated as the ratio of *A. baylyi* population to the total bacterial

164   population.

165   *Statistical analysis*

166   One-way analysis of variance (ANOVA) with Tukey's post hoc test/or T-test was

167   employed to examine the discriminating differences. All statistical analysis was

168   carried out in GraphPad Prism 6.

169

170   **Results and Discussion**

171   *Growth and kanamycin resistance gene of individual strains*

172   All the three bacterial strains (*A. baylyi* ADPWH_recA, *M. vanbaalenii* PYR-1 and *E.*

173   *coli* DH5α) had similar growth curves without kanamycin pressure (see ESI Figure

174   S1A). Cultivated in 10 mg/L kanamycin, only *A. baylyi* ADPWH_recA maintained

175   positive growth because of the expression of *kan^R* gene and resistance to kanamycin

176   (see ESI, Figure S1B). Neither *M. vanbaalenii* PYR-1 nor *E. coli* DH5α grew

177   post-exposure to 10 mg/L kanamycin. The results of qPCR further confirmed that the

178   high relative abundance of *kan^R* gene (kanR/16S) were only found in *A. baylyi*

179   ADPWH_recA (0.306 in medium without kanamycin and 0.275 in medium with 10

180   mg/L kanamycin respectively, no significant difference), whereas it was less than

181   0.001 or below the limit of detection for *M. vanbaalenii* PYR-1 or *E. coli* DH5α (see

182   ESI, Figure S2). It was further proved that kanamycin resistance gene is only

183   detectable in *A. baylyi* ADPWH_recA, but neither *M. vanbaalenii* PYR-1 nor *E. coli*

184   DH5α, and the latter two cannot tolerate kanamycin pressure. The active group of

185   kanamycin, 2-deoxystreptamine, impairs bacterial protein synthesis through binding

186   to prokaryotic ribosomes 30S subunit [22]. The *kan^R* encoding neomycin

187   phosphotransferase is an aminoglycoside-modifying enzyme, using ATP as donor to

188   modify the hydroxyl functions of 2-deoxystreptamine and inhibit its binding to

7

189   ribosomes[28]. The *kan*[R] gene is therefore a reliable molecular indicator in detecting the

190   kanamycin resistance.

IR spectral fingerprints of individual strains and microbiotas

192   The IR spectral fingerprint region (1800 - 900 cm[-1]) of the three strains and artificial

193   microbiotas are shown in Figure 1. The representative peaks of the biochemical

194   fingerprint include lipids (~1750 cm[-1]), Amide I (~1650 cm[-1]), Amide II (~1550 cm[-1]),

195   Amide III (~1260 cm[-1]), carbohydrate (~1155 cm[-1]), asymmetric phosphate stretching

196   vibrations ($v_{as}PO_2^-$; ~1225 cm[-1]), symmetric phosphate stretching vibrations ($v_sPO_2^-$;

197   ~1080 cm[-1]), glycogen (~1030 cm[-1]) and protein phosphorylation (~970 cm[-1])[20 21].

198   Past literatures[12,20,29,30] suggest the characteristic peaks given by the region can be

199   used as biomarkers to characterize microbial cell types (even at subspecies level) and

200   diagnose microbe-induced diseases.

201   However, the visual spectral differences with the mean spectra are almost

202   identical regardless of the bacterial species or community composition. For this

203   reason, we applied the cluster vectors after multivariate analysis (PCA-LDA) and the

204   second order differentiation baseline correction to further reveal the underlying

205   biochemical differences between each strain or microbiota. Based on the derived

206   spectral biomarkers from PCA-LDA (Figure 1B), all the microbiota samples showed

207   marked segregation (see ESI, Table S1). Characteristics associated with microbial

208   composition were observed in particular wavenumber-absorbance intensities. For

209   instance, the intensities at 980 cm[-1] and 1740 cm[-1] were increased with increasing

210   ratio of ARGs but fluctuated in some artificial microbiotas, particularly for microbiota

211   M3 (*M. vanbaalenii* PYR-1: *E. coli* DH5α: *A. baylyi* ADPWH_recA =

212   30%:20%:50%). Additionally, IR spectral analysis (Figure 2A) based on the second

213   order differentiation baseline correction and vector normalization highlighted several

214   key biomarkers. Two apparent shifts from ~1630 cm[-1] to ~1640 cm[-1] (Amide I) and

215   from ~1222 cm[-1] to ~1235 cm[-1] ($v_{as}PO_2^-$) associated with *A. baylyi* were regarded as

216   biomarkers for the presence of kanamycin resistance. These spectral alterations might

217   be attributed to the upregulated activities of the *kan*[R] encoding aminoglycoside

218   *O*-phosphotransferase, which contributes to microbial resistance by inactivating

219   kanamycin molecular via catalyzing ATP-dependent phosphorylation of specific

220   aminoglycoside hydroxyl groups[31]. Some other weaker discriminations included the

8

221   polysaccharide fingerprint region (1000-1150 cm^{-1}) and the protein absorbance region

222   (1500-1700 cm^{-1})[27]. These alterations were probably induced by the interference of

223   extracellular polymeric substances (EPS) produced by different species[32-34] and

224   resulted in the difficulties in distinguishing biomarkers from the PCA-LDA extracted

225   peaks. Based on the previous studies[32,35,36], we speculate that these extracellular

226   materials may interact with each other and generate new biochemical compositions

227   within the communities, influencing the discriminating peaks obtained with

228   spectrochemical interrogation.

229
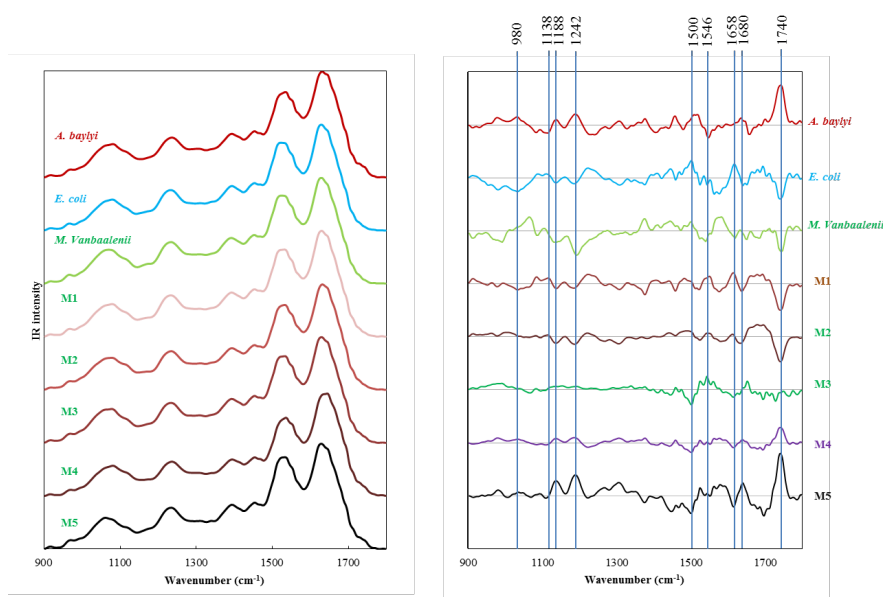
230

231   **Figure 1.** (A) Infrared spectra of *A. baylyi*, *M. vanbaalenii*, *E. coli* and five artificial

232   microbiotas (M1-M5). (B) Cluster vector plots after PCA-LDA, indicating significant

233   wavenumbers for the segregation between bacterial species and artificial microbiotas.

9
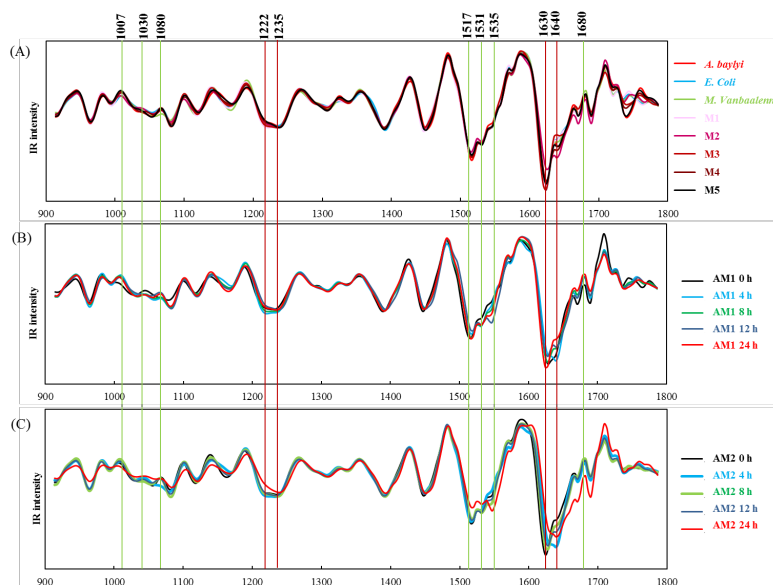
Figure 2. Class means spectra of pre-processed data based on second order differentiation baseline correction and vector normalization. (A) Processed spectra of *A. baylyi*, *M. vanbaalenii*, *E. coli* and five artificial microbiotas (M1-M5). (B) Processed spectra of AM1 at different time point in dynamic experiment. (C) Processed spectra of AM1 at different time point in dynamic experiment.

Predicting community composition in artificial microbiotas

Comparing to the IR spectra in the static tests, we observed identical spectral biomarkers in artificial community dynamics (Figure 2B and 2C) that the same shifts from ~1630 cm^{-1} to ~1640 cm^{-1} (Amide I) and from ~1222 cm^{-1} to ~1235 cm^{-1} ($v_{as}PO_2^-$) developed along with the time. The results indicated the consistent spectral biomarkers in both static and dynamic microbiotas in analyzing the phenotypic presence and abundance of kanamycin resistance gene in the targeted microbiota.

The PCA-LDA scores plot (Figure 3A) also illustrates a significant segregation of the different groups, associated with differing microbiota compositions. The control groups (*M. vanbaalenii*, *E. coli,* and *A. baylyi*) are clearly separated from each other. In contrast to *M. vanbaalenii* and *E. coli*, all the converted spectral values of *A. baylyi* are aligned as negative along linear discriminant one (LD1), likely attributed to its kanamycin resistance. Meanwhile, along with linear discriminant two (LD2), the group of *M. vanbaalenii* (Gram-positive bacteria) is located on the negative axis alone, separated from the other two groups (*E. coli* and *A. baylyi*), which are Gram-negative.

10

256 The five artificial microbiota samples (M1 to M5) are located inbetween, and their

257 distances to the control groups are correlated with their community compositions.

258
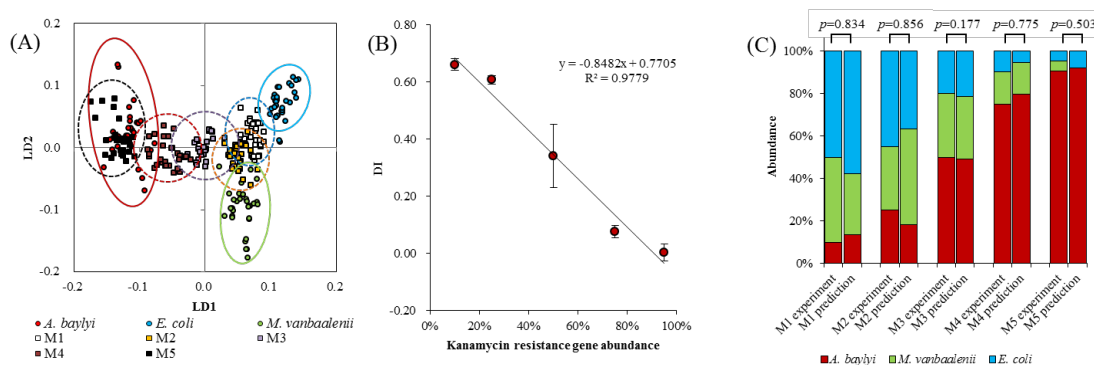
260 **Figure 3.** (A) Two-dimensional (LD1, LD2) scores plot after PCA-LDA of pure

261 microbial strains and artificial microbiotas with different composition. (B) Correlation

262 between kanamycin resistance gene abundance and group distance dispersion (D_I). (C)

263 Comparison of artificial microbiota composition between experimental data and

264 model prediction.

265

266 In order to predict the composition of artificial microbiota, the dispersion indicator

267 model[37] was carried out by transferring the dispersion analysis from the IR spectral

268 variables to the vectors (LD1 and LD2) and using D_I as the indicator, comparing to

269 the ARGs gene copy numbers quantified by qPCR as reference. This method used the

270 summarized spectral information from PCA-LDA which accounts for over 90% of

271 spectral variations in the present study, and was more conclusive than the limited

272 biomarkers from second order differentiation. Here, microbiotas with less abundance

273 of *A. baylyi* were further separated from the *A. baylyi* group, but closer to those of *E.*

274 *coli* and *M. vanbaalenii*, leading to an increasing D_I against the decreasing

275 kanamycin resistance (kanamycin resistance genes in *A. baylyi*). Figure 3B illustrates

276 the negative linear correlation between D_I and the abundance of *A. baylyi*

277 (kanamycin resistance gene abundance) within the artificial microbiotas ($D_I =$

278 $-0.8482 \times [kanamycin\ resistance\ gene] + 0.7705$). The high coefficient

279 (R^2=0.9779) suggests a good linear regression of D_I against kanamycin resistance.

280 The composition of each microbiota was, therefore, calculated from the D_I linear

11

281 regression based on PCA-LDA, as shown in Figure 3C. The results indicated that the

282 predicted microbial compositions had high similarity to their theoretical structure with

283 no significant differences found ($P > 0.05$). The standard deviation of microbiota M3

284 (middle point in Figure 3B) was greater than the others, possibly attributing to their

285 higher Shannon-Wiener index (1.02) than other microbiotas (0.35 to 0.94 for M1, M2,

286 M4 and M5). Shannon-Wiener index represents the diversity of microbial community,

287 and higher microbial diversity has been reported to increase complicated

288 intracommunity interaction[32]. It might cause huge variation of microbial chemical

289 composition, consequently leading to the difficulties in interrogating spectral

290 biomarkers and significant standard deviation in data prediction.

291 *Quantification of kanamycin resistance dynamics within microbiota*

292 Figure 4A illustrates the PCA-LDA scores plot of microbiotas post-exposure to

293 kanamycin, derived from the spectral dynamics of the artificial microbiotas (see ESI

294 Figure S3). All the interrogated communities exhibit a dramatic shift from the original

295 location as the exposure time increases. The *M. vanbaalenii* category moves towards a

296 different direction when compared to *A. baylyi* and *E. coli*, which might be attributed

297 to distinct cell structures between Gram-positive (*M. vanbaalenii*) and Gram-negative

298 bacteria (*A. baylyi* and *E. coli*). Specifically, there is only one lipid bilayer in the

299 membrane of Gram-positive bacteria, with a thick ring of peptidoglycan and teichoic

300 acid[38,39]. On the other hand, the cell membrane of Gram-negative bacteria contains

301 two lipid associated bilayers, which appear to increase the chance that the applied

302 treatments influence their structure[38,39]. The artificial microbiotas, AM1 and AM2,

303 follow similar trends as the *A. baylyi* and they come even closer to *A. baylyi* after

304 extended exposure to the kanamycin antibiotic. After PCA-LDA, the most

305 discriminating peaks were observed in Gram-negative bacteria and were attributed to

306 lipids (~ 1750 cm^{-1}), ν_{as}PO$_2^-$ (~ 1225 cm^{-1}) and ν_sPO$_2^-$ (~ 1080 cm^{-1}). Kanamycin's

307 antimicrobial mechanism is associated with aminoglycosides, interfering with

308 aminoacyl-tRNA recognition at the ribosomal A site and disrupting protein

309 expression[40]. Such a mechanism causes series of secondary effects, *e.g.*, membrane

310 damage. Our results are consistent with previous findings showing that the damage is

311 mainly linked to a broad range of alterations associated with the elements of

312 membranes, *e.g.*, proteins, supported by derived peaks the protein absorbance region

313 from 1500 to 1700 cm^{-1}, such as Amide II (~ 1517 cm^{-1}, ~ 1543 cm^{-1}) and Amide I

12

314 ($\sim$1650 cm^{-1}, $\sim$1680 cm^{-1})[16,21,41,42].

315   Applying the linear D_I regression model, we successfully predicted the dynamic

316 abundance of *A. baylyi* and kanamycin resistance within the microbiotas under

317 kanamycin antibiotic pressures. Both artificial microbiotas, AM1 (Figure 4B) and

318 AM2 (Figure 4C), had defined community composition at 0 h, with *A. baylyi*

319 (kanamycin resistance gene) accounting for 10% and 40% of the total population,

320 respectively. Post-exposure to kanamycin, the ARGs abundance from qPCR results

321 gradually increased to 85.0% in AM1 and 92.2% in AM2 after 12 h, which is

322 explained by the competitive advantages of bacteria with kanamycin resistance gene

323 in the community[43]. It therefore led to a faster growth of *A. baylyi* compared to other

324 strains and subsequent dominancy of *A. baylyi* within the microbiota. From the

325 dynamics of discriminant functions, the predicted ARGs abundance in both

326 microbiotas fitted efficiently with experimental data (Figure 4B and 4C). The linear

327 correlation at each time point did not show significant difference between predicted

328 and experimental ARGs abundance (Figure 4D), with a Pearson correlation

329 coefficient of 0.9487. The prediction *via* infrared spectroscopy coupled and

330 multivariate analysis fitted the experimental data better at higher ARGs abundance,

331 but was slightly lower than the qPCR results at low ARGs abundance, *e.g.*, 33% in

332 Figure 4D, which might underestimate the ARGs abundance to some extent. These

333 results not only prove that our model can be used for static community composition

334 and abundance/dynamics of kanamycin resistance gene, but they also evaluate the

335 impact of antibiotic pressure on kanamycin resistance gene transfer or dominancy.

336

337
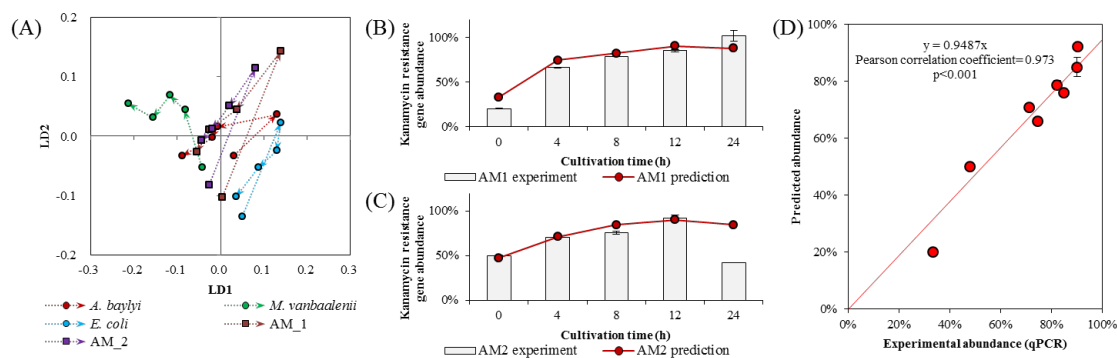
338 **Figure 4.** (A) Two-dimensional (LD1, LD2) scores plot after PCA-LDA of IR

339 dynamics of artificial microbiotas. Dots along with the arrow point in each colour

13

340 refer to the measurement at 0, 4, 8, 12 and 24 h, respectively. The prediction of

341 kanamycin resistance gene abundance is based on the dispersion among the

342 classification groups in PCA-LDA for artificial microbiotas AM1 (B) and AM2 (C).

343 (D) Regression correlation of kanamycin resistance gene abundance between

344 experimental data *via* qPCR and model prediction.

345

346 It is worth mentioning that less dispersion is observed for *A. baylyi* after exposure

347 because *A. baylyi* ADPWH_recA contains the *kan*[R] kanamycin resistance gene, which

348 is capable of tolerating kanamycin pressure. In the present study, the *kan*[R] kanamycin

349 resistance gene belongs to *npt* encoding neomycin phosphotransferase and shows high

350 similarity to *addA* encoding aminoglycoside phosphotransferase (aminoglycoside

351 kinase), which modifies the aminoglycosides by phosphoryl transfer, catalysing the

352 phosphate addition from ATP to 3'-hydroxyl group[40]. By expressing *kan*[R], *A. baylyi*

353 ADPWH_recA inactivates the interference of protein expression by kanamycin,

354 achieves fast recovery from suppression, and minimizes spectral alterations as

355 compared to others. It is confirmed by the presence of consistent shifts and

356 discriminating biomarkers in *A. baylyi* postexposure to kanamycin, including Amide I

357 ($\sim$1630 cm^{-1}, $\sim$1640 cm^{-1}) and $v_{as}PO_2^-$ ($\sim$1222 cm^{-1}, $\sim$1235 cm^{-1})[42].

358 An unexpected decline of kanamycin resistance gene was observed for AM2

359 artificial microbiota after 24 h exposure to kanamycin (42%, Figure 4C), but the

360 predicted kanamycin resistance by D_I regression model remained close to 100%. It

361 might be explained by the dramatically decreasing kanamycin concentration via the

362 metabolism of aminoglycoside modifying enzyme and the change in microbial

363 community structure. The functions of *kan*[R] encoding aminoglycoside kinase are

364 stabilizing a metaphosphate transition state and inactivating kanamycin[31], and the

365 spectral alterations represent the alignment disruption of ß-phosphate and γ-phosphate

366 by amide backbone. The declining kanamycin results in less inhibition on bacteria

367 without kanamycin resistance gene (*M. vanbaalenii* and *E. coli*), and their growth and

368 regeneration consequently reduce the abundance of A. *baylyi* and *kan*[R] gene.

369 Alternatively, the FTIR spectral alteration reflects such phenotypic changes of the

370 whole microbiota under the low kanamycin exposure, illustrating the fact that the

371 majority of microbial cells within the microbiota have the pseudo-resistance to

14

372   kanamycin. The spectrochemical interrogation therefore actually quantifies the

373   microbial phenotypic antibiotic resistance rather than the ARGs abundance only.

374   Infrared spectroscopy has demonstrated the ability to diagnose the phenotypic

375   alteration of the cellular components induced by kanamycin, hinting its potential

376   possibility for the application to other members of the aminoglycoside family. Our

377   findings indicate that this dispersion model coupled with PCA-LDA is a potential

378   approach for monitoring the population dynamics within a microbiota in real-time.

379   Additionally, the model applied in the present study summarizes the whole spectral

380   information derived from the multivariate analysis, rather than only several

381   biomarkers, showing its potential as a universal predicting tool for a broad spectrum

382   of antibiotics based on well-trained databases. Though only successfully applied in

383   the case of kanamycin through phosphotransferase resistance pathway, this technique

384   is also feasible for detecting *N*-acetyltransferases and *O*-nucleotidyltransferases,

385   which also belong to aminoglycoside-modifying enzymes assisted by

386   acetyl-coenzyme A and ATP respectively[22], attributing to their similar anti-kanamycin

387   mechanisms as *kan^R* encoding neomycin phosphotransferase. Future work should

388   refer to more comprehensive range of antibiotics and their mechanisms including

389   penicillin-class (*e.g.*, ampicillin and amoxicillin), which disrupts the synthesis of

390   peptidoglycan layer and inhibits bacterial cell wall synthesis [44], and tetracycline,

391   which inhibits the binding of aminoacyl-tRNA and suppresses protein expression[45].

392   For the urgent need to characterize antibiotic resistance in complex environmental

393   microbiota with spectroscopy, the primary challenges are raised as the lack of routine

394   protocols, reproducible computational analysis, and reliable database[10]. Validated in

395   the artificial microbiota, our work provides the solutions for the first two barriers by

396   distinguishing biomarkers representing antibiotic resistance from the numerous

397   biological fingerprints. A well-built dataset along with robust analytical models

398   coupled with spectroscopic methods are suggested to address the antibiotic resistance

399   dynamics in real environmental samples.

400   The present study indicates that infrared spectroscopy, in conjunction with

401   multivariate analysis, is a potential tool for diagnosing the phenotypic existence and

402   dynamics of ARGs within microbial communities. Our work employed ATR-FTIR

403   spectroscopy coupled with a dispersion model to quantify microbial kanamycin

404   resistance, based on secondary derivative and PCA-LDA. This method not only

15

405   quantified the static community composition of the artificial microbiotas but also

406   successfully predicted the population dynamics of microbial communities and

407   kanamycin resistance under antibiotic pressure. We also suggest that spectroscopic

408   techniques have great potential in real-time monitoring of microbiota of interest in

409   medical or environmental fields; this would provide an excellent opportunity to

410   visualize the vivid phenotypic transformation during a biological and biochemical

411   process rather than only intermittent snap-shots.

412

418

419

16

References
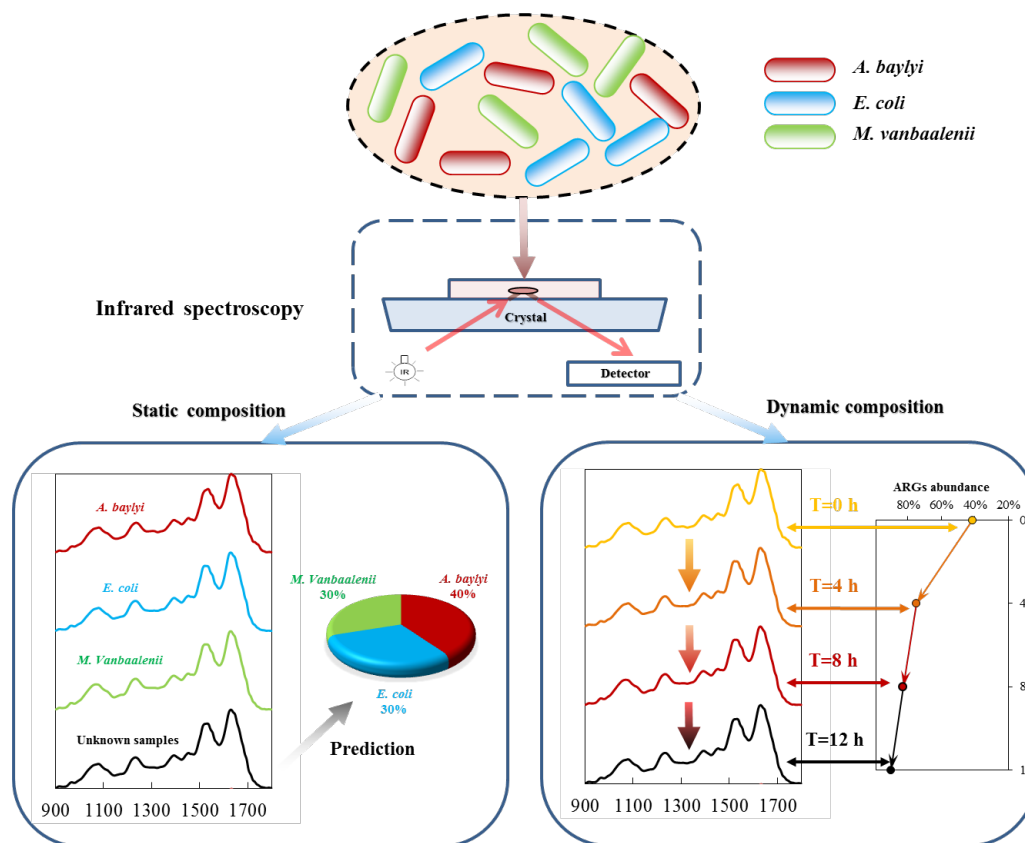
(1) Blair, J. M.; Webber, M. A.; Baylay, A. J.; Ogbolu, D. O.; Piddock, L. J. *Nat. Rev. Microbiol.* **2015**, *13*, 42-51.

(2) Chee-Sanford, J. C.; Aminov, R. I.; Krapac, I. J.; Garrigues-Jeanjean, N.; Mackie, R. I. *Appl. Environ. Microbiol.* **2001**, *67*, 1494-1502.

(3) Cantas, L.; Shah, S. Q. A.; Cavaco, L. M.; Manaia, C. M.; Walsh, F.; Popowska, M.; Garelick, H.; Burgmann, H.; Sorum, H. *Front. Microbiol.* **2013**, *4*.

(4) Potera, C. *Environ. Health Perspect.* **2013**, *121*, A255-A255.

(5) Smillie, C. S.; Smith, M. B.; Friedman, J.; Cordero, O. X.; David, L. A.; Alm, E. J. *Nature* **2011**, *480*, 241-244.

(6) Colomer-Lluch, M.; Imamovic, L.; Jofre, J.; Muniesa, M. *Antimicrob. Agents Chemother.* **2011**, *55*, 4908-4911.

(7) Riesenfeld, C. S.; Goodman, R. M.; Handelsman, J. *Environ. Microbiol.* **2004**, *6*, 981-989.

(8) Paraskevaidi, M.; Martin-Hirsch, P. L.; Kyrgiou, M.; Martin, F. L. *Mutagenesis* **2017**, *32*, 335-342.

(9) Kaern, M.; Elston, T. C.; Blake, W. J.; Collins, J. J. *Nat. Rev. Genet.* **2005**, *6*, 451-464.

(10) Jin, N. F.; Zhang, D. Y.; Martin, F. L. *Integr. Biol.* **2017**, *9*, 406-417.

(11) Naumann, D.; Helm, D.; Labischinski, H. *Nature* **1991**, *351*, 81-82.

(12) Dunn, W. B.; Ellis, D. I. *Trends Anal. Chem.* **2005**, *24*, 285-294.

(13) Freedman, B. G.; Zu, T. N. K.; Wallace, R. S.; Senger, R. S. *Biotechnol. J.* **2016**, *11*, 877-889.

(14) Sun, S. W.; Wang, X. T.; Gao, X.; Ren, L. H.; Su, X. Q.; Bu, D. B.; Ning, K. *BMC Bioinformatics* **2015**, *16*.

(15) Winnard, P. T.; Zhang, C.; Vesuna, F.; Kang, J. W.; Garry, J.; Dasari, R. R.; Barman, I.; Raman, V. *Oncotarget* **2017**, *8*, 20266-20287.

(16) Li, J. Y.; Strong, R.; Trevisan, J.; Fogarty, S. W.; Fullwood, N. J.; Jones, K. C.; Martin, F. L. *Environ. Sci. Technol.* **2013**, *47*, 10005-10011.

17

449 (17) Bankapur, A.; Krishnamurthy, R. S.; Zachariah, E.; Santhosh, C.; Chougule, B.;

450 Praveen, B.; Valiathan, M.; Mathur, D. *PLoS One* **2012**, *7*.

451 (18) Tao, Y. F.; Wang, Y.; Huang, S.; Zhu, P. F.; Huang, W. E.; Ling, J. Q.; Xu, J.

452 *Anal. Chem.* **2017**, *89*, 4108-4115.

453 (19) Siddhanta, S.; Paidi, S. K.; Bushley, K.; Prasad, R.; Barman, I. *Chemphyschem*

454 **2017**, *18*, 72-78.

455 (20) Baker, M. J.; Trevisan, J.; Bassan, P.; Bhargava, R.; Butler, H. J.; Dorling, K. M.;

456 Fielden, P. R.; Fogarty, S. W.; Fullwood, N. J.; Heys, K. A.; Hughes, C.; Lasch, P.;

457 Martin-Hirsch, P. L.; Obinaju, B.; Sockalingum, G. D.; Sule-Suso, J.; Strong, R. J.;

458 Walsh, M. J.; Wood, B. R.; Gardner, P.; Martin, F. L. *Nat. Protoc.* **2014**, *9*,

459 1771-1791.

460 (21) Martin, F. L.; Kelly, J. G.; Llabjani, V.; Martin-Hirsch, P. L.; Patel, II; Trevisan,

461 J.; Fullwood, N. J.; Walsh, M. J. *Nat. Protoc.* **2010**, *5*, 1748-1760.

462 (22) Mingeot-Leclercq, M. P.; Glupczynski, Y.; Tulkens, P. M. *Antimicrob. Agents*

463 *Chemother.* **1999**, *43*, 727-737.

464 (23) Sadovskaya, I.; Vinogradov, E.; Li, J. J.; Hachani, A.; Kowalska, K.; Filloux, A.

465 *Glycobiology* **2010**, *20*, 895-904.

466 (24) Delorenzo, V.; Herrero, M.; Jakubzik, U.; Timmis, K. N. *J. Bacteriol.* **1990**, *172*,

467 6568-6572.

468 (25) Song, Y.; Li, G.; Thornton, S. F.; Thompson, I. P.; Banwart, S. A.; Lerner, D. N.;

469 Huang, W. E. *Environ. Sci. Technol.* **2009**, *43*, 7931-7938.

470 (26) Trevisan, J.; Angelov, P. P.; Scott, A. D.; Carmichael, P. L.; Martin, F. L.

471 *Bioinformatics* **2013**, *29*, 1095-1097.

472 (27) Butler, H. J.; McAinsh, M. R.; Adams, S.; Martin, F. L. *Anal. Methods* **2015**, *7*,

473 4059-4070.

474 (28) Yenofsky, R. L.; Fine, M.; Pellow, J. W. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*,

475 3435-3439.

476 (29) Bosch, A.; Serra, D.; Prieto, C.; Schmitt, J.; Naumann, D.; Yantorno, O. *Appl.*

477 *Microbiol. Biotechnol.* **2006**, *71*, 736-747.

18

478   (30) Mariey, L.; Signolle, J. P.; Amiel, C.; Travert, J. *Vib. Spectrosc.* **2001**, *26*,

479   151-159.

480   (31) Wright, G. D. *Curr. Opin. Microbiol.* **1999**, *2*, 499-503.

481   (32) Flemming, H. C.; Wingender, J. *Nat. Rev. Microbiol.* **2010**, *8*, 623-633.

482   (33) Holman, H. Y. N.; Miles, R.; Hao, Z.; Wozei, E.; Anderson, L. M.; Yang, H.

483   *Anal. Chem.* **2009**, *81*, 8564-8570.

484   (34) Stewart, P. S. *Int. J. Med. Microbiol.* **2002**, *292*, 107-113.

485   (35) Hoiby, N.; Bjarnsholt, T.; Givskov, M.; Molin, S.; Ciofu, O. *Int. J. Antimicrob.*

486   *Agents* **2010**, *35*, 322-332.

487   (36) Karunakaran, E.; Mukherjee, J.; Ramalingam, B.; Biggs, C. A. *Appl. Microbiol.*

488   *Biotechnol.* **2011**, *90*, 1869-1881.

489   (37) Li, H. B.; Martin, F. L.; Zhang, D. Y. *Anal. Chem.* **2017**, *89*, 3909-3918.

490   (38) Ede, S. M.; Hafner, L. M.; Fredericks, P. M. *Appl. Spectrosc.* **2004**, *58*, 317-322.

491   (39) Morones, J. R.; Elechiguerra, J. L.; Camacho, A.; Holt, K.; Kouri, J. B.; Ramirez,

492   J. T.; Yacaman, M. J. *Nanotechnology* **2005**, *16*, 2346-2353.

493   (40) Boehr, D. D.; Thompson, P. R.; Wright, G. D. *J. Biol. Chem.* **2001**, *276*,

494   23929-23936.

495   (41) Heys, K. A.; Riding, M. J.; Strong, R. J.; Shore, R. F.; Pereira, M. G.; Jones, K.

496   C.; Semple, K. T.; Martin, F. L. *Analyst* **2014**, *139*, 896-905.

497   (42) Movasaghi, Z.; Rehman, S.; Rehman, I. U. *Appl. Spectrosc. Rev.* **2008**, *43*,

498   134-179.

499   (43) Hibbing, M. E.; Fuqua, C.; Parsek, M. R.; Peterson, S. B. *Nat. Rev. Microbiol.*

500   **2010**, *8*, 15-25.

501   (44) Strominger, J. L.; Park, J. T.; Thompson, R. E. *J. Biol. Chem.* **1959**, *234*,

502   3263-3268.

503   (45) Connell, S. R.; Trieber, C. A.; Dinos, G. P.; Einfeldt, E.; Taylor, D. E.; Nierhaus,

504   K. H. *EMBO J.* **2003**, *22*, 945-953.

505

506

19

507

Table of Content Graphic

Infrared spectroscopy

Static composition

Dynamic composition

508

509

510

For TOC only

511

512