1 # Evaluation of Mobile Phase Characteristics on Three Zwitterionic
2 Columns in HILIC mode for Liquid Chromatography-High Resolution
3 Mass Spectrometry based Untargeted Metabolite profiling of
4 *Leishmania* Parasites

5 Rong Zhang[1,2], David G. Watson[♯1], Lijie Wang[1], Gareth D. Westrop[1], Graham H. Coombs[1], Tong
6 Zhang[♯*1]

7 1. Strathclyde Institute of Pharmacy and Biomedical Sciences, 161, Cathedral Street, Glasgow,
8 Scotland, UK, G4 0RE.
9 2. Institute of Clinical Pharmacology, Guangzhou University of Chinese Medicine, No. 12
10 Jichang Road, Guangzhou, China, 510405

11 ♯ joint senior authors

12

13 E-mails: zhangrong@gzucm.edu.cn; d.g.watson@strath.ac.uk; lijie.wang@strath.ac.uk;
14 gareth.westrop@strath.ac.uk; graham.coombs@strath.ac.uk; tong.zhang.101@strath.ac.uk;

15

16

17 *Corresponding author:

18 Tong Zhang,

19 Strathclyde Institute of Pharmacy and Biomedical Sciences,

20 161 Cathedral Street,

21 Glasgow G4 0RE

22 Telephone : +441415482651

23 Fax: +441415526443

24 E mail: tong.zhang.101@strath.ac.uk

Abstract

It has been reported that HILIC column chemistry has a great effect on the number of detected metabolites in LC-HRMS-based untargeted metabolite profiling studies. However, no systematic investigation has been carried out with regard to the optimisation of mobile phase characteristics. In this study using 223 metabolite standards, we explored the retention mechanisms on three zwitterionic columns with varied mobile phase composition, demonstrated the interference from poor chromatographic peak shapes on the output of data extraction, and assessed the quality of chromatographic signals and the separation of isomers under each LC condition. As expected, on the ZIC-cHILIC column the acidic metabolites showed improved chromatographic performance at low pH which can be attributed to the opposite arrangement of the permanently charged groups on this column in comparison with the ZIC-HILIC column. Using extracts from the protozoan parasite *Leishmania*, we compared the numbers of repeatedly detected LC-HRMS features under different LC conditions with putative identification of metabolites not amongst the standards being based on accurate mass (±3 ppm). Besides column chemistry, the pH of the mobile phase plays a key role in not only determining the retention mechanisms of solutes but also the output of the LC-HRMS data processing. Fast evaporation of ammonium carbonate produced less ion suppression in ESI source and consequently improved the detectability of the metabolites in low abundance in comparison with other ammonium salts. Our results show that the combination of a ZIC-pHILIC column with an ammonium carbonate mobile phase, pH=9.2, at 20 mM in the aqueous phase or 10 mM in both aqueous and organic mobile phase components, provided the most suitable LC conditions for LC-HRMS-based untargeted metabolite profiling of *Leishmania* parasite extracts. The signal reliability of the mass spectrometer used in this study (Exactive Orbitrap) was also investigated.

Keywords: HILIC-HRMS (Hydrophilic Interaction Liquid Chromatography-High Resolution Mass Spectrometry); zwitterionic stationary phase; Metebolomics; untargeted metabolite profiling; Leishmania.

1. Introduction

Untargeted metabolite profiling is a major challenge in metabolomics because of the wide variation in the physicochemical properties and abundance of the metabolites in a biological sample. Due to the rapid development of instrumentation in High Resolution Mass Spectrometry (HRMS) during the last two decades, HRMS-based analytical platforms are now able to achieve high sensitivity and mass resolution and this has greatly extended the coverage of metabolites in untargeted metabolite profiling [1, 2]. HRMS is usually interfaced with a separation technique providing retention/migration times as additional identity information for metabolites and reducing ion suppression during the ionisation process [3-6]. More importantly, the separation step offers the possibility of distinguishing isomeric metabolites; this is often unachievable with HRMS alone. In comparison with Gas Chromatography (GC) and Capillary Electrophoresis (CE), Liquid Chromatography (LC) demonstrates superiority in the requirements for sample preparation, metabolite coverage range and reproducibility [4, 5, 7]. Therefore it is the most commonly used separation technique in HRMS-based metabolomics studies.

Reversed Phase (RP) LC was the first LC separation technique hyphenated with HRMS and has been widely employed in metabolomics studies. However, its limitation is in the retention and separation of polar metabolites and these are major components in all biological samples. Hydrophilic Interaction Liquid Chromatography (HILIC) has been increasingly applied in recent HRMS-based metabolomics studies [8-11]. It not only enables superior retention of hydrophilic compounds but also improves their ionisation efficiency with ESI. The general retention mechanism in HILIC is believed to be the partitioning of the analytes between an organic-rich mobile phase and the aqueous layer formed at the surface of the polar stationary phase. A great number of HILIC columns are now commercially available and they show very different chromatographic performance [12, 13]. Therefore many column comparison/characterisation studies have been performed to investigate the retention mechanisms involved in HILIC [14-20]. They have often provided a theoretical understanding of chromatographic interactions in HILIC, but they are not a particularly helpful guide for column selection in metabolomics studies. A few column evaluation studies addressing the practical concerns in LC-HRMS based metabolomics studies have been reported and have provided information on the number of repeatedly detected LC-HRMS features, the

85 separation of isomers and the quality of the chromatographic peak shapes [21-23]. In these
86 studies, however, the effects of mobile phase characteristics were not sufficiently
87 investigated.

88 Current HILIC columns can be categorised into three groups based on the stationary phase
89 chemistry: bare silica/neutral, charged and zwitterionic phases [11, 12]. Compared with the
90 other two types of HILIC columns, the zwitterionic stationary phase is believed to have more
91 potential for untargeted metabolite profiling studies. It has been demonstrated by many
92 column characterisation studies that, depending on the chemical nature of the analyte and
93 mobile phase composition, electrostatic interactions with permanently charged moieties and
94 hydrogen bonding adsorption on the silica gel substrate could occur on the zwitterionic
95 stationary phase [15, 18, 20, 24]. As there are also hydrophilic interactions, the solutes can
96 elute with a mixed retention/separation mechanism in HILIC mode providing a unique
97 selectivity in the separation of polar/ionisable metabolites. The first zwitterionic stationary
98 phase in HILIC was fabricated by bonding sulfobetaine moieties on a silica-based substrate.
99 The sulfonic group at the distal end of the zwitterionic moiety generates a negative surface
100 charge which will produce an electrostatic attraction of positively charged metabolites at low
101 pH values. Later the same moiety was bonded onto a polymer-based substrate to improve the
102 pH tolerance and this polymer-based stationary phase has demonstrated improved
103 chromatographic performance for negatively charged metabolites at high pH values [22, 25].
104 Both of the sulfobetaine zwitterionic stationary phases have been widely applied under HILIC
105 conditions in untargeted metabolite profiling studies for various biological samples [26-29].
106 Recently a new zwitterionic stationary phase bonded with phosphorylcholine moieties has
107 been introduced onto the market [30]. According to the manufacturer its opposite
108 arrangement of the permanently charged groups to sulfobetaine will provide a different
109 selectivity for charged solutes in HILIC mode. To date, there have been no reported
110 applications of this new type of zwitterionic stationary phase in metabolomics, its application
111 has only been reported in proteomic studies [31].

112 *Leishmania* is a group of protozoan parasites that can cause various disease manifestations
113 depending on the infecting species. Our previous studies have demonstrated the potential of
114 metabolomics for elucidating differences between *Leishmania* species with regard to their
115 metabolomes [28, 32, 33]. In these studies, sample preparation for LC-HRMS was optimised

116 for quenching of metabolism, cell washing and disruption, and metabolite extraction but
117 there remained scope for further development of LC methods regarding stationary and
118 mobile phases [34]. In the present study we have systematically evaluated the three
119 zwitterionic stationary phases in HILIC mode under different mobile phase conditions with
120 223 metabolite standards. Cell extracts of *Leishmania major* promastigotes and extracellular
121 medium after growth were used for assessment of each chromatographic condition with
122 respect to the practical concerns in data processing for a metabolomics study. In addition the
123 complementarity of RP LC to HILIC in the coverage of metabolites was investigated as was the
124 reliability of the LC-HRMS signals for comparative analyses using a series of diluted biological
125 samples.

126 **2. Experimental**

127 **2.1 Chemicals and solvents**

128 HPLC grade acetonitrile (ACN), chloroform and methanol were purchased from Fisher
129 Scientific, UK. HPLC grade water was produced by a Direct-Q 3 Ultrapure Water System from
130 Millipore, UK. AnalaR grade formic acid (98%) was obtained from BDH-Merck, UK. Ammonium
131 carbonate, ammonium formate, ammonium acetate, ammonium hydroxide solution (30-33%)
132 and all standard compounds were purchased from Sigma-Aldrich, UK.

133 **2.2 Preparation of metabolite standard solutions**

134 Each metabolite standard was prepared at 1 mg/ml with HPLC grade methanol and water (1:1,
135 v/v) as the stock solution and stored at -20°C. 100 µl was taken from each stock solution,
136 about 56 metabolites were mixed and then the solution was made up to 10 ml with
137 acetonitrile. Consequently, the final concentration for each metabolite standard was 10
138 µg/ml and 223 metabolite standards were distributed into four mixed metabolite standard
139 solutions (detailed in supplementary material 1 (SM1)). In order to avoid identity confusion,
140 isomers were distributed into different standard solutions and in-source fragments were also
141 carefully verified.

142 **2.3 Sample preparation of *L. major* cell extracts and spent media**

143 *Leishmania major* promastigote were inoculated at 2.5 x 10^5 cells/ml in replicate 10 ml
144 cultures at 26°C. Promastigotes were harvested after 6 days for metabolite extraction. The
145 cell metabolite extraction was performed as described in our previous study [34] and the
146 spent media metabolite extraction was performed as following: to 75 µl of spent medium
147 was added 300 µl of cold chloroform/methanol (20/60, v/v) followed by incubation for 1 h in
148 a Thermomixer (14,000 rpm, 4ºC). After centrifugation at 12,000 g for 10 min at 4ºC, the
149 supernatant was recovered and stored at -70ºC until analysed. 200 µl of cell extract in
150 chloroform/methanol/water (20/60/20, v/v/v) was obtained from each biological replicate. In
151 order to ensure the equality of the samples tested under each LC condition, the obtained cell
152 extracts from all biological replicates were thoroughly mixed together, redistributed into 25
153 aliquots and stored at -80°C until analysis. The same sample reconstruction was also carried
154 out to the spent media extracts.

155 **2.4 LC-HRMS conditions**

156 Three zwitterionic HILIC columns (ZIC-HILIC and ZIC-pHILIC, both 150×4.6 mm, 5 µm; ZIC-
157 cHILIC, 150×4.6 mm, 3 µm) and a reversed phase column (ACE C18-AR, 150×4.6 mm, 5 µm)
158 were obtained from HiChrom (Reading UK). A Dionex Ultimate 3000 HPLC system (Camberley,
159 UK) combined with an Exactive (Orbitrap) mass spectrometer from Thermo Fisher Scientific
160 (Bremen, Germany) was employed as the LC-HRMS platform in this study. The MS parameter
161 settings were described in our previous study [27] and the details of the mobile phase
162 compositions and the gradient programs are described in Table 1. The flow rate of 0.3 ml/min
163 was used under all LC conditions. In order to equilibrate the column for each LC condition,
164 two solvent blanks (pure ACN) were run in advance and followed by three consecutive
165 injections of the cell extract sample and then the spent medium sample and the four mixed
166 metabolite standard solutions were run subsequently. The samples were kept at 4 ˚C during
167 the experiment and a new set were used for each LC condition.

168 **2.5 Data processing**

169 The LC-HRMS raw data of the mixed metabolite standard solutions were processed using
170 ToxID 2.1 (Thermo Fisher Scientific Inc., Hemel Hempstead, UK) with ±3 ppm mass accuracy
171 tolerance with both ESI positive and negative modes. The generated extracted ion
172 chromatograms of metabolite standards under different LC conditions were visually evaluated

173 with respect to peak shapes and are summarised with their retention times in SM 1. MZMine
174 2.10 was used for processing the LC-HRMS raw data of the parasite samples. The procedure
175 and the settings were the same as described in our previous study [27] and the putative
176 identification was also based on our in-house database used before. All representative
177 chromatograms were produced by using Xcalibur 2.2 (Thermo Fisher Scientific Inc., Hemel
178 Hempstead, UK).

179 **2.6 Statistical analysis of LC conditions**

180 The retention times of the metabolite standards generated by ToxID 2.1 were visually
181 evaluated and corrected and then imported to SIMCA-P 13 (Umetrics, Sweden) for Principal
182 Component Analysis (PCA) of 20 LC conditions. Prior to PCA, the data were mean-centered
183 and unit variance (UV) scaled. Other statistical analysis were performed using the statistical
184 functions in Excel (Microsoft Office 2010) associated with a few macros coded by Visual Basic
185 for Application (VBA, Microsoft Office 2010).

186 **2.7 Preparation of diluted parasite samples**

187 To prepare a 3 times diluted sample, 100 µl of the original extract solution was diluted with
188 200 µl chloroform/methanol/water (20/60/20, v/v/v). A similar process was used to achieve 9,
189 27 and 81 times dilutions.

190 **3. Results and discussion**

191 **3.1 Evaluation of chromatographic factors influential to retention mechanisms**

192 We evaluated five chromatographic factors in our study: stationary phase chemistry; pH of
193 mobile phase; nature of mobile phase additives (counter ions present or not); mobile phase
194 additives added only in the aqueous mobile phase component or in both organic and aqueous
195 components; and mobile phase buffer concentration. The ratio of aqueous and organic
196 solvents in the mobile phase of the eluting program was not evaluated because this is the
197 most influential chromatographic factor in HILIC mode and its effect is widely known. As
198 described in Table 1, a gradient program of 80% ACN linearly decreasing to 20% in 30 min was
199 applied for all the tested HILIC conditions. In order to evaluate the retention behaviour of 232
200 metabolite standards under different LC conditions, Principal Component Analysis (PCA) was

201 carried out based on their retention times. In the initial score plot, C18-FA was assigned far
202 away from others as an outlier (data not shown), as expected because of the very different
203 reversed phase retention mechanism. After excluding this set of conditions, the score plot
204 generated by all HILIC conditions is shown in Figure 1. It is clear that the stationary phase
205 chemistry and the pH of the mobile phase dominate the pattern. These two factors were
206 represented by the colours and the shapes and almost reflected by the first and the second
207 components, respectively, in Figure 1. The next most influential factor was the mobile phase
208 additive used. At low pH (squares) the conditions with addition of formic acid (FA) in the
209 mobile phases were separated from the ones with ammonium formate (AF) on both ZIC-HILIC
210 and ZIC-cHILIC columns. The same situation could be observed with the HILIC conditions at
211 high pH (stars). Regarding the question of whether to add buffer into only the aqueous phase
212 or into both solvents, it seemed that the differential effect was greater at medium and high
213 pH (a conclusion based on the increased distance between each pair of two relevant
214 conditions).

215 Zwitterionic metabolites like amino acids and their derivatives were almost insensitive in
216 terms of specific retention behaviour to these HILIC factors. For example, as shown in Figure
217 2, the retention times and the peak shapes of L-glutamine remained fairly constant under
218 varying conditions except for a small decrease in retention time at high pH on the ZIC-pHILIC
219 column. In contrast, however, variation in column and mobile phase conditions caused readily
220 observable effects with basic and acidic metabolites (tertiary/quaternary amines and organic
221 acids) such as choline and maleic acid (Figure 2). Under the same FA condition (0.1% formic
222 acid in both aqueous and organic solvents: low pH without counter ions) the retention time of
223 choline (a completely positively charged metabolite) was significantly less on the ZIC-cHILIC
224 column compared to the ZIC-HILIC column, whereas that for maleic acid (a partially
225 negatively charged metabolite) was significantly increased. This could be due to the
226 electrostatic attraction and repulsion between the charged solutes and the charged surface of
227 the zwitterionic stationary phases. It is believed that on the ZIC-cHILIC column the positively
228 charged surface (phosphorylcholine) should attract negatively charged solutes but repulse
229 positively charged solutes, with the opposite occurring on the ZIC-HILIC column. It should be
230 also noticed that on the ZIC-HILIC column the peak shape of maleic acid was broad and split
231 which might produce wrong/confusing outcome in the peak extraction algorithms used in

232 processing metabolomics data. This is most likely due to repulsion by the negatively charged
233 groups on the surface of the stationary phase causing loss of peak integrity. On the ZIC-pHILIC
234 column, which has a sulphonic acid group towards the surface of the phase, at high pH, with
235 ammonium carbonate modifier, all three analytes have good peak shapes and the better peak
236 shapes being at higher ionic strength with 10 mM buffer in reservoirs A and B. The retention
237 time of choline decreases at higher ionic strength due to increased competition from the
238 ammonium ion. In contrast, the retention time of the maleic acid increases at higher ionic
239 strength due to less electrostatic repulsion due to masking of the negative charge by
240 increased levels of the ammonium ion. Often chromatographers only incorporate modifier
241 into the aqueous phase reservoir (A), partly because the modifiers used in LC-MS are not that
242 soluble in 100% acetonitrile. The outcome of this approach was studied for the ZIC-cHILIC
243 column and it was found that incorporation of ammonium acetate into A alone led to poorer
244 peak shapes due to lower ionic strength overall whereas having the modifier in both
245 reservoirs resulted in better peak shapes due to the same factors as discussed for the ZIC-
246 pHILIC column. The optimal peak shapes on the ZIC-cHILIC column were with ammonium
247 formate in reservoirs A and B. This might be due to a higher level of ionisation of the two
248 buffer ions at high levels of acetonitrile since formic acid will be less ion suppressed in an
249 environment with low dielectric constant than acetic acid as it has a lower pKa value and thus
250 be better able to donate a proton to ammonia and keep in its protonated form. The ZIC-
251 pHILIC column, in which the stationary phase is based on an organic polymer rather than silica
252 gel, was the only one of the three zwitterionic HILIC columns suitable for testing at high pH
253 (9.2). It was found that, consistent with the data of Figure 1, the retention behaviour of the
254 metabolite standards were quite distinct on this column under these conditions even for the
255 zwitterionic metabolites such as L-glutamine (Figure 2). However, the unknown surface
256 chemistry of the polymer substrate obstructed the understanding of this stationary phase to
257 some extent. Apart from hydrophilic partitioning, ion exchange interaction plays an important
258 role. For instance, the decrease in the retention time of choline with increasing the buffer
259 concentration from 5 mM to 10 mM is typical of ion exchange interaction. Compared with
260 partially ionised status at low and medium pH, the two carboxylic groups in maleic acid are
261 completely deprotonated at pH 9.2 which should improve the electrostatic attraction to the
262 positively charged betaine group thus causing it to elute later at high pH in comparison to low
263 and medium pH. Moreover, ammonium ions in the mobile phase will associate with the

264 negatively charged sulfonate groups facilitating the interaction of negatively charged solutes
265 with the betaine group which is closer to the surface of the stationary phase support. This
266 hypothesis is supported by the increase of the retention time of maleic acid with increasing
267 the buffer concentration because the maleate ions are able to stay longer in the stationary
268 phase if the charge repulsion by sulfonate groups is reduced by their association with
269 ammonium ions. The peak shape of maleic acid was also improved when the buffer
270 concentration was increased, which could be attributed to reduced charge repulsion. This
271 phenomenon at high pH on the ZIC-pHILIC column could be observed with all completely
272 negatively charged metabolites (almost all acidic metabolites) including diphosphate or
273 triphosphate nucleotides (see SM 1).

274 **3.2 Examples of poor chromatographic peaks and their interference to data processing**

275 Elucidation of the retention mechanisms is not the aim of LC optimisation in LC-MS-based
276 metabolomics studies. In untargeted metabolite profiling, optimisation of the number of
277 metabolites detected is the highest priority in selecting a LC-MS method. However, this
278 number can be potentially determined by the data processing. In our experience, the quality
279 of the chromatographic peaks generated has a great effect on the output of the data
280 processing. Poor peak shapes are more likely to interfere with peak extraction (peak picking)
281 and alignment in the data processing and consequently confuse the identification and
282 produce wrong comparative information - both potentially resulting in misleading biological
283 interpretation. This issue might be overlooked by biologists because of a lack of
284 chromatographic knowledge and understanding of the algorithms used in data processing.
285 Only few specific demonstrations of this issue can be readily found in the literature [22, 23].
286 Therefore we addressed the issue in this study and we present some examples and explaining
287 their potential for interference with data processing. By visually checking the
288 chromatographic signals generated by 223 metabolite standards under different LC
289 conditions, it was possible to divide the poor chromatographic peaks into six categories:
290 absence of a distinct peak; multiple peaks; double peaks; split peaks; broad peaks; and tailing
291 peaks. Some chromatograms of metabolites are shown in Figure 3 to demonstrate typical
292 poor peak shapes in each of these categories and we also show a comparative chromatogram
293 of the same metabolite under different LC conditions where a good peak shape was
294 produced. The worst case is generation of what amounts to baseline noise, which can be

295 described as a bunch of zigzag/splitting/irregular peaks covering a wide time range. The
296 example shown in Figure 3 is putrescine under pHILIC-AC-10 mM-both conditions. These
297 irregular patterns are usually non-repeatable even between two consecutive injections.
298 During the data processing of this type of peak many features with the same m/z will be
299 generated by the peak extraction algorithm; some of them would be removed by noise
300 and/or peak shape filter algorithms, depending on the software used, which would itself
301 introduce an artefact. After alignment, the remaining multiple features would confuse the
302 subsequent identification because they are generated by a single metabolite; moreover,
303 wrong comparative outcomes would be produced because of their irregular and fluctuating
304 profile. The same issue would emerge if the chromatographic signal looked like the peak for
305 isocitrate under cHILIC-AF conditions in Figure 3. In this case, however, several peaks could be
306 observed repeatedly at certain retention times across samples which might produce correct
307 peak extraction and alignment but still confuse the identification. If the chromatographic
308 signal splits or forms two closely eluting features as in the case of maltose, features which are
309 due to the equilibrium between and the α- and β-anomers of this reducing sugar under the
310 pHILIC-AC-5mM conditions used, there would be an uncertain output of peak extraction. It
311 may treat as two or one feature depending on the settings of the algorithm; it certainly would
312 be confused with another situation in which two isomers elute very closely. Theoretically it is
313 expected that in the former situation the chromatographic signal should be treated as one
314 feature, but two features in the latter situation. Practically, however, there is no such ideal
315 setting in any peak extraction algorithm to avoid this confusion in LC-MS data processing for
316 untargeted analysis. Some metabolites can generate two peaks apart from each other such as
317 nicotinate under pHILIC-AC-10mM conditions as also shown in Figure 3. In this case the
318 chromatographic signal would be certainly treated as two features which would confuse the
319 identification. Sometimes the shape and size of the peaks can vary from sample to sample
320 depending on the composition of the sample solvent. Compared to the poor peak shapes
321 described above, broad and tailing peaks might be acceptable in some cases, especially if their
322 peak shapes are repeatable, because a single feature could be generated with proper settings
323 (e.g. maximum peak span) in the peak extraction algorithm. However, if the peak is
324 extensively broad or tailing with zigzag profiling across the top, like the pyridoxamine under
325 pHILIC-AC-10mM conditions example in Figure 3, there would be high possibility of excluding
326 the feature or splitting the peak into multiple features in the peak extraction algorithm.

327 Sometimes a combination of several poor peak shapes can be observed in one chromatogram
328 and a brief comment on the chromatographic signal of each metabolite standard under
329 different LC conditions is described in SM 1. The ideal peak shape for data processing is a
330 single and sharp peak, like the ones in the right column in Figure 3. A little tailing or broadness
331 could be acceptable although it might introduce inaccuracy in peak area integration.

332 **3.3 Evaluation of LC conditions using metabolite standards**

333 The number of metabolite standards detected under each LC condition, regardless of peak
334 shape, is presented as a yellow strip in Figure 4. The main reason for a metabolite standard to
335 be undetected is because it is unable to ionise under the conditions applied. This is primarily
336 determined by the pH value of the mobile phase and the pKa value(s) of the metabolite and
337 not the other LC parameters. The other two possible reasons are either that the metabolite
338 was not able to elute within the given run time or that it was not stable in the testing solution
339 and so degraded. For example, oxaloacetate will rapidly degrade to pyruvate in the
340 water/methanol solution. As can be seen in Figure 4, all LC conditions showed quite similar
341 numbers for the detected metabolite standards. As demonstrated and explained above,
342 however, they would not be treated correctly by the data processing algorithms if their peak
343 shapes are poor. Therefore it is more meaningful to compare between the different LC
344 conditions the number of metabolites showing good peak shapes rather than the total
345 number of detected metabolites. The number of metabolites exhibiting good peaks shapes is
346 given as a purple strip for each LC condition in Figure 4. It is clear that the numbers of
347 compounds detected dropped down dramatically under some LC conditions but not much
348 under others and the fall in detectable metabolites seemed to be pH related (Figure 4). In
349 order to elucidate the effect the LC conditions are highlighted with stars for high pH; triangles
350 for medium pH and squares for low pH mobile phases in Figure 4 and also further classified
351 the metabolites based on their detection according to ESI polarity: green strip if the signal was
352 observed in both modes; red only in negative mode and blue only in positive mode. It can be
353 observed that the green strips cover more than half length of the purple strip for each LC
354 condition which represents the majority of the metabolites. More metabolites are detected
355 for the ZIC-HILIC and the ZIC-cHILIC columns when increasing the pH from low to medium
356 (squares to triangles) and more still for the ZIC-pHILIC column with increasing the pH from
357 medium to high (triangles to stars) and the increase in the buffer concentration. It should be

358 also noted that with low pH mobile phases the numbers of the metabolite standards observed
359 only in ESI positive mode (the blue strip) were similar or slightly higher than in negative mode
360 (the red strip), whereas with medium and high pH mobile phases these numbers decreased in
361 ESI positive mode while greatly increased in negative mode. This change is attributed to the
362 deterioration of the chromatographic signals of basic metabolites and the improvement of
363 the peak shapes of acidic metabolites (e.g. organic acids with multiple carboxylic groups and
364 diphosphate or triphosphate nucleotides), respectively. With respect to the number of
365 metabolite standards showing good peak shapes (the purple strip), the best LC conditions
366 would be the ZIC-pHILIC column with high pH mobile phases and buffer concentrations (pH,
367 9.2; 10 mM or 20 mM in both solvents and 20 mM only in aqueous solvent).

368 The other necessary evaluation in LC condition optimisation for untargeted metabolite
369 profiling is the separation of isomeric metabolites; chromatographic separation is the only
370 reliable way to distinguish these compounds. There were totally 32 groups of isomeric
371 metabolites tested in this study and based on their acidity/basicity they were categorised as
372 acidic, basic, zwitterionic and neutral compounds (SM 2). In general, there were no
373 chromatographic factors showing clear influence on the separations. As shown in Figure 5,
374 more groups of isomeric metabolites were separated on the ZIC-pHILIC column especially with
375 high pH mobile phases. It appeared that the ionisable isomers would be more likely to be
376 separated in their completely ionised modes because ion-exchange could be involved to
377 greater extent in their interaction between mobile and stationary phase in HILIC mode. For
378 example, betaine is permanently positively charged; therefore it was separated from valine
379 where the ionisation of the amine group is dependent on pH underalmost all LC conditions.
380 This is also reflected by the observation that more acidic isomers (blue column in Figure 5)
381 were separated at medium or high pH than at low pH. For the neutral isomers (sugars) the
382 separations were difficult to achieve because hydrophilic partitioning was the only
383 chromatographic interaction for them under all LC conditions. The improvement in the
384 separation of acidic isomers on the ZIC-cHILIC column at medium pH could be attributed to
385 the unique selectivity for the negatively charged ions resulting from the positively charged
386 stationary phase surface formed by betaine groups and the smaller particle size (µm) might be
387 also contributing. It should be noted that the separation of some isomer groups was limited
388 by their poor peak shapes, such as broadening and tailing.

3.4 Evaluation of LC conditions using a biological extract

A cell extract and spent medium extract of the protozoan parasite *Leishmania major* were prepared as described in experimental section and tested under each LC condition with triplicate injections. The raw data were processed using MZMine 2.10 with the settings described in our previous study [27]. Of most interest was the number of LC-MS features generated under each LC condition. After removing the interference from the solvent blank the remaining LC-MS features were filtered again by the relative standard deviation (<25%) of the integrated peak areas obtained from three consecutive injections. By searching an in-house database the features were putatively identified based by accurate mass (±3 ppm). The numbers of total (Total: yellow strip) and the putatively identified (ID: purple strip) features under each LC condition obtained from the *L. major* cell extract and the spent medium are shown in Figures 6 and 7, respectively. According to their ESI detection polarity, the features were classified into positive (Pos: blue strip) and negative (Neg: red strip) and if two features were detected in both modes with mass difference of 2.0146 ± 0.001 m/z within an Rt window of ±0.2 min they were considered as a single feature (Both: green strip) and counted once in the calculation of total and identified features. These data are the number of repeatable features in the data processing without considering their real identities. Thus false-positive features caused by poor chromatographic signals as explained above, MS non-proton adducts (any MS signals not generated by M ± H), ESI in-source fragments, complex ions (e.g. dimers: 2M ± H) and salt cluster ions were all possibly included in counts. It is easily seen in both Figure 6 and 7 that the total numbers of repeatable features generated under C18-FA conditions from both cell extract and spent medium samples were much lower than for any of the HILIC conditions. This could be explained by the serious ion-suppression effect due to the many polar metabolites eluting early and together under this condition, consequently leading to poor repeatability in the MS response and so exclusion from the total by the RSD filter. For instance, after removing the interference from the blank 141 features eluting between 4.6 and 6.5 min remained for the cell extract sample under the C18-FA condition in ESI positive mode; whereas only 59 of them survived through the 25% RSD filter. Unstable ionisation efficiency caused by high portion of aqueous at the beginning of elution gradient under C18-FA condition could also obstruct the MS signal repeatability of the polar metabolites.

420 Among the HILIC conditions, it is obvious that the total number of metabolites detected
421 (yellow) are higher under the conditions of the ZIC-pHILIC column but only with a high pH
422 mobile phase. The same column with medium pH mobile phases, especially for cell extract
423 samples, did not out-perform the other columns, which means high pH might be the key
424 factor to achieve higher coverage of metabolites. A similar observation has been reported in a
425 previous column comparison study by Bajad *et al.* [23] who observed that an aminopropyl
426 column at pH 9.45 was the best LC condition for measuring soluble intracellular metabolites.
427 Based on the explanation previous outlined, the rise of the number of metabolites detected
428 only in ESI negative mode (red) results from the enhanced detection and/or the improved
429 peak shape of acidic metabolites like nucleotides in cell extract samples at high pH.
430 Metabolite numbers detected in both ESI modes (green) and putatively identified metabolites
431 (purple) also are greater under these conditions. It has been reported that ammonium
432 hydrogen carbonate is an excellent buffer for LC-MS analysis because of its good volatility
433 [35]. Ammonium carbonate was used in this study and its evaporation in ESI source appeared
434 to be faster than ammonium formate and ammonium acetate because much higher
435 background ion signals (dimer of formate: 91.00368 m/z, dimer of acetate: 119.03498m/z in
436 ESI negative mode and dimer of ACN: 83.06037 m/z in ESI positive mode) were observed if
437 ammonium formate or ammonium acetate were present but such background was not
438 observed for ammonium carbonate and also there was less evidence of deposition of
439 condensed salt in the ion source. Therefore it is likely that faster evaporation of ammonium
440 carbonate produced less ion-suppression effect and so facilitated the detection of the
441 metabolites present in low abundance. It was expected that MS responses of basic
442 metabolites would decrease with increasing mobile phase pH. In fact, however, the number
443 of metabolites detected only in ESI positive mode (blue) remained at a similar level or even
444 slightly increased, especially with high ammonium carbonate concentrations (AC-10mM-both
445 and AC-20mM). It might be due to the enhanced opportunity to generate ammonium adducts
446 ($M + NH_4^+$) under these conditions. Another finding was that the total number of detected
447 features under the conditions where the buffer was present in both solvents was always
448 slightly lower than when it was present only in aqueous solvent regardless of the other
449 factors.

450 Considering all the evaluations above using standard metabolites and parasite samples, the
451 best HILIC condition could be either pHILIC-AC-10 mM-both or pHILIC-AC-20 mM. However, it
452 should be noted that the poor chromatographic signals of highly basic metabolites like
453 polyamines could not be avoided under these conditions and these would produce false-
454 positive features in the data processing. The complementary coverage of metabolites under
455 C18-FA condition to these two HILIC conditions was also investigated. Approximately 80 and
456 167 features were exclusively detected under C18-FA conditions for the cell extract and the
457 spent medium samples, respectively. However, based on their m/z only 17 and 36 of these
458 could be putatively identified as metabolites suggesting that many of these features might be
459 false-positives. However, some isomers showed improved resolution under the C18-FA
460 conditions, for example leucine and isoleucine.

461 **3.5 Investigation of reliability of LC-HRMS signals**

462 In an untargeted metabolite profiling study, it is optimal to record as many small but real and
463 to remove as much noise as possible. The application of signal-to-noise ratio sometimes is not
464 suitable for LC-HRMS data because in many cases there is no noise at all if the extracted ion
465 window is only ±3 ppm. Therefore the differentiation between small true signals and noise is
466 generally determined by the settings of two parameters in the peak extraction algorithm:
467 minimum peak height and noise level. In order to find the proper signal intensity for these
468 two settings, the original samples of cell extract and spent medium of *L. major* species were
469 diluted as described in experimental section and measured under the pHILIC-AC-10 mM-both
470 condition. By tracing some real signals confirmed by standard metabolites we would find the
471 intensity level at which they disappeared or were undistinguishable from the noise. The LC-
472 HRMS signals of three nucleotides in ESI negative mode were extracted with a ±3 ppm
473 window for the original and the diluted samples and are shown in Figure 8 with the labelled
474 retention times (RT), peak areas (PA) and heights (NL). The reason for selecting these
475 nucleotides was that there is no effect of ionisation efficiency on MS signals intensity because
476 they have a similar ionisable group. As can be seen above, at the intensity of 1.50E4 all the LC-
477 HRMS signals could be confirmed as real. Below 1.00E4, there was one signal (GTP: 1/3) which
478 looked like a chromatographic peak and two signals (UDP: 1/27 and GTP: 1/9) which were
479 barely recognised as chromatographic peaks. Below or close to 1.00E3, there were almost no
480 signals (GTP: 1/27 and 1/81) or only noises (UDP: 1/81). The evaluation was carried out with

481 more confirmed signals in this way and finally the noise level and the minimum peak height
482 were determined as 1.00E3 and 1.00E4, respectively. It should be noted that the MS response
483 could vary between instruments or with different settings, therefore we recommend that a
484 trial experiment such as this should always be carried out before running a whole set of
485 samples.

486 In LC-HRMS-based metabolomics studies, the level of a specific metabolite in a biological
487 sample is reflected by its chromatographic peak area or height. When comparing the levels of
488 this metabolite across multiple samples there is a question raising that whether the calculated
489 ratios are able to reflect a true fold difference, especially with weak LC-HRMS signals.
490 According to the intensity (peak height), the generated LC-HRMS features from the original
491 sample were classified as three levels: low (1.00~3.00E4), medium (3.00~9.00E4) and high
492 (>9.00E4) and based on the known dilution factors the reliability of the ratios calculated by
493 peak areas at different levels was investigated and the results are shown in Figure 9. At each
494 level the majority of the total LC-HRMS features showed R^2 value greater than 0.9. However,
495 the number of features falling into the tolerance range (70-130%) of correct ratios
496 dramatically deceased with increased dilution. At high level (green strip) only about 20% of
497 the features were in the expected ratio range with 81 times dilution. At low level (blue strip)
498 about 70% of the features even could not truly reflect 3-fold difference. This means in general
499 there was a reasonable linear dynamic trend with dilution but that the calculated ratios could
500 not practically reflect a real fold difference with large dilution factors especially when the LC-
501 HRMS signals are low, which could be due to the loss of peak shape as can be observed as
502 UDP: 1/27 and GTP: 1/9 in Figure 8. Therefore in real studies the fold difference of a
503 metabolite across samples might be highly unreliable if the ratio is large plus calculated from
504 weak LC-HRMS signals. This is where targeted screens based on tandem mass spectrometry
505 are superior.

506 4. Conclusion

507 In this study on three zwitterionic columns the effect of varied HILIC factors on the retention
508 mechanism and the output of the data processing for metabolomics studies were evaluated
509 using metabolite standards and biological cell extract samples. In general, as well as
510 stationary phase chemistry the mobile phase pH plays an important role in determining the

retention behaviours of the metabolites. Due to the fact that ion-exchange is greatly involved with hydrophilic partition in the retention mechanism on these zwitterionic columns the chemical nature and the concentration of the buffer can also contribute. The ZIC-cHILIC column was tested for its applicability in metabolomics for the first time and in comparison with the ZIC-HILIC column it demonstrated improved chromatographic performance at low and medium pH, especially for acidic metabolites. With consideration of the quality of chromatographic signals, the separation of isomers and the number of LC-HRMS features generated in the biological samples, the optimum HILIC condition was using the ZIC-pHILIC column at pH 9.2 with 20 mM ammonium carbonate only in the aqueous mobile phase components. This can be mainly attributed to the high pH and the fast evaporation of the ammonium carbonate during the ESI process. By applying this method, clear differences between the intracellular metabolomes of three *Leishmania* species have been discovered and also between the composition of their suspending media after growth (manuscript in submission). With the Exactive Orbitrap MS used in this study, the minimum peak height and the noise level in data processing should be set to 1.00E4 and 1.00E3 respectively and care should be taken with regard to the reliability of weak LC-HRMS signals and avoiding multiplying these up unrealistically to obtain erroneous and misleading metabolite levels in whole cells.

References

[1] S.B. Milne, T.P. Mathews, D.S. Myers, P.T. Ivanova, H.A. Brown, Sum of the parts: mass spectrometry-based metabolomics, Biochemistry, 52 (2013) 3829-3840.

[2] S. Forcisi, F. Moritz, B. Kanawati, D. Tziotis, R. Lehmann, P. Schmitt-Kopplin, Liquid chromatography-mass spectrometry in metabolomics research: mass analyzers in ultra high pressure liquid chromatography coupling, Journal of chromatography. A, 1292 (2013) 51-65.

[3] H.G. Gika, G.A. Theodoridis, R.S. Plumb, I.D. Wilson, Current practice of liquid chromatography-mass spectrometry in metabolomics and metabonomics, Journal of pharmaceutical and biomedical analysis, 87 (2014) 12-25.

[4] N.L. Kuehnbaum, P. Britz-McKibbin, New Advances in Separation Science for Metabolomics: Resolving Chemical Diversity in a Post-Genomic Era, Chem Rev, 113 (2013) 2437-2468.

[5] G.A. Theodoridis, H.G. Gika, E.J. Want, I.D. Wilson, Liquid chromatography-mass spectrometry based global metabolite profiling: a review, Analytica chimica acta, 711 (2012) 7-16.

[6] R. Ramautar, G.W. Somsen, G.J. de Jong, CE-MS for metabolomics: developments and applications in the period 2010-2012, Electrophoresis, 34 (2013) 86-98.

544 [7] J.M. Buscher, D. Czernik, J.C. Ewald, U. Sauer, N. Zamboni, Cross-platform comparison of methods
545 for quantitative metabolomics of primary metabolism, Analytical chemistry, 81 (2009) 2135-2143.

546 [8] S. Cubbon, C. Antonio, J. Wilson, J. Thomas-Oates, Metabolomic applications of HILIC-LC-MS, Mass
547 spectrometry reviews, 29 (2010) 671-684.

548 [9] K. Spagou, H. Tsoukali, N. Raikos, H. Gika, I.D. Wilson, G. Theodoridis, Hydrophilic interaction
549 chromatography coupled to MS for metabonomic/metabolomic studies, Journal of separation science,
550 33 (2010) 716-727.

551 [10] D. Rojo, C. Barbas, F.J. Ruperez, LC-MS metabolomics of polar compounds, Bioanalysis, 4 (2012)
552 1235-1243.

553 [11] T. Zhang, D.G. Watson, High Performance Liquid Chromatographic Approaches to Mass
554 Spectrometry Based Metabolomics, Current Metabolomics, 1 (2013) 58-83.

555 [12] P. Jandera, Stationary and mobile phases in hydrophilic interaction chromatography: a review,
556 Analytica chimica acta, 692 (2011) 1-25.

557 [13] M.R. Gama, R.G.D. Silva, C.H. Collins, C.B.G. Bottoli, Hydrophilic interaction chromatography, Trac-
558 Trend Anal Chem, 37 (2012) 48-60.

559 [14] A. Periat, B. Debrus, S. Rudaz, D. Guillarme, Screening of the most relevant parameters for
560 method development in ultra-high,performance hydrophilic interaction chromatography, Journal of
561 Chromatography A, 1282 (2013) 72-83.

562 [15] A. Kumar, J.C. Heaton, D.V. McCalley, Practical investigation of the factors that affect the
563 selectivity in hydrophilic interaction chromatography, Journal of Chromatography A, 1276 (2013) 33-
564 46.

565 [16] Y. Kawachi, T. Ikegami, H. Takubo, Y. Ikegami, M. Miyamoto, N. Tanaka, Chromatographic
566 characterization of hydrophilic interaction liquid chromatography stationary phases: Hydrophilicity,
567 charge effects, structural selectivity, and separation efficiency, Journal of Chromatography A, 1218
568 (2011) 5903-5919.

569 [17] N.P. Dinh, T. Jonsson, K. Irgum, Probing the interaction mode in hydrophilic interaction
570 chromatography, Journal of Chromatography A, 1218 (2011) 5880-5891.

571 [18] R.I. Chirita, C. West, S. Zubrzycki, A.L. Finaru, C. Elfakir, Investigations on the chromatographic
572 behaviour of zwitterionic stationary phases used in hydrophilic interaction chromatography, Journal of
573 Chromatography A, 1218 (2011) 5939-5963.

574 [19] B. Preinerstorfer, S. Schiesel, M. Lammerhofer, W. Lindner, Metabolic profiling of intracellular
575 metabolites in fermentation broths from beta-lactam antibiotics production by liquid
576 chromatography-tandem mass spectrometry methods, Journal of chromatography. A, 1217 (2010)
577 312-328.

578 [20] D.V. McCalley, Study of the selectivity, retention mechanisms and performance of alternative
579 silica-based stationary phases for separation of ionised solutes in hydrophilic interaction
580 chromatography, Journal of Chromatography A, 1217 (2010) 3408-3417.

581 [21] D.P. Kloos, H. Lingeman, W.M. Niessen, A.M. Deelder, M. Giera, O.A. Mayboroda, Evaluation of
582 different column chemistries for fast urinary metabolic profiling, Journal of chromatography. B,
583 Analytical technologies in the biomedical and life sciences, 927 (2013) 90-96.

584 [22] T. Zhang, D.J. Creek, M.P. Barrett, G. Blackburn, D.G. Watson, Evaluation of coupling reversed
585 phase, aqueous normal phase, and hydrophilic interaction liquid chromatography with Orbitrap mass
586 spectrometry for metabolomic studies of human urine, Analytical chemistry, 84 (2012) 1994-2001.

587 [23] S.U. Bajad, W. Lu, E.H. Kimball, J. Yuan, C. Peterson, J.D. Rabinowitz, Separation and quantitation
588 of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass
589 spectrometry, Journal of chromatography. A, 1125 (2006) 76-88.

590 [24] R.I. Chirita, C. West, A.L. Finaru, C. Elfakir, Approach to hydrophilic interaction chromatography
591 column selection: Application to neurotransmitters analysis, Journal of Chromatography A, 1217
592 (2010) 3091-3104.

593 [25] E. Johnsen, S.R. Wilson, I. Odsbu, A. Krapp, H. Malerod, K. Skarstad, E. Lundanes, Hydrophilic
594 interaction chromatography of nucleoside triphosphates with temperature as a separation parameter,
595 Journal of chromatography. A, 1218 (2011) 5981-5986.

596 [26] T. Pluskal, T. Nakamura, A. Villar-Briones, M. Yanagida, Metabolic profiling of the fission yeast S.
597 pombe: quantification of compounds under different temperatures and genetic perturbation,
598 Molecular bioSystems, 6 (2010) 182-198.

599 [27] T. Zhang, D.G. Watson, L. Wang, M. Abbas, L. Murdoch, L. Bashford, I. Ahmad, N.Y. Lam, A.C. Ng,
600 H.Y. Leung, Application of Holistic Liquid Chromatography-High Resolution Mass Spectrometry Based
601 Urinary Metabolomics for Prostate Cancer Detection and Biomarker Discovery, PloS one, 8 (2013)
602 e65880.

603 [28] R. t'Kindt, R.A. Scheltema, A. Jankevics, K. Brunker, S. Rijal, J.C. Dujardin, R. Breitling, D.G. Watson,
604 G.H. Coombs, S. Decuypere, Metabolomics to unveil and understand phenotypic diversity between
605 pathogen populations, PLoS neglected tropical diseases, 4 (2010) e904.

606 [29] L. Chen, L. Zhou, E.C. Chan, J. Neo, R.W. Beuerman, Characterization of the human tear
607 metabolome by LC-MS/MS, Journal of proteome research, 10 (2011) 4876-4882.

608 [30] W. Jiang, G. Fischer, Y. Girmay, K. Irgum, Zwitterionic stationary phase with covalently bonded
609 phosphorylcholine type polymer grafts and its applicability to separation of peptides in the hydrophilic
610 interaction liquid chromatography mode, Journal of Chromatography A, 1127 (2006) 82-91.

611 [31] S. Di Palma, R. Raijmakers, A.J.R. Heck, S. Mohammed, Evaluation of the Deuterium Isotope Effect
612 in Zwitterionic Hydrophilic Interaction Liquid Chromatography Separations for Implementation in a
613 Quantitative Proteornic Approach, Analytical chemistry, 83 (2011) 8352-8356.

614 [32] R.A. Scheltema, S. Decuypere, R. T'Kindt, J.C. Dujardin, G.H. Coombs, R. Breitling, The potential of
615 metabolomics for Leishmania research in the post-genomics era, Parasitology, 137 (2010) 1291-1302.

616 [33] A.M. Silva, A. Cordeiro-da-Silva, G.H. Coombs, Metabolic variation during development in culture
617 of Leishmania donovani promastigotes, PLoS neglected tropical diseases, 5 (2011) e1451.

618 [34] R. t'Kindt, A. Jankevics, R.A. Scheltema, L. Zheng, D.G. Watson, J.C. Dujardin, R. Breitling, G.H.
619 Coombs, S. Decuypere, Towards an unbiased metabolic profiling of protozoan parasites: optimisation
620 of a Leishmania sampling protocol for HILIC-orbitrap analysis, Analytical and bioanalytical chemistry,
621 398 (2010) 2059-2069.

622 [35] A. Espada, A. Rivera-Sagredo, Ammonium hydrogencarbonate, an excellent buffer for the analysis
623 of basic drugs by liquid chromatography-mass spectrometry at high pH, Journal of chromatography. A,
624 987 (2003) 211-220.

625

626

627

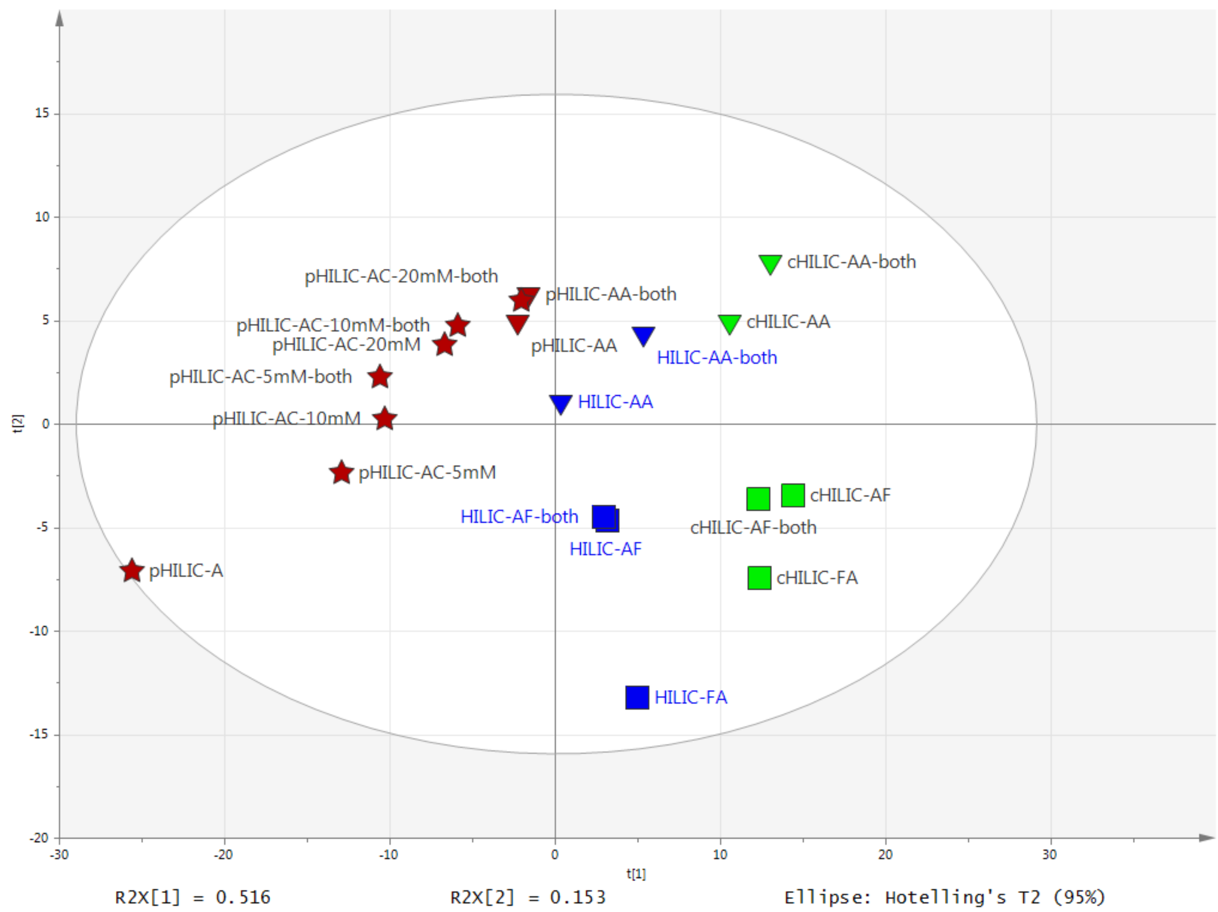
628 Table 1 Mobile phase conditions

Mobile phase		Column			
A	**B**	**ZIC-HILIC**	**ZIC-cHILIC**	**ZIC-pHILIC**	**C18-AR**
0.1% (v/v) FA in H_2O*	0.1% (v/v) FA in ACN	√	√	-	√
20mM AF pH=3 in H_2O*	ACN	√	√	-	-
10mM AF pH=3 in H_2O/ACN (9:1)[&]	10mM AF pH=3 in H_2O/ACN (1:9)	√	√	-	-
20mM AA pH=6.8 in H_2O*	ACN	√	√	√	-
10mM AA pH=6.8 in H2O/ACN (9:1)[&]	10mM AA pH=6.8 in H_2O/ACN (1:9)	√	√	√	-
0.05% (v/v) A in H_2O*	0.05% (v/v) A in ACN	-	-	√	-
5mM AC pH=9.2 in H_2O*	ACN	-	-	√	-
10mM AC pH=9.2 in H_2O*	ACN	-	-	√	-
20mM AC pH=9.2 in H_2O*	ACN	-	-	√	-
5mM AC pH=9.2 in H_2O/ACN (9:1)[&]	5mM AC pH=9.2 in H_2O/ACN (1:9)	-	-	√	-
10mM AC pH=9.2 in H_2O/ACN (9:1)[&]	10mM AC pH=9.2 in H_2O/ACN (1:9)	-	-	√	-
20mM AC pH=9.2 in H_2O/ACN (8:2)[#]	20mM AC pH=9.2 in H_2O/ACN (2:8)	-	-	√	-

629
630 FA: formic acid AF: ammonium formate AA: ammonium acetate
631 A: ammonium hydroxide AC: ammonium carbonate
632 *: Gradient program: mobile phase A was increased from 20% to 80% in 30mins and held at 92% for 5mins in
633 HILIC mode and A was decreased from 95% to 10% in 30mins and held at 10% for 5mins in RP LC mode.
634 &: Gradient program: mobile phase A was increased from 12.5% to 87.5% in 30mins and held at 100% for 5mins
635 #: Gradient program: mobile phase A was increased from 0% to 100% in 30mins and held at 100% for 10mins
636 √: The column was tested with this mobile phase.
637 -: The column was not tested with this mobile phase.
638

639

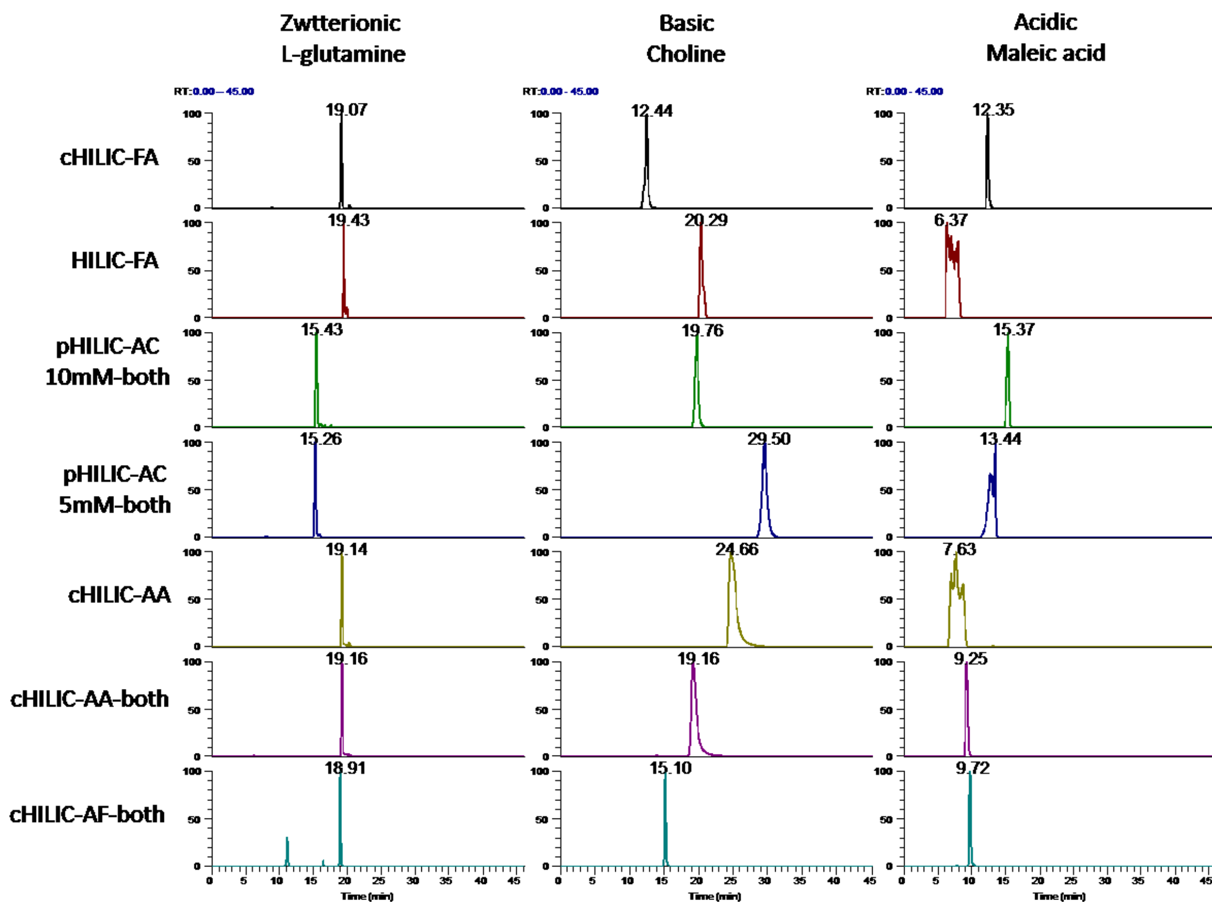
640 **Figure 1**. Principal Component Analysis (PCA) of 19 HILIC conditions with retention times of 641 223 metabolite standards (the first two components explain 66.9% of the total variance). 642 Column classification by colours: cHILIC, green; HILIC, blue; pHILIC, red. pH classification by 643 shapes: low, square; medium, triangle; high, stars. For more detailed HILIC conditions see 644 Table 1.



645

646

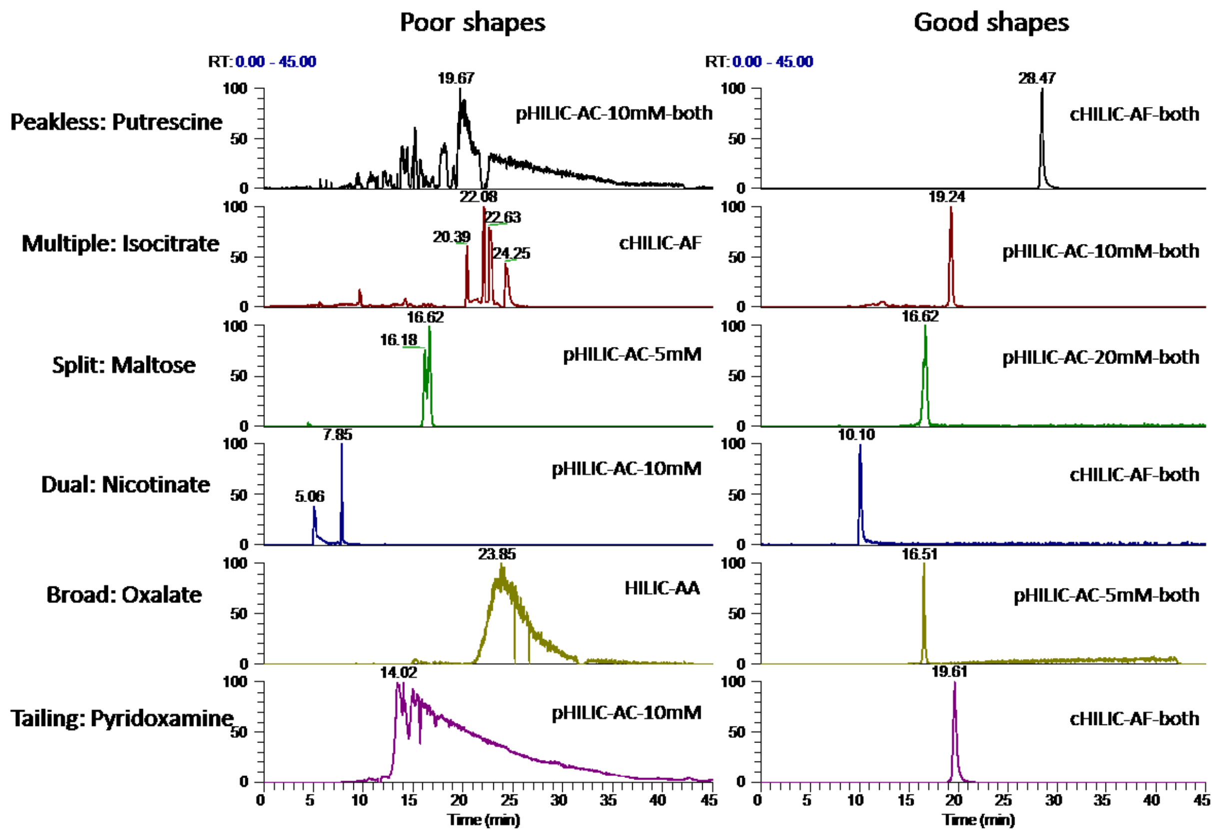
647 **Figure 2**. Extracted ion chromatograms (XIC: ±5 ppm) of one zwitterionic (L-glutamine), one
648 basic (choline) and one acidic (maleic acid) metabolite under different HILIC conditions.



649

650

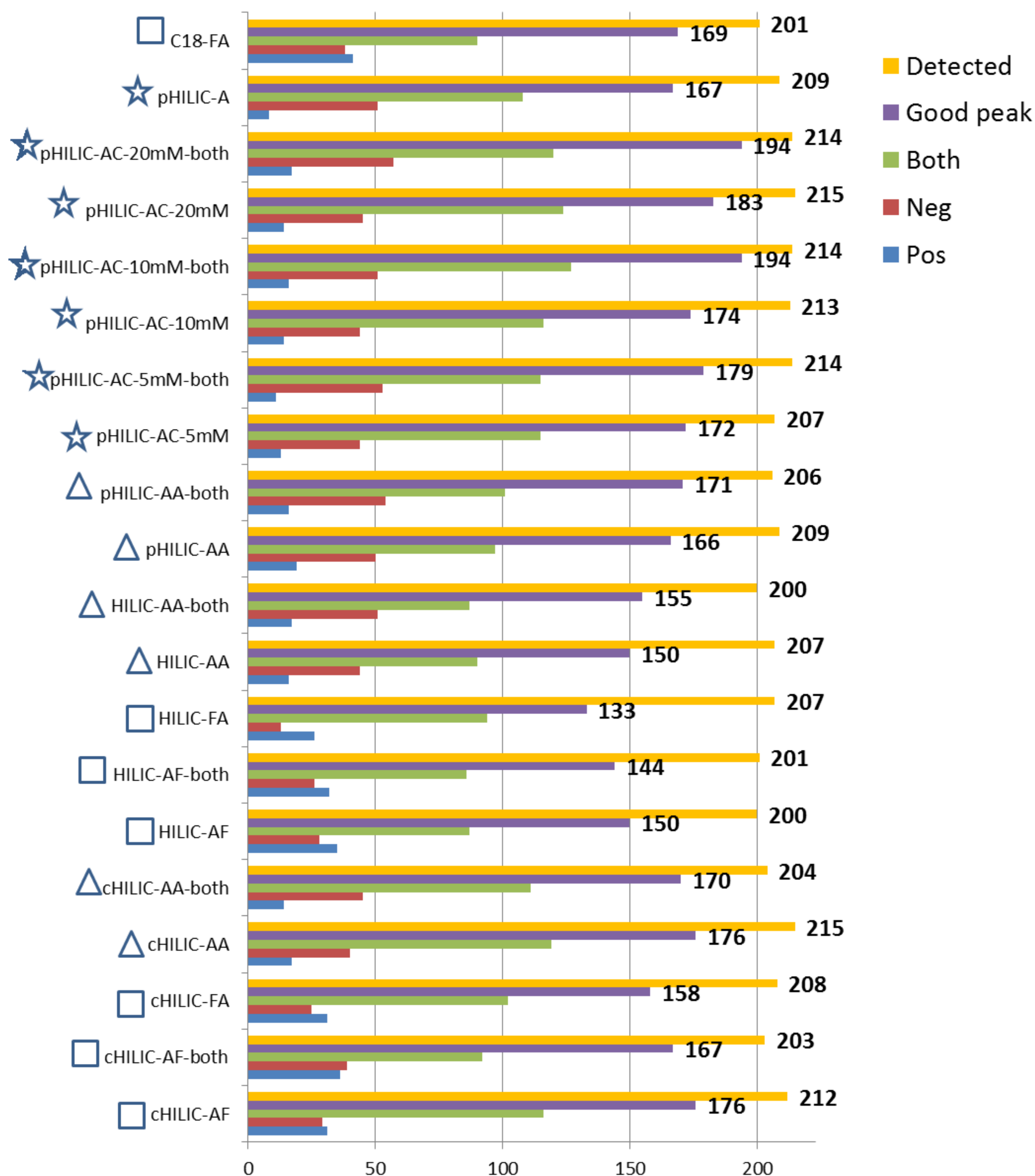
651 **Figure 3.** Examples of different types of poor chromatographic peak shapes compared with
652 good peaks for the same metabolite standard.



653

654

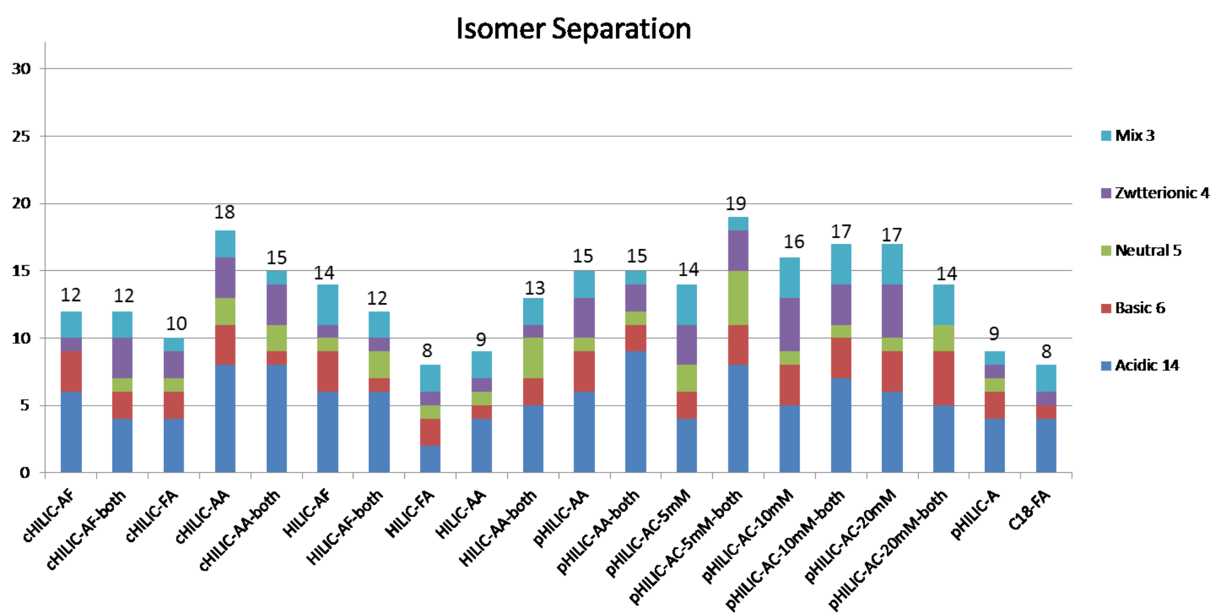
655 **Figure 4.** Bar chart of the numbers of the total detected metabolite standards (yellow); the
656 metabolite standards with good peak shapes (purple); the metabolite standards with good
657 peak shapes and detected in both ESI modes (green); the metabolite standards only detected
658 in negative mode (red) and in positive mode (blue) under different LC conditions. pH
659 classification by shapes: low, square; medium, triangle; and high, stars.



660

661

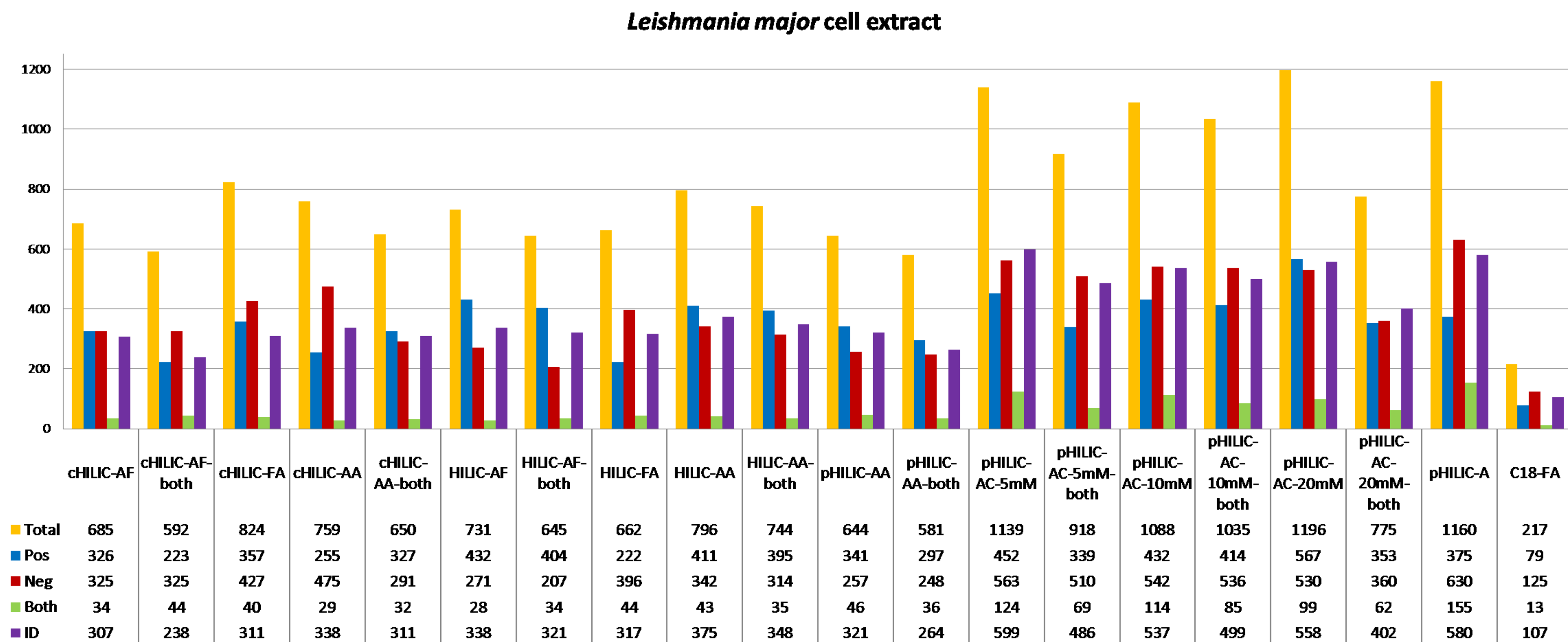
662 **Figure 5.** Column chart of the numbers of separated isomer groups under different LC
663 conditions with classification of their acidity/basicity (light blue, mixed; purple, zwitterionic;
664 green, neutral; red, basic; and deep blue, acidic).



665

666

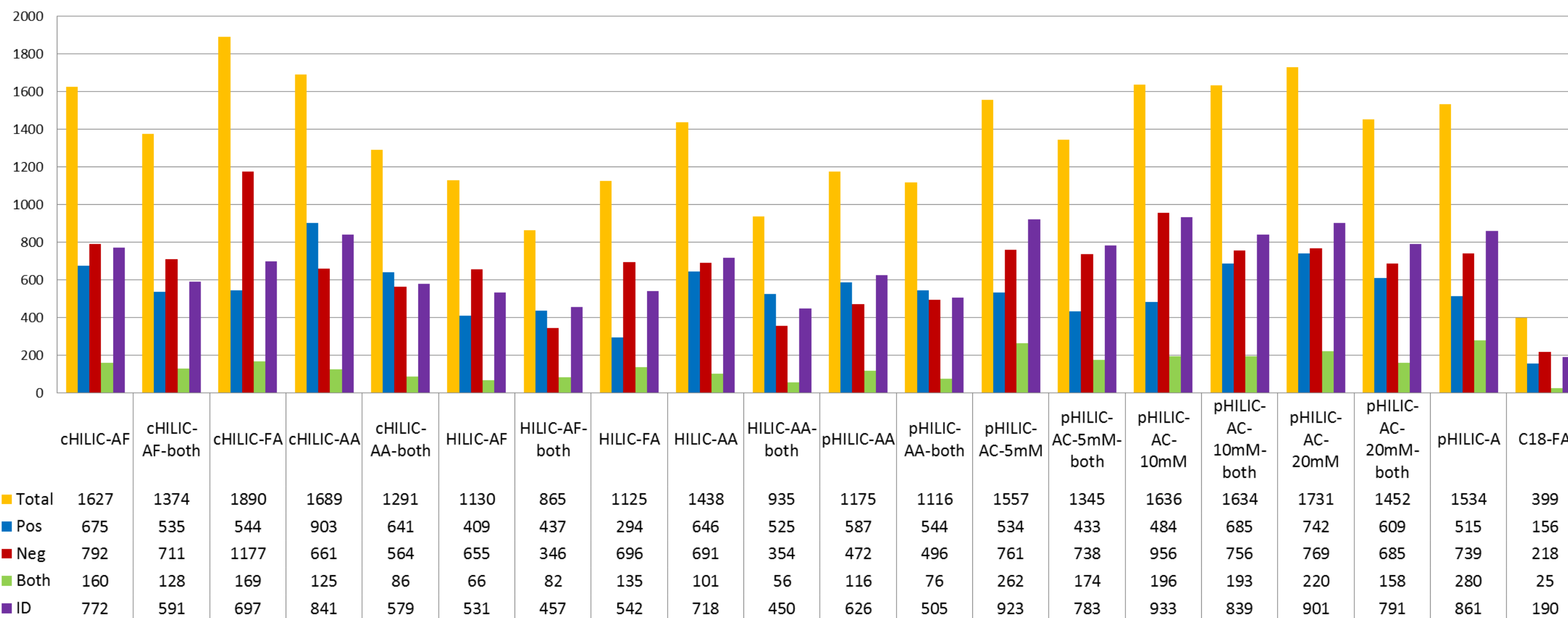
Figure 6. The numbers of total detected (yellow), only detected in ESI positive mode (blue), only detected in negative mode (red), detected in both modes (green), and putatively identifiable (purple) LC-HRMS features in the *L. major* extract sample under different LC conditions.



	cHILIC-AF	cHILIC-AF-both	cHILIC-FA	cHILIC-AA	cHILIC-AA-both	HILIC-AF	HILIC-AF-both	HILIC-FA	HILIC-AA	HILIC-AA-both	pHILIC-AA	pHILIC-AA-both	pHILIC-AC-5mM	pHILIC-AC-5mM-both	pHILIC-AC-10mM	pHILIC-AC-10mM-both	pHILIC-AC-20mM	pHILIC-AC-20mM-both	pHILIC-A	C18-FA
Total	685	592	824	759	650	731	645	662	796	744	644	581	1139	918	1088	1035	1196	775	1160	217
Pos	326	223	357	255	327	432	404	222	411	395	341	297	452	339	432	414	567	353	375	79
Neg	325	325	427	475	291	271	207	396	342	314	257	248	563	510	542	536	530	360	630	125
Both	34	44	40	29	32	28	34	44	43	35	46	36	124	69	114	85	99	62	155	13
ID	307	238	311	338	311	338	321	317	375	348	321	264	599	486	537	499	558	402	580	107

672

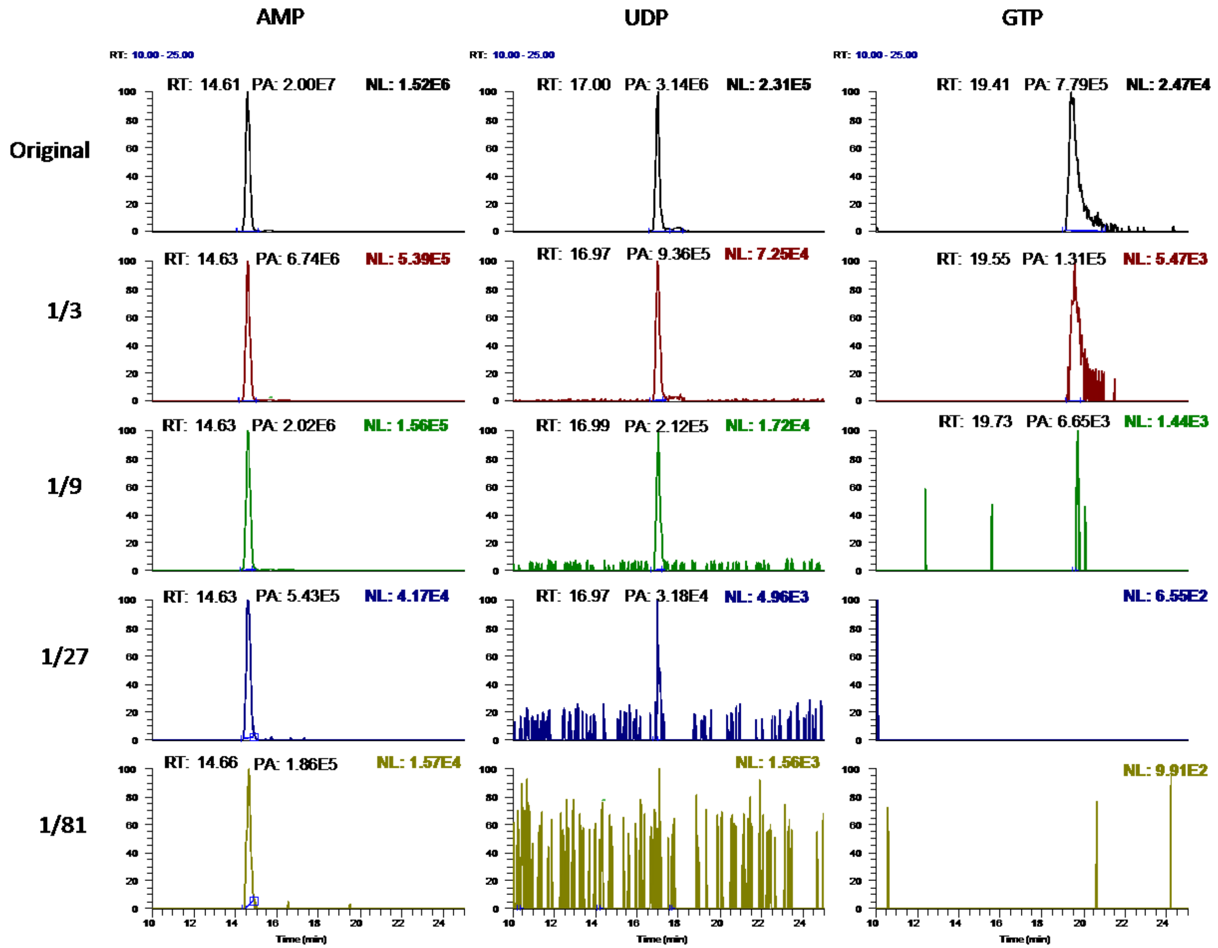
Figure 7. The numbers of total detected (yellow), only detected in ESI positive mode (blue), only detected in negative mode (red), detected in both modes (green), and putatively identifiable (purple) LC-HRMS features in the *L. major* spent medium sample under different LC conditions.



	cHILIC-AF	cHILIC-AF-both	cHILIC-FA	cHILIC-AA	cHILIC-AA-both	HILIC-AF	HILIC-AF-both	HILIC-FA	HILIC-AA	HILIC-AA-both	pHILIC-AA	pHILIC-AA-both	pHILIC-AC-5mM	pHILIC-AC-5mM-both	pHILIC-AC-10mM	pHILIC-AC-10mM-both	pHILIC-AC-20mM	pHILIC-AC-20mM-both	pHILIC-A	C18-FA
Total	1627	1374	1890	1689	1291	1130	865	1125	1438	935	1175	1116	1557	1345	1636	1634	1731	1452	1534	399
Pos	675	535	544	903	641	409	437	294	646	525	587	544	534	433	484	685	742	609	515	156
Neg	792	711	1177	661	564	655	346	696	691	354	472	496	761	738	956	756	769	685	739	218
Both	160	128	169	125	86	66	82	135	101	56	116	76	262	174	196	193	220	158	280	25
ID	772	591	697	841	579	531	457	542	718	450	626	505	923	783	933	839	901	791	861	190

677

678 **Figure 8.** Extracted ion chromatograms (XIC: ±5 ppm) of AMP (the left column), UDP (the
679 middle column) and GTP (the right column) from the *L. major* extract samples with different
680 dilution factors (labelled on the left for each row). RT, retention time; PA, peak area; NL, peak
681 intensity (height).



682

683

Figure 9. The numbers of LC-HRMS signals at low (1~3E4: blue), medium (3~9E4: red) and high (>9E4: green) intensity level in the original *L. major* extract (A) and spent medium sample (B) and in the diluted samples (3x, 9x, 27x and 81x dilutions). The data illustrate the numbers of the signals at each level showing R^2>0.9 across 1, 3, 9, 27 and 81 dilutions and the numbers falling into the expected ratio range under each dilution.