



**HAL**  
open science

## Optimization of ion trajectories in a dynamically harmonized Fourier-Transform Ion Cyclotron Resonance cell using a Design of Experiments strategy

Julien Maillard, Justine Ferey, Christopher P. Rüger, Isabelle Schmitz-Afonso, Soumeya Bekri, Thomas Gautier, Nathalie Carrasco, Carlos Afonso, Abdellah Tebani

### ► To cite this version:

Julien Maillard, Justine Ferey, Christopher P. Rüger, Isabelle Schmitz-Afonso, Soumeya Bekri, et al.. Optimization of ion trajectories in a dynamically harmonized Fourier-Transform Ion Cyclotron Resonance cell using a Design of Experiments strategy. *Rapid Communications in Mass Spectrometry*, 2020, 34 (7), pp.e8659. 10.1002/rcm.8659 . insu-02395300

**HAL Id: insu-02395300**

**<https://insu.hal.science/insu-02395300v1>**

Submitted on 8 Mar 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Optimization of ion trajectories in a dynamically  
2 harmonized Fourier-Transform Ion Cyclotron  
3 Resonance cell using a Design of Experiments  
4 strategy  
5

6 Julien MAILLARD<sup>1,2</sup>, Justine FERREY<sup>2</sup>, Christopher P. Ruger<sup>2</sup>, Isabelle SCHMITZ-  
7 AFONSO<sup>2</sup>, Soumeya BEKRI<sup>3</sup>, Thomas GAUTIER<sup>1</sup>, Nathalie CARRASCO<sup>1</sup>, Carlos  
8 AFONSO<sup>2</sup> and Abdellah TEBANI<sup>2,3\*</sup>

9  
10 <sup>1</sup> LATMOS/IPSL, Universite Versailles St Quentin, UPMC Universite Paris 06, CNRS, 11  
11 blvd d'Alembert, F-78280 Guyancourt, France

12 <sup>2</sup> Universite de Rouen, Laboratoire COBRA UMR 6014 & FR 3038, IRCOF, 1 Rue  
13 Tesniere, 76821 Mont St Aignan Cedex, France

14 <sup>3</sup> Department of Metabolic Biochemistry, Rouen University Hospital, Rouen, 76000,  
15 France

16

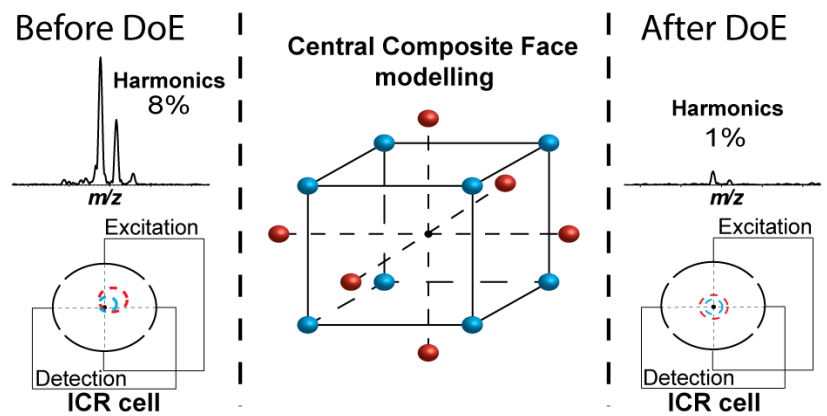
## 17 **Abstract**

18 The optimization of the ion trajectories in the ion cyclotron resonance (ICR) cell of a  
19 Fourier transform ICR mass spectrometer is a crucial step to obtain the best dynamic range,  
20 mass resolution, and mass accuracy. With the recent introduction of the dynamically  
21 harmonized cell, the complexity of tuning expanded drastically, and a fine-tuning of the  
22 DC voltages is required to optimize the ion cloud movement. This adjustment is typically  
23 performed manually.

24 Here, we propose a computational method based on a design of experiments (DoE) strategy  
25 to overcome the limits of classical manual tuning. This DoE strategy was exemplarily  
26 applied on a 12T FTICR equipped with a dynamically harmonized ICR cell. The  
27 chemometric approach, based on a composite central face design (CCF), was first applied  
28 on a reference material (sodium trifluoroacetate) allowing for the evaluation of the primary  
29 cell parameters. Eight factors were identified related to shimming and gating. The summed  
30 intensity of the signal corresponding to the even harmonics was defined as one quality  
31 criteria.

32 Consequently, the DoE response allowed for rapid and complete mapping of cell  
33 parameters resulting in an optimized parameter set. The new set of cell parameters was  
34 applied to the study of an ultra-complex sample. Here, Tholins, an ultra-complex mixture  
35 that mimics the haze present on Titan, was chosen. We observed a substantial improvement  
36 in mass spectrometric performance. The sum of signals related to harmonics was decreased  
37 by a factor of three (from 4% for conventional tuning to 1.3%). Furthermore, the dynamic  
38 range was also increased and lead to an increase of attributed peaks by 13%.

39 **Graphical abstract**



40

41

## 42 1. Introduction

43 Fourier Transform Ion Cyclotron Resonance mass spectrometer (FTICR) is well known to  
44 be the most powerful mass spectrometer by reaching unprecedented resolving power, mass  
45 accuracy and sensitivity levels<sup>1,2</sup>. It turned out to be the most appropriate instrument for the  
46 characterization of ultra-complex mixtures such as petroleum<sup>3-7</sup>, natural organic matter<sup>8</sup> or  
47 extraterrestrial organic matter<sup>9,10</sup>. In order to detect the cyclotron motion, FTICR uses the  
48 combination of a static electric field which is hyperbolic near the center and a strong  
49 homogeneous static magnetic field (up to 21 Tesla – so far). The combination of both fields  
50 retains ions in the cell. Ions in the magnetic field travel according to a cyclotron motion  
51 defined by equation 1.

$$52 \quad \omega_0 = \frac{qB_0}{m}$$

53 Equation 1: Cyclotron frequency ( $\omega_0$ ) of a charged particle in a homogeneous magnetic  
54 field.  $q$  is the charge of the particle,  $B_0$  the strength of the magnetic field and  $m$  the mass of  
55 the particle.

56  
57 In addition, the ion packet undergoes a drift of a trajectory equal to  $\mathbf{E} \times \mathbf{B}$  where  $\mathbf{E}$  is the  
58 electric field and  $\mathbf{B}$  is the magnetic field. Both of these movements allow the ion packet to  
59 be trapped in the cell<sup>2,11</sup>.

60 The shape of the ICR cell was improved steadily over time from first cubic geometry<sup>12</sup> to  
61 a variety of spherical designs varying in electrode shape and number<sup>13</sup>. As an example, the  
62 orthorhombic capacitively coupled cell developed by Steven Beu in 1991<sup>14</sup> or the infinity  
63 cell which was designed by Caravatti et al. in 1991 are cited here<sup>15</sup>, the later remaining the  
64 most commonly used nowadays. In recent instruments, the so-called dynamically  
65 harmonized cell is frequently deployed<sup>2</sup>. For the purpose of radiofrequency excitation and  
66 detection of ions, the cell is composed of several electrodes. In case of the dynamically

67 harmonized cell, these excitation and detection electrodes are specially shaped to  
68 compensate the quadrupolar electric field over the entire cell's body. It allows to produce  
69 a volume where, over the orbit of the ions, the average field is quadrupolar at all cyclotron  
70 radii. Thus, the frequency drifts of ions of the same mass over time along the Z axis are  
71 reduced. This makes possible to record transients for a longer time than in the previous  
72 cells and thus significantly improve performance

73 Optimized detection in terms of highest resolving power, sensitivity and dynamic range is  
74 obtained when the ion cloud is perfectly centered in the middle of the magnet. However, in  
75 practice, the imperfections of the alignment cause a bias in the position of the ion packets  
76 and thus a deviation in the cyclotron motion synchronization. This issue is solved by adding  
77 another set of segments to the dynamically harmonized cell in order to smooth  
78 imperfections by applying additional DC (Direct Current) voltages<sup>11</sup>. With this approach,  
79 it is possible to decrease the abundance of even-harmonic signals and, thus, to minimize  
80 magnetron motion. However, the values for the additionally applied DC voltages have to  
81 be tuned manually, which requires specially trained personnel and is tedious and time-  
82 consuming. In the present work, an experimental design approach is proposed in order to  
83 overcome the drawbacks of the manual ICR cell tuning. The approach seeks to optimize  
84 the DC voltages and reach the best set of parameters. Design of experiments (DoE) is a  
85 rigorous procedure in which defined modifications are induced based on a given set of  
86 variables. The response of the system is recovered to create a mathematical model, which  
87 allows a prediction of a set of parameters optimizing the response value<sup>16-18</sup>. Some example  
88 of mass spectrometer's optimization using DoE can be cite, such as the work performed by  
89 Lemonakis et al (2016) and Tebani et al (2016)<sup>19,20</sup>. Here, this method is applied on a  
90 defined response, the sum of the abundance of the second harmonic of sodium  
91 trifluoroacetate cluster.

92 After optimization, the obtained parameters are applied to the analysis of a complex  
93 mixture, Tholins. This synthetic material is used to understand the chemistry occurring in  
94 Titan's atmosphere by artificially mimicking the brown haze surrounding the largest moon  
95 of Saturn<sup>21-25</sup>. More importantly, Tholins have shown to be ultra-complex mixtures with  
96 high isobaric complexity and thousands of elemental compositions <sup>26</sup>. Tholins spectra  
97 obtained after DoE and after the usual manual method are compared. The potential of the  
98 developed method for complex mixtures analysis is highlighted evaluating the mass  
99 spectrometric performance, i.e., resolving power, the abundance of harmonics and number  
100 of detected signals.

101

## 102 **2. Experimental methods**

### 103 *2.1. Tholins production*

104 Tholins were produced following the PAMPRE procedure (French acronym for Aerosols  
105 Microgravity Production by Reactive Plasmas) presented elsewhere<sup>24,27</sup>. Briefly, in a  
106 stainless steel tubular reactor, a continuous gas mixture composed of 95 v-% nitrogen and  
107 5 v-% methane was injected through polarized electrodes and evacuated by a primary  
108 vacuum pump system. A Radio Frequency – Capacity Coupled Plasma (RF-CCP)  
109 discharge was established in the gas mixture deploying an RF of 13.56 MHz. The pressure  
110 in the plasma discharge was maintained at  $0.9 \pm 0.1$  mbar, and the reaction took place at  
111 room temperature (293 K).

### 112 *2.2. Instrumentation*

#### 113 *2.2.1. Fourier transform ion cyclotron resonance mass spectrometer*

114 All analyses were performed on a FTICR Solarix XR from Bruker equipped with a 12T  
115 superconducting magnet and a dynamically harmonized ICR cell. The following ion  
116 transfer parameters were used for both electrospray (ESI) and laser desorption ionization  
117 (LDI) analyses in positive ion mode: capillary exit 150V, deflector plate 200V, funnel1  
118 150V, skimmer1 25V, funnel RF amplitude 60Vpp, octopole frequency 5MHz, octopole  
119 RF amplitude 350Vpp, lower cut-off of the quadrupole at  $m/z$  120, time-of-flight 0.7ms,  
120 frequency TOF 6MHz, TOF RF amplitude 200Vpp, side kick offset -1V, front and back  
121 trapping plate 1.75V.

122 For ESI the following parameters were used: capillary 4.5kV, spray shield -500V, dry gas  
123 flow  $2\text{L}\cdot\text{min}^{-1}$ , dry gas temperature  $180^\circ\text{C}$ , nebulizer gas flow 0.5bar. All analyses were  
124 recorded on 20 scans with a quadrupole accumulation time of 0.1s. Electrospray spectra



125 were recorded with a mass range from  $m/z$  98 to 1,200 and a transient length of 2.2s. Sodium  
126 trifluoroacetate (NaTFA, Sigma Aldrich) at  $0.1\text{mg}\cdot\text{ml}^{-1}$  in ACN/H<sub>2</sub>O (50/50, v/v) was used  
127 as a standard for the design of experiments approach.

128 Complex Tholins mixture analyses was recorded applying LDI. The solid sample was  
129 deposited on a MALDI plate following previous work<sup>28</sup>. The third harmonic of a Nd:YAG  
130 laser at 355nm delivering a maximum output of 0.5mJ (Smartbeam II, Bruker) was used to  
131 ionize the samples with the following parameters: laser power 15%, laser shots per scan  
132 50, shots frequency 1kHz, plate offset 100V. The recorded mass spectra combined 800  
133 scans to achieve high signal-to-noise. LDI spectra were recorded with a mass range from  
134  $m/z$  110 to 1200 and a transient length of 4.3s.

### 135 *2.3. Experimental design*

#### 136 *2.3.1. Software*

137 DataAnalysis 4.4 (Bruker Daltonics GmbH, Bremen) was used to process all mass  
138 spectrometric analyses, including  $m/z$ -calibration and response recovering (abundance of  
139 the harmonic signals).

140 MODDE 11 software (Umetrics, Sartorius Stedim Data Analytics AB, Umea, Sweden) was  
141 used to perform the consecutive data analyses and modeling. The initial step is the  
142 definition of the response (i.e., user observation). Then, in a second step, factors influencing  
143 the response are defined. Afterward, several choices of experimental design shapes can be  
144 chosen. The software will automatically create the DoE given to the user as a coded matrix  
145 (**Table 1**) and a real matrix (**Table S1**). This matrix defines the experiments to be recorded  
146 and from which the response has to be extracted. Then, MODDE propose an automatic  
147 fitting method (e.g., partial least squares regression - PLS) allowing the visualization of the

148 quality and the predictability of the DoE. Finally, the optimizer tool was used to predict the  
149 set of parameters giving the best response result.

### 150 2.3.2. Factorial designs

151 The tuning of the ICR cell is established with eight principal parameters, four shimming  
152 parameters (shimming 0°, shimming 180°, shimming 90° and shimming 270°) and four  
153 gating settings (gated 0°, 90°, 180°, and 270°)<sup>11</sup>. The set of gating parameters represents  
154 the values of DC voltages applied during the injection in the ICR cell in order to correctly  
155 trap the ion packets. The set of shimming parameters represents the values of DC voltages  
156 applied for the excitation and the detection of the ions packets in the ICR cell. A simplified  
157 description of the ICR cell is given in **figure 1**. Other parameters of the ICR cell are not  
158 included (e.g., back and front plate voltages, sidekick offset) as they are easily and rapidly  
159 manually tunable. This significantly reduces the number of experiments required for the  
160 DoE approach, considerably decreasing the optimization time.

161 Parameters were grouped in pairs leading to four couples of parameters: shimming 0°-180°  
162 (Shi0), shimming 90°-270° (Shi90), gated 0°-180° (Gat0) and gated 90°-270° (Gat90)  
163 simplifying the optimization process compared to the conventional manual procedure. The  
164 values of these couples change symmetrically, meaning that when the value of one part is  
165 increased, the other part decreases about the same value. Initially, all values are set to 1.5V  
166 (default settings). As example, increasing the value of the shimming 0° about 50mV to  
167 1.55V results at a value of the shimming 180° of 1.45V. In addition to these four coupled  
168 parameters, the sweep excitation power (SEP) was added as a factor parameter for the first  
169 step of the optimization strategy. SEP has a crucial impact on the ion trajectories as it  
170 defines their radius. After defining the assessed experimental factors, it was necessary to  
171 define the desired response as a result of the optimization. Following the classical

172 optimization method, the intensity ratio between the ion of the NaTFA at  $m/z$  702.86324  
173 and its second harmonic at  $m/z$  351.43164 was recovered.

174 The optimization of the dynamically harmonized cell parameters was divided into two  
175 steps. The first screening step was performed with wide parameter intervals to evaluate the  
176 effects of each parameter on the defined response. The second part, named optimization  
177 experimental design, focused on narrower intervals to reach the best set of parameters. The  
178 applied chemometric approach for the two steps is based on a Composite Central  
179 Composite face design (CCF) which takes into account interactions between parameters.  
180 The coded CCF for the first step is given in **Table 1** formatted according to the required  
181 software input. Basically, the table reflects the conducted experimental plan, where -1  
182 means that the parameter was set at the respective low level (e.g., for Shi 0 at 1.465V) and  
183 +1 at the respective high level. Defined intervals for each parameter are given in **Table 2**  
184 for the screening and optimization step. These intervals have been defined quickly by  
185 varying each of the parameters until the intensity of the harmonic reaches a minimum level.  
186 This step takes about 5 minutes.

### 187 *2.3.3. Multivariate modeling*

188 Partial Least Squares regression (PLS) was used to model each CCF design and to find the  
189 best set of parameters regarding the response chosen using the optimizer tool implemented  
190 in MODDE 11. Thus, PLS coefficients are used to estimate factor effects. A positive value  
191 of the coefficient indicates an increase of the response upon increasing the corresponding  
192 parameter, whereas a negative value indicates a decrease of the response upon increasing  
193 the corresponding parameter. The larger the absolute value of the PLS coefficient, the  
194 higher the effect of the parameter on the response. Analysis of variance (ANOVA) applied  
195 to the sum of squares of the PLS regression coefficient was used to assess the significance

196 of the factor effect. For a statistically significant effect of a factor, the model  $p$ -value of  
197 ANOVA should not exceed 0.05<sup>29</sup>.

### 198 **3. Results and discussion**

#### 199 *3.1. Screening experimental design using NaTFA*

200 The design of experiments approach was performed using NaTFA to minimize the  
201 abundance of signals related to harmonics. Briefly, the influence of four pairs of factors  
202 and the SEP (Shi0, Shi90, Gat0, Gat90, and Sweep Excitation Power) was evaluated using  
203 an even harmonic peak intensity of the ion  $m/z$  702.86324 at  $m/z$  351.43164 as the response.  
204 A Central Composite Face design (CCF) was chosen as it is generally applicable for this  
205 type of optimization problems.

206 **Figure 2a** visualizes the effect of the different factors on the even harmonic abundance in  
207 the screening step. SEP, Shi0, and Shi90 revealed the most significant effects on the  
208 intensity of the second harmonic peak, with contribution factors of 28.0%, 15.4%, and  
209 33.3% respectively. SEP played a positive effect with high values leading to high  
210 abundance of the harmonic. In contrast, Shi0 and Shi90 showed a negative response, with  
211 high values resulting in low abundance of the harmonic. In addition, the interaction of Shi0  
212 and Shi90 provided a significant positive effect, which is related to a high separation effect  
213 of these parameters. Hence these two couples of parameters have two very different effects  
214 on the abundance of the harmonic and, thus, on the ion cloud trajectories.

215 The prediction plot obtained based on the 27 experimental data points, between the  
216 predicted and observed values of the abundance of the second harmonic, is shown in **Fig.**  
217 **2b**. This representation allows to access the quality of prediction with this model based on  
218 the initial data set. The cumulative modeled variation ( $R^2X = 99.1\%$ ) and the cumulative  
219 predicted variation ( $Q^2Y = 62.8\%$ ) reflect the explained variance and the predictability of

220 the model. The near to one  $R^2X$  value proves the validity of the model. The  $Q^2Y$  value  
221 (which is above 50%) proves that the model can give predictable values of parameters for  
222 the defined response<sup>17,18,29</sup>. However, this value could be improved to obtain better  
223 optimization results, which will be the aim of the following part of the optimization  
224 experimental design. The values of the regression coefficients and their respective  $p$ -values  
225 are given in **Table S-3** (Supplementary data). The fitted models' metrics, ANOVA and  
226 validation results of each model are presented in **Tables S-4** and **S-5** (Supplementary data).  
227 Using this built model, the instrumental parameters were optimized to reach the lowest  
228 harmonic intensity. The predicted optimized values are presented in **Table 2**. Nevertheless,  
229 the obtained set of predicted values reach the extremity of each defined interval. Therefore,  
230 an additional optimization step was performed focusing on tight intervals for Shi0, Shi90,  
231 Gat0, and Gat90. It should be noted that the SEP and shimming values are two decorrelated  
232 parameters. In the rest of this study, the SEP was fixed at a value of 30% in order to focus  
233 on fine tuning. This value is a percentage of the maximum excitation amplitude.

### 234 *3.2. Optimization experimental design using NaTFA*

235 As described above, for this step, the sweep excitation power (SEP) was removed from the  
236 studied factors to focus only on the cell parameters. It turned out, that Shi0 and Shi90 had  
237 a major negative response on the second harmonic abundance (**Fig. 2c**), which is consistent  
238 with the results obtained in the previous screening step. Coded regression coefficients,  
239 ANOVA and model validation results are given in **Tables S-2, S-5, and S-6**  
240 (Supplementary data). The fitted model shows a cumulative modeled variation of  $R^2X =$   
241 99.5% and a cumulative predicted variation of  $Q^2Y = 88.9%$  as showed in **Fig. 2d**. These  
242 results improved compared to the screening step and proved the high predictability of the  
243 movement of ions in the dynamically harmonized cell. Using this built model, the

244 optimized set of parameters provided an abundance of 1.3% of second harmonics compared  
245 to the base signal. In comparison, the non-optimized parameters (manufacturer default  
246 values) provided 8% and manually optimized parameters provided 4% of harmonic  
247 abundance, respectively.

248 This first study proved the possible optimization of the movement of ions in the ICR cell  
249 by lowering harmonic peaks. This DoE methodology was validated by the study of a  
250 NaTFA standard solution typically used for instrument calibration and optimization.  
251 However, the high interest of optimizing the ion trajectories in the ICR cell is related to the  
252 description of complex mixtures with unknown constituents. Consequently, the set of  
253 parameters yielded by the DoE methodology have been applied to study Titan tholins. The  
254 results were compared to those obtained after the manual optimization.

### 255 *3.3. Application for the analysis of complex mixtures: Titan tholins*

256 The largest moon of Saturn, Titan, is surrounded by a thick and nitrogen-rich fog with some  
257 fraction of methane<sup>30-33</sup>. Understanding how this haze is produced along with its molecular  
258 composition are crucial steps from a planetary and prebiotic chemistry perspective<sup>34,35</sup>.  
259 Tholins are analogs of Titan's haze. They are laboratory-produced and used as a proxy on  
260 Earth as there is currently no sample return option. This material is an extremely complex  
261 mixture comparable to the complexity of petroleum and as such ideal for evaluating the  
262 tuning of the dynamically harmonized cell parameters<sup>26</sup>. **Figure 3** shows an overview of a  
263 tholins mass spectrum, containing more than 50,000 peaks. This mass spectrum was  
264 generated with LDI ionization together with the optimized cell tuning parameters resulting  
265 from the DoE approach on NaTFA. The insert in **Figure 3** presents the molecular  
266 complexity on a single nominal mass revealing more than 100 resolved peaks. In this study,

267 we used this ultra-complex sample to compare the effects of the above-described  
268 optimization strategy with the conventional method relying on manual optimization.

269 The comparison between the detected species applying the parameters given by manual  
270 tuning and DoE optimization is illustrated in **Figure 4**. The mass spectrum at the top was  
271 recorded after the conventional manual optimization of the ICR cell. This manual  
272 optimization can take several hours easily to obtain high performance due to the manual  
273 iteration process. The spectrum at the bottom was recorded after deploying the optimized  
274 parameter set given by the newly developed DoE strategy, which takes approximately one  
275 hour. It should be mentioned that the shimming parameters are re-optimized once a year to  
276 compensate instrumental's drift. Except for the cell parameters, the two spectra were  
277 recorded with the exact same conditions. The number of detected species is significantly  
278 higher with the DoE optimization. **Figure 4**, giving a small fraction of the mass  
279 spectrometric complexity revealed thirteen additional species (indicated with a red star) for  
280 the DoE optimized spectrum compared to the spectrum resulted after manual tuning. The  
281 attribution of these new peaks is given in **table S-7**. This observation can be induced by  
282 two effects: 1) The ion packets is more centered in the cell thanks to the computational  
283 method, inducing a higher dynamic range or, 2) in the spectrum after manual adjustment,  
284 some peaks are in coalescence with their close neighbor. This effect is not present after the  
285 DoE tuning, which explain why more peaks are observed in this case. In order to reinforce  
286 this, through attributions, it can be justified that these peaks are not harmonics. Manual  
287 optimization is a difficult and limited process. Indeed, it is impossible for the human eye to  
288 take into account the interactions between the various parameters. Therefore, the location  
289 of the ions is not perfectly centered and could still be improved. Furthermore, this method  
290 is a time-consuming iterative process based on trial and error steps to reach acceptable  
291 values specified by the manufacturer (here below 6% for the second harmonic). The above

292 described DoE strategy, performed with NaTFA, allows a complete mapping of the  
293 different parameter effects along with their interactions, which allows modeling of the  
294 system. Thus, this strategy leads to a more fine-tuning and a greater reduction of harmonics  
295 compared to the conventional manual attempt.

296 Different metrics obtained for DoE originated and manually recorded spectra are listed in  
297 **Table 3**. These results were obtained after recording 5 replicates for each tuning. As  
298 illustrated in **Figure 4** applying the DoE generated parameters significantly increased the  
299 number of observed peaks. The comparison with the manually tuned cell revealed 13%  
300 more peaks attributed after the DoE tuning. The resolving power is also slightly increased  
301 after the experimental design optimization: approx. 1,440,000 with the DoE method and  
302 1,350 000 with the manual method at  $m/z$  400. A slight increase is also observed for the  
303 average mass accuracy of the elemental composition assignment. All these results are  
304 summarized in **Figure 5**.

305 The application of the new predicted dynamically harmonized cell parameters on Titan's  
306 tholins showed a clear enhancement on almost every metric of the recorded spectrum in  
307 comparison with the spectrum generated with the non-tuned cell but also with the manually  
308 tuned cell. Furthermore, this work will be useful for the implementation of quadrupolar ion  
309 detection, which commonly induces higher order harmonics with increased abundance.  
310 This newly developed optimization strategy is highly time efficient compared to the  
311 conventional method, taking only one hour including the recording of the spectra, data  
312 processing, and data analysis. It is a universal approach and can be applied to any complex  
313 mixture such as petroleum and biological sample materials. Those ultra-complex mixtures  
314 usually raises high analytical challenges regarding experimental optimization.

315



316 **4. Conclusions**

317 The developed design of experiments (DoE) strategy has proven to be applicable for  
318 generating optimized parameter sets for a dynamically harmonized ICR cell. We could  
319 show that the approach results in settings leading to a reduction of harmonics to levels not  
320 reached by manual tuning and far below the manufacturer default values. Moreover, the  
321 optimization process takes significantly less time and can be performed by any instrument  
322 user. This substantial improvement allows for a better performance of the FTICR by  
323 significantly increasing all metrics of the recorded spectrum, i.e., mass accuracy, resolution  
324 and dynamic range, for both the standard as well as the ultra-complex mixture. Finally, this  
325 computational procedure based on a composite face design could be applied for any other  
326 mass spectrometric parameter optimization problem. The authors believe this will allow for  
327 a more transparent and more structured design when performing method development.

328

329

330 **5. Acknowledgments**

331 N.C. thanks the European Research Council for funding via the ERC PrimChem project  
332 (grant agreement No. 636829.).

333 Financial support from the National FT-ICR network (FR 3624 CNRS) for conducting the  
334 research is also gratefully acknowledged.

335 This work was also supported at COBRA laboratory by the European Regional  
336 Development Fund (ERDF N°31708), the Région Normandie, and the Labex SynOrg  
337 (ANR-11-LABX-0029) together with Normandie Université (NU), the Centre National de  
338 la Recherche Scientifique (CNRS), Université de Rouen Normandie (URN) and Innovation  
339 Chimie Carnot (I2C).

340

341 **6. References**

- 342 1. Smith RD, Cheng X, Bruce JE, Hofstadler SA, Anderson GA. TRAPPING,  
343 DETECTION AND REACTION OF VERY LARGE SINGLE MOLECULAR-  
344 IONS BY MASS-SPECTROMETRY. *Nature*. 1994;369(6476):137-139.
- 345 2. Nikolaev EN, Boldin IA, Jertz R, Baykut G. Initial experimental characterization  
346 of a new ultra-high resolution FTICR cell with dynamic harmonization. *Journal of*  
347 *the American Society for Mass Spectrometry*. 2011;22(7):1125-1133.
- 348 3. McKenna AM, Purcell JM, Rodgers RP, Marshall AG. Heavy Petroleum  
349 Composition. 1. Exhaustive Compositional Analysis of Athabasca Bitumen HVGO  
350 Distillates by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A  
351 Definitive Test of the Boduszynski Model. *Energy & Fuels*. 2010;24(5):2929-2938.
- 352 4. Smith DF, Podgorski DC, Rodgers RP, Blakney GT, Hendrickson CL. 21 Tesla FT-  
353 ICR Mass Spectrometer for Ultrahigh-Resolution Analysis of Complex Organic  
354 Mixtures. *Analytical chemistry*. 2018;90(3):2041-2047.
- 355 5. Rodgers RP, McKenna AM. Petroleum analysis. *Analytical chemistry*.  
356 2011;83(12):4665-4687.
- 357 6. Hughey CA, Rodgers RP, Marshall AG. Resolution of 11 000 Compositionally  
358 Distinct Components in a Single Electrospray Ionization Fourier Transform Ion  
359 Cyclotron Resonance Mass Spectrum of Crude Oil. *Analytical chemistry*. 2002;74.
- 360 7. Marshall AG, Rodgers RP. Petroleomics: chemistry of the underworld. *Proceedings*  
361 *of the National Academy of Sciences of the United States of America*.  
362 2008;105(47):18090-18095.
- 363 8. Li L, Fang Z, He C, Shi Q. Separation and characterization of marine dissolved  
364 organic matter (DOM) by combination of Fe(OH)<sub>3</sub> co-precipitation and solid phase  
365 extraction followed by ESI FT-ICR MS. *Analytical and bioanalytical chemistry*.  
366 2019;411(10):2201-2208.
- 367 9. Schmitt-Kopplin P, Gabelica Z, Gougeon RD, et al. High molecular diversity of  
368 extraterrestrial organic matter in Murchison meteorite revealed 40 years after its  
369 fall. *Proceedings of the National Academy of Sciences of the United States of*  
370 *America*. 2010;107(7):2763-2768.
- 371 10. Schmitt-Kopplin P, Harir M, Kanawati B, Tziozis D, Hertkorn N, Gabelica Z.  
372 Chemical footprint of the solvent soluble extraterrestrial organic matter occluded in  
373 soltmany ordinary chondrite. *Meteorites*. 2012:79-92.
- 374 11. Jertz R, Friedrich J, Kriete C, Nikolaev EN, Baykut G. Tracking the Magnetron  
375 Motion in FT-ICR Mass Spectrometry. *Journal of the American Society for Mass*  
376 *Spectrometry*. 2015;26(8):1349-1366.
- 377 12. Guan S, Marshall AG. Ion traps for Fourier transform ion cyclotron resonance mass  
378 spectrometry: principles and design of geometric and electric configurations.  
379 *International Journal of Mass Spectrometry and Ion Processes*. 1995;146-147:261-  
380 296.
- 381 13. Marshall AG, Hendrickson CL, Jackson GS. Fourier transform ion cyclotron  
382 resonance mass spectrometry: A primer. *Mass Spectrometry Reviews*.  
383 1998;17(1):1-35.
- 384 14. Beu SC, Laude DA. Open trapped ion cell geometries for Fourier transform ion  
385 cyclotron resonance mass spectrometry. *International Journal of Mass*  
386 *Spectrometry and Ion Processes*. 1992;112(2-3):215-230.

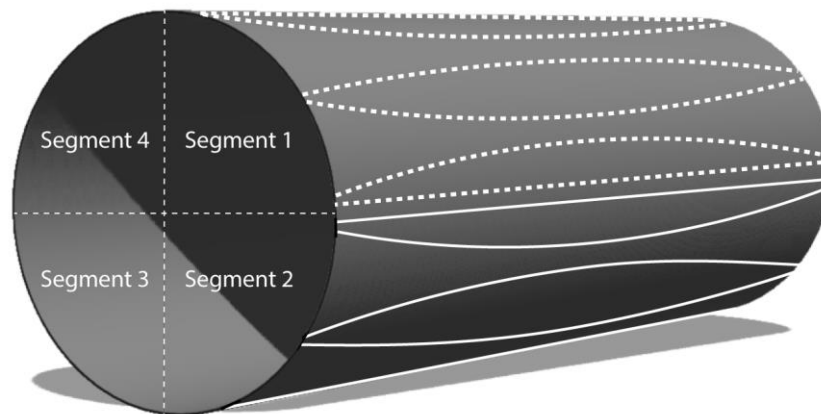
- 387 15. Caravatti P, Allemann M. The ‘infinity cell’: A new trapped-ion cell with  
388 radiofrequency covered trapping electrodes for fourier transform ion cyclotron  
389 resonance mass spectrometry. *Organic Mass Spectrometry*. 1991;26(5):514-518.
- 390 16. Gerretzen J, Szymanska E, Jansen JJ, et al. Simple and Effective Way for Data  
391 Preprocessing Selection Based on Design of Experiments. *Analytical chemistry*.  
392 2015;87(24):12096-12103.
- 393 17. Eliasson M, Rannar S, Madsen R, et al. Strategy for optimizing LC-MS data  
394 processing in metabolomics: a design of experiments approach. *Analytical  
395 chemistry*. 2012;84(15):6869-6876.
- 396 18. Eriksson L. Design of Experiments: Principles and Applications  
397 *Umetrics*. 2008.
- 398 19. Lemonakis N, Skaltsounis AL, Tsaibopoulos A, Gikas E. Optimization of  
399 parameters affecting signal intensity in an LTQ-orbitrap in negative ion mode: A  
400 design of experiments approach. *Talanta*. 2016;147:402-409.
- 401 20. Tebani A, Schmitz-Afonso I, Rutledge DN, Gonzalez BJ, Bekri S, Afonso C.  
402 Optimization of a liquid chromatography ion mobility-mass spectrometry method  
403 for untargeted metabolomics using experimental design and multivariate data  
404 analysis. *Analytica chimica acta*. 2016;913:55-62.
- 405 21. Khare BN, Sagnan C, Zumbege JE, Sklarew DS, Nagy B. Organic Solids Produced  
406 by Electrical Discharge in Reducing Atmospheres Tholin Molecular Analysis.  
407 *Icarus*. 1981;48:290-297.
- 408 22. Coll P, Coscia D, Smith N, et al. Experimental laboratory simulation of Titan's  
409 atmosphere: aerosols and gas phase. *Planetary and Space Science*. 1999;47:1331-  
410 1340.
- 411 23. Somogyi A, Oh CH, Smith MA, Lunine JJ. Organic environments on Saturn's  
412 moon, Titan: simulating chemical reactions and analyzing products by FT-ICR and  
413 ion-trap mass spectrometry. *Journal of the American Society for Mass  
414 Spectrometry*. 2005;16(6):850-859.
- 415 24. Szopa C, Cernogora G, Boufendi L, Correia JJ, Coll P. PAMPRE: A dusty plasma  
416 experiment for Titan's tholins production and study. *Planetary and Space Science*.  
417 2006;54(4):394-404.
- 418 25. Imanaka H, Smith MA. Formation of nitrogenated organic aerosols in the Titan  
419 upper atmosphere. *Proceedings of the National Academy of Sciences*. 2010;107.
- 420 26. Maillard J, Carrasco N, Schmitz-Afonso I, Gautier T, Afonso C. Comparison of  
421 soluble and insoluble organic matter in analogues of Titan's aerosols. *Earth and  
422 Planetary Science Letters*. 2018;495:185-191.
- 423 27. Gautier T, Carrasco N, Buch A, Szopa C, Sciamma-O'Brien E, Cernogora G. Nitrile  
424 gas chemistry in Titan's atmosphere. *Icarus*. 2011;213(2):625-635.
- 425 28. Barrere C, Hubert-Roux M, Lange CM, et al. Solvent-based and solvent-free  
426 characterization of low solubility and low molecular weight polyamides by mass  
427 spectrometry: a complementary approach. *Rapid communications in mass  
428 spectrometry : RCM*. 2012;26(11):1347-1354.
- 429 29. Montgomery D. Design and Analysis of Experiments. *John Wiley & Sons*. 2008.
- 430 30. Niemann HB, Atreya SK, Bauer SJ, et al. The abundances of constituents of Titan's  
431 atmosphere from the GCMS instrument on the Huygens probe. *Nature*.  
432 2005;438(7069):779-784.
- 433 31. Fulchignoni M, Ferri F, Angrilli F, et al. In situ measurements of the physical  
434 characteristics of Titan's environment. *Nature*. 2005;438(7069):785-791.

- 435 32. Lebreton JP, Witasse O, Sollazzo C, et al. An overview of the descent and landing  
436 of the Huygens probe on Titan. *Nature*. 2005;438(7069):758-764.
- 437 33. Israel G, Szopa C, Raulin F, et al. Complex organic matter in Titan's atmospheric  
438 aerosols from in situ pyrolysis and analysis. *Nature*. 2005;438(7069):796-799.
- 439 34. Sagan C. The origin of life in a cosmic context. *Origins of Life*. 1974;5(3-4):497-  
440 505.
- 441 35. Sagan C, Thompson WR, Khare BN. Titan: a laboratory for prebiological organic  
442 chemistry. *Accounts of Chemical Research*. 2002;25(7):286-292.

443

444

445 **7. Figure and tables**

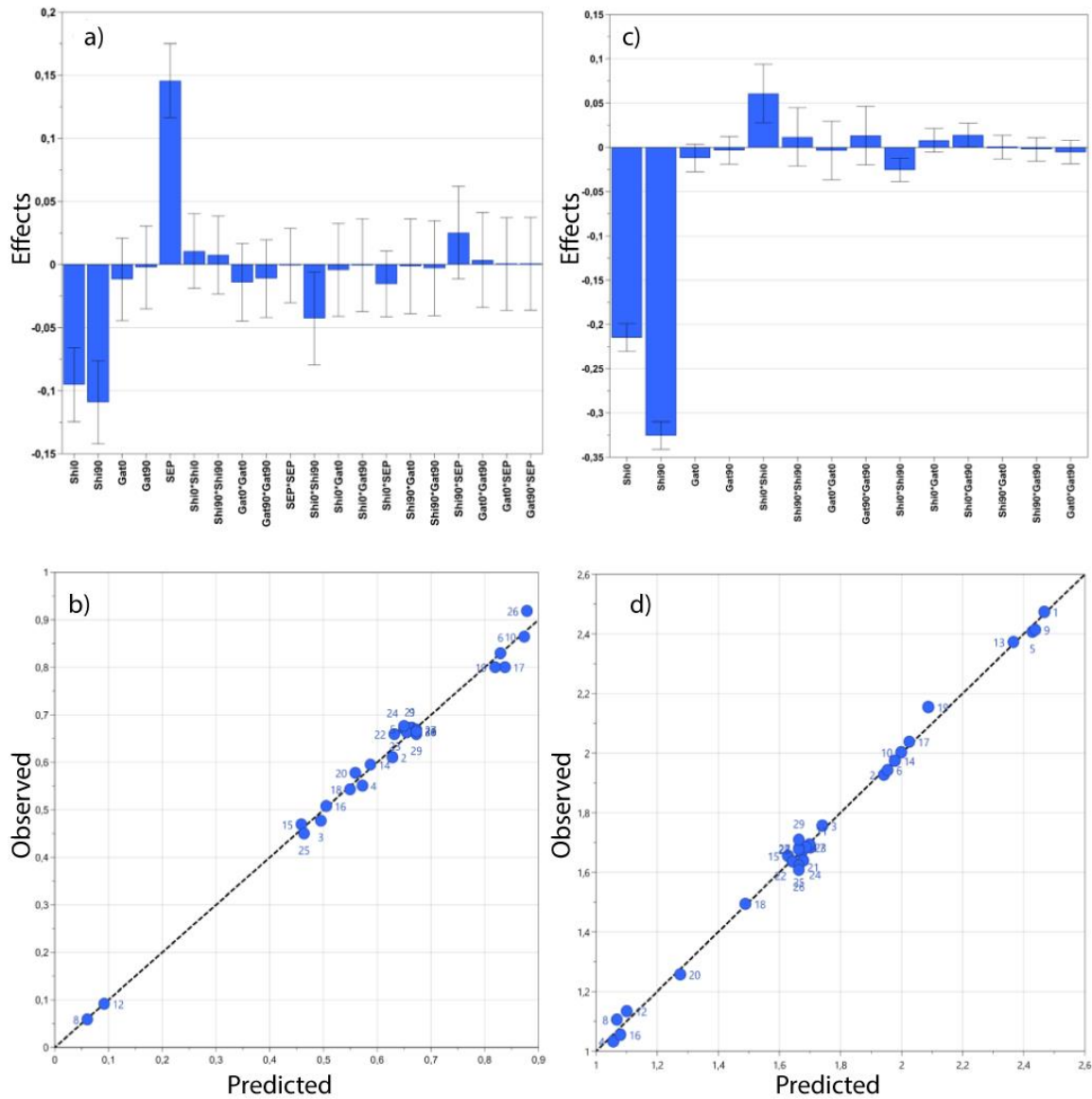


446

447 Figure 1: Schematic representation of a dynamically harmonized ICR cell. The cell is  
448 divided into four segments representing each adjustable parameter: Segment 1 (in dashed  
449 lines) is for shimming and gating  $0^\circ$ , segment 2 (in solid lines) for shimming and gating  
450  $90^\circ$ , segment 3 for shimming and gating  $180^\circ$  and segment 4 for shimming and gating  $270^\circ$ .

451

452

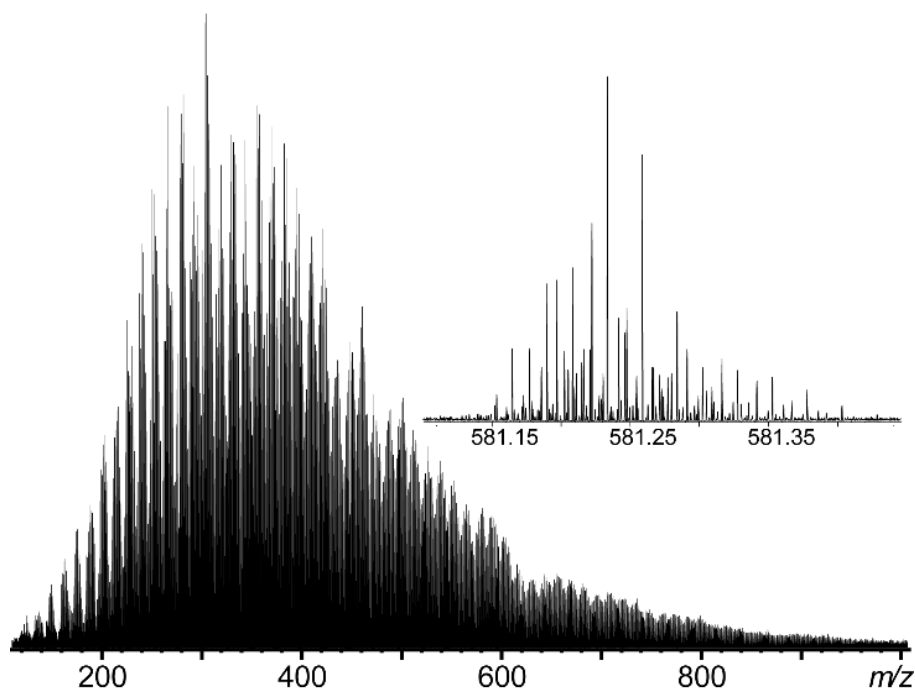


453

454 Figure 2: a) Observation of the effects of the different factors on the even harmonic  
 455 magnitude for the screening step, b) Prediction plot between the predicted and observed  
 456 values of the second harmonic magnitude for the screening step, c) Observation of the  
 457 effects of the different factors on the second harmonic magnitude for the optimization step,  
 458 d) Prediction plot between the predicted and observed values of the second harmonic  
 459 magnitude for the optimization step.

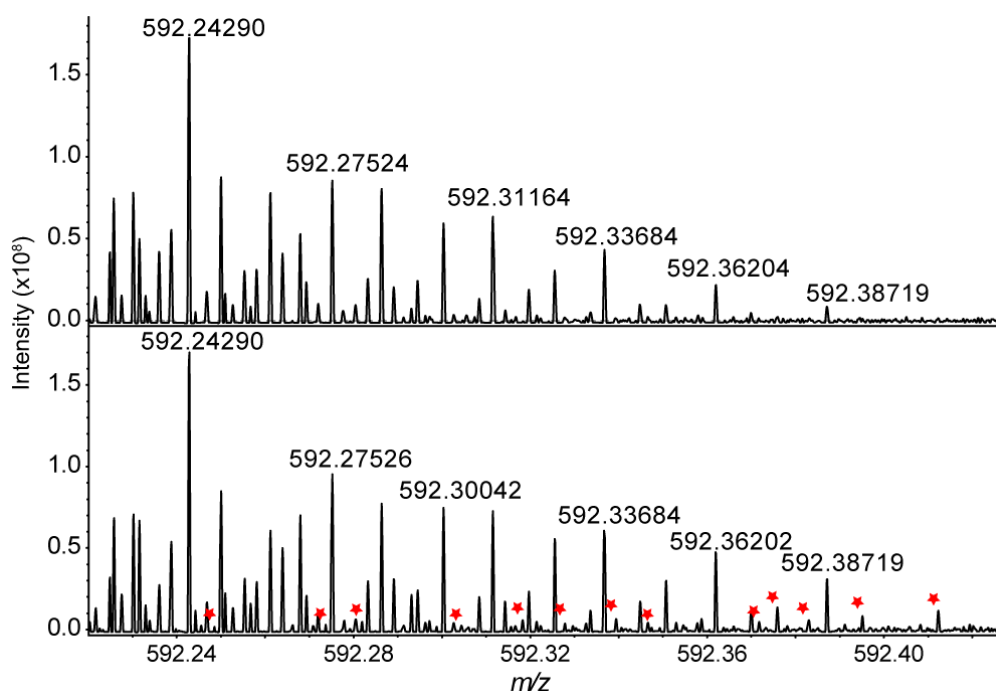
460

461



462

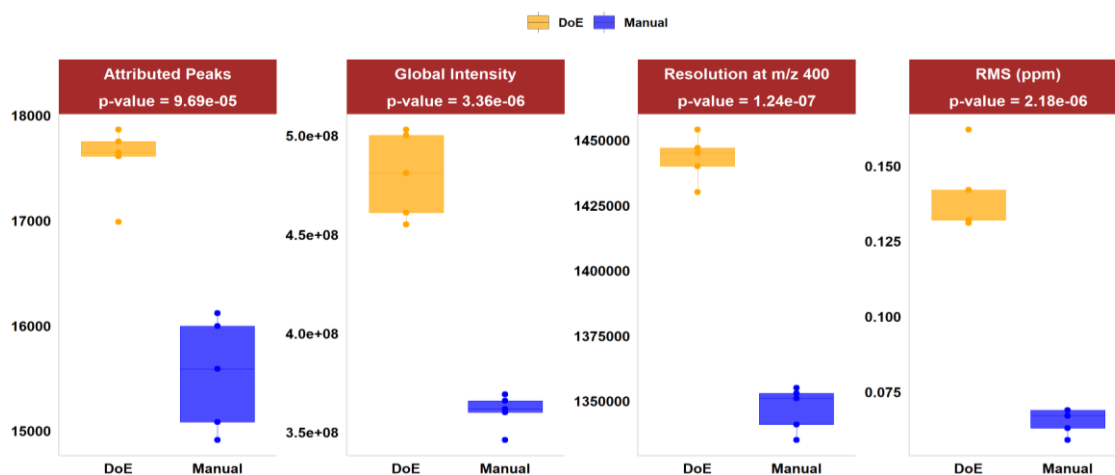
463 Figure 3: LDI spectrum in positive mode of a tholins sample obtained after deploying the  
 464 cell parameter optimized with the DoE approach, revealing more than 50,000 peaks. The  
 465 inset visualized a close zoom between  $m/z$  581.15 and  $m/z$  581.40 illustrating the high  
 466 isobaric complexity.



467

468 Figure 4: (Top) Zoom on the tholins spectrum obtained after a manual optimization  
 469 (shimming) of the ICR cell parameters. (Bottom) Tholins spectrum obtained after applying  
 470 the parameters given by the DoE optimization. Red dots indicate species found after DoE  
 471 optimization and not detected after manual optimization.  
 472





473  
 474 Figure 5: Comparison of attributed peaks, global intensity, resolution and RMS between  
 475 the manually optimized spectrum and the DoE spectrum of a tholins sample.  
 476

477 **Table 1: Coded composite face design (CCF) design matrix for experimental**  
 478 **screening on NaTFA**

Run N°	Run Order	Shi0	Shi90	Gat0	Gat90	SEP
1	10	-1	-1	-1	-1	1
2	28	1	-1	-1	-1	-1
3	23	-1	1	-1	-1	-1
4	9	1	1	-1	-1	1
5	1	-1	-1	1	-1	-1
6	24	1	-1	1	-1	1
7	20	-1	1	1	-1	1
8	12	1	1	1	-1	-1
9	29	-1	-1	-1	1	-1
10	8	1	-1	-1	1	1
11	6	-1	1	-1	1	1
12	5	1	1	-1	1	-1
13	25	-1	-1	1	1	1
14	16	1	-1	1	1	-1
15	26	-1	1	1	1	-1
16	17	1	1	1	1	1
17	2	-1	0	0	0	0
18	13	1	0	0	0	0
19	15	0	-1	0	0	0
20	19	0	1	0	0	0
21	4	0	0	-1	0	0
22	14	0	0	1	0	0
23	18	0	0	0	-1	0
24	22	0	0	0	1	0
25	3	0	0	0	0	-1
26	7	0	0	0	0	1
27	27	0	0	0	0	0
28	31	0	0	0	0	0
29	30	0	0	0	0	0
30	11	0	0	0	0	0
31	21	0	0	0	0	0

479

480

481 **Table 2: Predicted values for each factor for the screening and optimization**  
 482 **step**

483 **Screening step**

Factor	Intervals	Predicted Value	Factor contribution (%)
Shim0	1.465 to 1.535 (V)	1.535V	15.4
Shim90	1.465 to 1.540 (V)	1.540V	33.3
Gat0	1.4 to 1.620 (V)	1.620V	12.6
Gat90	1.450 to 1.550 (V)	1.550V	10.7
Sweep	20 to 40 (%)	20%	28.0

488 **Optimization step**

Factor	Intervals	Predicted Value	Factor contribution (%)
Shim0	1.530 to 1.550 (V)	1.542V	53.1
Shim90	1.530 to 1.550 (V)	1.550V	43.4
Gat0	1.600 to 1.640 (V)	1.600V	1.9
Gat90	1.540 to 1.570 (V)	1.551V	1.7

492  
 493 **Table 3. Comparison of the mass spectrometric response of the LDI tholins**  
 494 **spectra deploying manufacturer default values (before tuning), conventional**  
 495 **manual tuning (manual), and parameters given by the DoE optimization**  
 496 **method. p-values refers to the comparison between manual Tuning and the**  
 497 **DoE strategy. (p-value > 0.05 is considered as significant)**

	DoE (Mean)	Manual (Mean)	p-value	Fold Change
<b>Attributed Peaks</b>	1.76 10 <sup>+04</sup>	1.55 10 <sup>+04</sup>	9.69 10 <sup>-05</sup>	1.13
<b>Global Intensity</b>	4.80 10 <sup>+08</sup>	3.61 10 <sup>+08</sup>	3.36 10 <sup>-06</sup>	1.33
<b>Resolution at m/z 400</b>	1.44 10 <sup>+06</sup>	1.35 10 <sup>+06</sup>	1.24 10 <sup>-07</sup>	1.07
<b>RMS (ppm)</b>	0.140	0.0654	2.18 10 <sup>-06</sup>	2.14

498

499

500 **Supplementary Information for**  
501 Optimization of ion trajectories in a dynamically harmonized  
502 Fourier-Transform Ion Cyclotron Resonance cell using a Design of  
503 Experiments strategy

504

505 Julien MAILLARD<sup>1,2</sup>, Justine FERREY<sup>2</sup>, Christopher P. Ruger<sup>2</sup>, Isabelle SCHMITZ-  
506 AFONSO<sup>2</sup>, Soumeya BEKRI<sup>3</sup>, Thomas GAUTIER<sup>1</sup>, Nathalie CARRASCO<sup>1</sup>, Carlos  
507 AFONSO<sup>2</sup> and Abdellah TEBANI<sup>2,3\*</sup>

508

509 <sup>1</sup> LATMOS/IPSL, Universite Versailles St Quentin, UPMC Universite Paris 06, CNRS, 11  
510 blvd d'Alembert, F-78280 Guyancourt, France

511 <sup>2</sup> Universite de Rouen, Laboratoire COBRA UMR 6014 & FR 3038, IRCOF, 1 Rue  
512 Tesniere, 76821 Mont St Aignan Cedex, France

513 <sup>3</sup> Department of Metabolic Biochemistry, Rouen University Hospital, Rouen, 76000,  
514 France

515 **This PDF file includes:**

516 Tables S-1 to S-6

517 **Captions:**

518 • **Table S-1:** Fractional Factorial design matrix with response values for the  
519 screening step.

520 • **Table S-2:** Fractional Factorial design matrix with response values for the  
521 optimization step.

522 • **Table S-3:** PLS model regression coefficients generated by the screening step.

523 • **Table S-4:** ANOVA test results of the PLS model generated by the screening step.

524 • **Table S-5:** PLS model regression coefficients generated by the optimization step.

525 • **Table S-6:** ANOVA test results of the PLS model generated by the optimization  
526 step.

**Table S-1: Fractional Factorial design matrix with response values the screening step.**

Run N°	Run Order	Shim0 (V)	Shim90 (V)	Gat0 (V)	Gat90 (V)	SEP (%)	Second harmonic (%)
1	10	1.465	1.465	1.400	1.450	40	
2	28	1.535	1.465	1.400	1.450	20	4.083
3	23	1.465	1.540	1.400	1.450	20	2.996
4	9	1.535	1.540	1.400	1.450	40	3.558
5	1	1.465	1.465	1.620	1.450	20	4.694
6	24	1.535	1.465	1.620	1.450	40	6.77
7	20	1.465	1.540	1.620	1.450	40	
8	12	1.535	1.540	1.620	1.450	20	1.143
9	29	1.465	1.465	1.400	1.550	20	4.714
10	8	1.535	1.465	1.400	1.550	40	7.312
11	6	1.465	1.540	1.400	1.550	40	
12	5	1.535	1.540	1.400	1.550	20	1.237
13	25	1.465	1.465	1.620	1.550	40	
14	16	1.535	1.465	1.620	1.550	20	3.937
15	26	1.465	1.540	1.620	1.550	20	2.946
16	17	1.535	1.540	1.620	1.550	40	3.228
17	2	1.465	1.503	1.510	1.500	30	6.332
18	13	1.535	1.503	1.510	1.500	30	3.494
19	15	1.500	1.465	1.510	1.500	30	6.305
20	19	1.500	1.540	1.510	1.500	30	3.788
21	4	1.500	1.503	1.400	1.500	30	4.716
22	14	1.500	1.503	1.620	1.500	30	4.56
23	18	1.500	1.503	1.510	1.450	30	4.599
24	22	1.500	1.503	1.510	1.550	30	4.747
25	3	1.500	1.503	1.510	1.500	20	2.819
26	7	1.500	1.503	1.510	1.500	40	8.281
27	27	1.500	1.503	1.510	1.500	30	4.659
28	31	1.500	1.503	1.510	1.500	30	4.627
29	30	1.500	1.503	1.510	1.500	30	4.567
30	11	1.500	1.503	1.510	1.500	30	4.629
31	21	1.500	1.503	1.510	1.500	30	4.624

527

528

**Table S-2: Fractional Factorial design matrix with response values for the optimization step.**

Run N°	Run Order	Shim0 (V)	Shim90 (V)	Gat0 (V)	Gat90 (V)	Second harmonic (%)
1	27	1.530	1.530	1.600	1.540	2.475
2	2	1.550	1.530	1.600	1.540	1.928
3	20	1.530	1.550	1.600	1.540	1.757
4	4	1.550	1.550	1.600	1.540	1.033
5	5	1.530	1.530	1.640	1.540	2.407
6	23	1.550	1.530	1.640	1.540	1.942
7	12	1.530	1.550	1.640	1.540	1.684
8	24	1.550	1.550	1.640	1.540	1.107
9	19	1.530	1.530	1.600	1.570	2.414
10	6	1.550	1.530	1.600	1.570	2.003
11	26	1.530	1.550	1.600	1.570	1.693
12	3	1.550	1.550	1.600	1.570	1.135
13	1	1.530	1.530	1.640	1.570	2.373
14	11	1.550	1.530	1.640	1.570	1.975
15	28	1.530	1.550	1.640	1.570	1.654
16	29	1.550	1.550	1.640	1.570	1.054
17	17	1.530	1.540	1.620	1.555	2.039
18	18	1.550	1.540	1.620	1.555	1.493
19	8	1.540	1.530	1.620	1.555	2.156
20	10	1.540	1.550	1.620	1.555	1.256
21	21	1.540	1.540	1.600	1.555	1.665
22	22	1.540	1.540	1.640	1.555	1.637
23	16	1.540	1.540	1.620	1.540	1.684
24	15	1.540	1.540	1.620	1.570	1.639
25	25	1.540	1.540	1.620	1.555	1.623
26	14	1.540	1.540	1.620	1.555	1.608
27	13	1.540	1.540	1.620	1.555	1.682
28	7	1.540	1.540	1.620	1.555	1.676
29	9	1.540	1.540	1.620	1.555	1.708

**Table S-3: PLS model regression coefficients generated by the screening step.**

**Second harmonic**

	<b>Coeff. SC</b>	<b>Std. Err.</b>	<b>P</b>
<b>Shim0</b>	-0.08870	0.01303	0.00049
<b>Shim90</b>	-0.11241	0.01455	0.00025
<b>Gat0</b>	-0.00704	0.01455	0.64563
<b>Gat90</b>	0.00193	0.01455	0.89872
<b>SEP</b>	0.14123	0.01303	0.00004
<b>Shi0*Shi0</b>	-0.00072	0.01310	0.95773
<b>Shi90*Shi90</b>	0.00442	0.01369	0.75791
<b>Gat0*Gat0</b>	-0.01040	0.01368	0.47577
<b>Gat90*Gat90</b>	-0.00789	0.01368	0.58525
<b>SEP*SEP</b>	0.00569	0.01310	0.67912
<b>Shi0*Shi90</b>	-0.04575	0.01633	0.03106
<b>Shi0*Gat0</b>	-0.00388	0.01633	0.81987
<b>Shi0*Gat90</b>	-0.00052	0.01633	0.97560
<b>Shi0*SEP</b>	-0.01725	0.01162	0.18826
<b>Shi90*Gat0</b>	-0.00212	0.01669	0.90321
<b>Shi90*Gat90</b>	-0.00107	0.01669	0.95099
<b>Shi90*SEP</b>	0.03057	0.01633	0.11029
<b>Gat0*Gat90</b>	0.00019	0.01668	0.99123
<b>Gat0*SEP</b>	-0.00032	0.01633	0.98493
<b>Gat90*SEP</b>	0.00031	0.01633	0.98555

N = 27	Q2 = 0.628		RSD = 0.03919
DF = 6	R2 = 0.991		
Comp. = 4	R2 adj. = 0.959		Confidence = 0.95

530

**Table S-4: ANOVA test results of the PLS model generated by the screening step.**

**Second harmonic**

	<b>DF</b>	<b>SS</b>	<b>MS (variance)</b>	<b>F</b>	<b>p</b>
Total	27	11.0417	0.40895		
Total corrected	26	0.972344	0.0373979		
Regression	20	0.96313	0.0481565	31.3577	<b>0</b>
Residual	6	0.00921428	0.00153571		

N = 27	Q2 = 0.628	Cond. no. = 12.05
DF = 6	R2 = 0.991	RSD = 0.03919
Comp. = 4	R2 adj. = 0.959	

531  
532  
533  
534

**Table S-5: PLS model regression coefficients generated by the optimization step.**

<b>Second harmonic</b>			
	<b>Coeff. SC</b>	<b>Std. Err.</b>	<b>P</b>
<b>Constant</b>	1,66208	0,0126069	4,58E-23
<b>Shim0</b>	-0,214759	0,0072904	5,37E-14
<b>Shim90</b>	-0,325479	0,00729039	1,69E-16
<b>Gat0</b>	-0,0120755	0,0072904	0,119882
<b>Gat90</b>	-0,00339587	0,0072904	0,648523
<b>Shi0*Shi0</b>	0,0608795	0,0154223	0,00145892
<b>Shi90*Shi90</b>	0,0117455	0,0154223	0,458943
<b>Gat0*Gat0</b>	-0,00371134	0,0154223	0,813317
<b>Gat90*Gat90</b>	0,0132228	0,0154223	0,405672
<b>Shi0*Shi90</b>	-0,0255319	0,00619991	0,00104456
<b>Shi0*Gat0</b>	0,00798444	0,0061999	0,218685
<b>Shi0*Gat90</b>	0,0138399	0,0061999	0,0424476
<b>Shi90*Gat0</b>	0,000118387	0,0061999	0,985034
<b>Shi90*Gat90</b>	-0,0022854	0,0061999	0,717926
<b>Gat0*Gat90</b>	-0,00542158	0,0061999	0,39662
N = 29	Q2 = 0,889		
DF = 14	R2 = 0,995		RSD = 0,03858
Comp. = 5	R2 adj. = 0,991		Confidence = 0,95

535

**Table S-6: ANOVA test results of the PLS model generated by the optimization step.**

<b>Second harmonic</b>					
	<b>DF</b>	<b>SS</b>	<b>MS (variance)</b>	<b>F</b>	<b>p</b>
Total	29	92,3573	3,18473		
Total corrected	28	4,41765	0,157773		
Regression	14	4,39681	0,314058	211,033	0
Residual	14	0,0208347	0,00148819		
	N = 29	Q2 = 0,889		Cond. no. =	6,106
	DF = 14	R2 = 0,995		RSD =	0,03858
	Comp. = 5	R2 adj. = 0,991			

536

537



538 **Table S-7: Attribution of new peaks detected after the DoE optimization**

539

<b>Measured <i>m/z</i></b>	<b>Molecular formula</b>	<b>Err. (ppm)</b>
592.24866	C25H30N13O5	0.13
592.27374	C26H32N13O3 <sup>18</sup> O	-0.04
592.28203	C27H35N10O5 <sup>13</sup> C	-0.05
592.31836	C38H38N7	-0.06
592.32791	C19H41N14O7 <sup>15</sup> N	-0.14
592.33949	C25H43N10O6 <sup>13</sup> C	0.03
592.34673	C19H43N14O7 <sup>13</sup> C	-0.01
592.37013	C33H43N10 <sup>13</sup> C	-0.11
592.37597	C26H47N10O5 <sup>13</sup> C	-0.12
592.38314	C30H46N11O2	-0.17
592.39528	C35H47N8 <sup>13</sup> C	-0.17
592.41238	C37H50N7	-0.11