Off-line two-dimensional liquid chromatography separation for the quality control of saponins samples from Quillaja Saponaria

Lucile Lecas, Sylvie Nuccio, René Vaumas, Karine Faure

Off-line two-dimensional liquid chromatography separation for the quality control of saponins samples from *Quillaja Saponaria*

Lucile Lecas[1,2], Sylvie Nuccio[2], René de Vaumas[2], Karine Faure[1*]

[1]Université de Lyon, CNRS, Université Claude Bernard Lyon 1, Institut des Sciences Analytiques, UMR 5280, 5 rue de la Doua, F-69100 VILLEURBANNE, France

[2]Extrasynthese, Impasse Jacquard, F-69730 GENAY, France

*Corresponding author: karine.faure@isa-lyon.fr

Running title: off-line LC × LC for Quil-A quality control

Non-standard abbreviations:

1D LC: one-dimensional LC

[1]D: first dimension

[2]D: second dimension

RPLC: Reversed-phase liquid chromatography

HILIC: Hydrophilic interaction chromatography

QC: Quality control

LC x LC: comprehensive two-dimensional liquid chromatography

Keywords

Two-dimensional chromatography; Reversed-phase liquid chromatography; Hydrophilic interaction chromatography; Saponins; off-line optimization

Abstract

Quil-A is a purified extract of saponins with strong immunoadjuvant activity. While isolated molecules have been tested in clinical trials, Quil-A is mostly used as a totum of the Quillaja Saponaria bark extract. Quality control of the extract stability is usually based on the monitoring of specific saponins, whereas the comparison of samples with an initial chromatogram seems more appropriate. A reference fingerprint based on comprehensive two-dimensional liquid chromatography offers a rapid detection of non-conform samples. To fulfill quantity control constraints, off-line configuration was promoted. Hence, reversed-phase LC × reversed-phase LC and hydrophilic interaction chromatography × reversed-phase LC methods with single-quadrupole MS detection were kinetically optimized. The reversed-phase LC × reversed-phase LC method used a pH switch between dimensions to maximize orthogonality. Despite diagonalization, it led to a peak capacity of 831 in two hours. Moreover, the combination of hydrophilic interaction chromatography and reversed-phase LC offered a larger orthogonality but a lower, yet satisfactory peak capacity of 673. The advantages of both methods were illustrated on degraded samples, where the reversed-phase LC × reversed-phase LC contour plot highlighted the loss of fatty-acid chains, while the hydrophilic interaction chromatography × reversed-phase LC method attested enzymatic loss of sugar moieties.

1. Introduction

Plant extracts enter the composition of many cosmetic and pharmaceutical products and their bioactivity is most of the time the result of the combination of a large number of molecules [1]. The quality control (QC) of natural ingredients is very important at every stage of the production, from the incoming raw material to the stability of the plant material during storage. In general, the quantitative determination of selected markers is performed to ensure the quality of plant extracts. However, the lack of individual standards may reduce the number of monitored molecules. Moreover, the assay of such markers may not fully reflect the bioactivity of the totum. As such, the European guidelines for herbal medicine recommend to consider the whole of the plant extract as active substance [2]. Metabolite profiling using highly resolutive techniques provides a better understanding of the chemical profile. Unfortunately, in an industrial context, these powerful tools are often reserved for research purposes and are not available in routine QC laboratories. Qualitative chromatographic profiling of the samples with a comparison to a validated reference material may help rejecting non-conform samples in a timely manner. As so, the European guidelines state that the stability of the plant material "should be demonstrated by means of appropriate fingerprint chromatograms and its content should remain comparable to the initial fingerprint" [2].

Saponins are a class of naturally occurring molecules presenting surface-active or detergent properties. Saponins are commercially used in many industrial contexts: as foaming agent and emulsifier in food and cosmetics [3], as a supplement in the agri-food industry [4] or by pharmaceutical companies as vaccines adjuvants since some specific saponins are able to modulate the immune response [5, 6]. The bark of the tree *Quillaja Saponaria Molina* represents

one of the main saponin source as it contains up to 10 % of saponins. *Quillaja* saponins present unique structural features with a triterpenic aglycone core, usually quillaic acid but also phytolaccinic acid or gypsogenin, substituted with two saccharide groups that are further extended by oligosaccharides and with different degrees of acylation (generic structure available in supplementary material, Figure S1). This results in a very significant molecular complexity. Hundreds *Quillaja* saponins were discerned so far, using a combination of NMR [7], two-dimensional liquid chromatography coupled with high resolution mass spectrometry (LC × LC – HRMS) [8] or LC – MS/MS [9]. In 2019, 60 structures were identified from *Quillaja Saponaria* bark extracts [6]. The commercial Quil-A® product is a highly purified extract of *Quillaja Saponaria Molina* bark that presents a strong immunoadjuvant activity. It is considered as a reference material for many industrial uses, especially for animal vaccines [10]. Because of the large number of structurally close molecules present in such sample, the fingerprint chromatogram used in quality control has to offer a maximal peak capacity that one-dimensional LC methods hardly provide. Off-line comprehensive LC × LC can be implemented in routine QC laboratories, thanks to the limited instrumental requirements. Off-line 2D separations can offer as much peak capacity as online LC × LC, provided the overall separation is over 2 hours [11]. Moreover, off-line transfer allows fraction solvent evaporation, hence limiting band broadening due to injection effects and the subsequent dilution. In LC × LC, the increase in peak capacity over 1D LC is directly linked to the orthogonality offered by the two selected dimensions. Several LC × LC methods have been developed for the separation of saponins. While online HILIC x HILIC separation carried in 130 min has shown limited orthogonality [8], a RPLC × RPLC combination has been suggested, in which an improved orthogonality is provided thanks to a pH switch between dimensions [12]. Unfortunately, this off-line method lasts 12 hours, which is not compatible with a QC routine. On

the other hand, the combination of HILIC and RPLC mechanisms has been suggested to separate saponins [13], but has never been attempted on *Quillaja Saponaria*.

In this paper, an off-line comprehensive two-dimensional liquid chromatography method was developed in respect with QC constraints. Two combinations, namely RPLC × RPLC with pH switch and HILIC × RPLC were optimized with a maximum chromatographic duration of 120 min, taking into account the instrumental limitations of the fraction collection system. These two methods were evaluated in terms of orthogonality, peak capacity and dilution factor. Finally, the utility of the initial chromatographic fingerprint of Quil-A sample was demonstrated for the control of hydrolytically and enzymatically degraded samples.

2. Material and methods

2.1. Chemicals

LC-MS grade acetonitrile (ACN), methanol (MeOH) and 2-butanol (2-BuOH; ≥ 99.5%), disodium phosphate (Na_2HPO_4) and citric acid were purchased from Sigma Aldrich (Isle d'Abeau, France). Formic acid (FA) (LC-MS grade) and ammonium acetate (AA) (analytical reagent grade) were obtained from Fisher scientific (Waltham, MA, USA). Sodium bicarbonate was from Laurylab (Brindas, France) and hydrochloric acid from Carlo Erba Reagents (Val-de-Reuil, France). Water was purified by an Elga purification system (Veolia, Paris, France). Mobile phases (all proportion expressed as v/v) were filtered through 0.2 µm PTFE membranes (Merck, Germany) before use.

2.2. Sample preparation

The dry sample Quil-A® (InvivoGen, San Diego, USA) was prepared in 100 % water at 50 mg/mL for RPLC × RPLC, and at 12.5 mg/mL in 75/25 2-butanol/water (v/v) for HILIC × RPLC. All samples were stored at 4°C until use.

Quil-A sample was hydrolyzed by two different processes to mimic potential degradation. On one hand, under mild basic conditions: 20 mg of $NaHCO_3$ were added to 300 mg of Quil-A in 50 mL of MeOH/water (50/50 v/v). The mixture was heated under reflux for 1h before neutralization with HCl and filtration on a 0.22 µm membrane. The sample was then evaporated to dryness and reconstituted in 100 % water at 50 mg/mL. On the other hand, Quil-A was enzymatically hydrolyzed using Rapidase® Revelation Aroma (La Littorale, Servian, France): 20 mg of enzyme preparation was added to 10 mg of Quil-A in 5 mL of citrate-phosphate buffer (Na_2HPO_4 0.2 M and citric acid 0.1 M) at pH 5. The mixture was left for 2 h on a plate shaker incubator (50°C; 1 000 rpm). The sample was then evaporated to dryness and reconstituted in 1 mL of 75/25 2-butanol/water (v/v). Protein insoluble fraction was discarded by centrifugation.

2.3. Instrumentation and conditions

2.3.1. Instrumentation

One-dimensional LC as well as method developments and second dimension LC-UV-MS were performed on an Agilent (Santa Clara, CA, USA) UHPLC (1290 Infinity) (V_{dwell} 0.20 mL) connected via a 1:1 split to a photodiode array detector set at 205 nm (external variance 13 µl²) and to an Agilent single quadrupole with ESI source in negative mode (external variance 26 µL²). The spray chamber was set with a drying gas flow of 12 L/min (N_2), nebulizer pressure at 35 psig, drying gas temperature at 350°C and capillary voltage at 3 000 V. The fragmentor was at 70 V. The scan range was m/z 300 to 2 000. Compounds with m/z below 1 000 were considered as non-saponins Optimized ^{1}D separations for fractions collection were performed on an Acquity I-Class (Waters, Milford, MA, USA) (V_{dwell} 0.11 mL; external variance 5.6 µL²) equipped with a fraction collector (Fraction Manager). Agilent equipments were controlled by OpenLab and Waters equipment by Empower. A homemade Matlab program (S. Heinisch and F. Rouviere) was used to generate contour plots.

136

137 *2.3.2. 1D-reversed-phase LC separation*

138 The one-dimensional RLPC analysis was conducted with a Waters BEH C18 column (100 mm x 2.1
139 mm, 1.7 µm) using H_2O with 0.1 % FA (A) and ACN with 0.1 % FA (B) and a gradient elution as
140 follows : 0 min/15 % B, 117.1 min/42 % B, 120.0 min/42 % B, 120.5 min/95 % B, 121.8 min/95 %
141 B, 123.0 min/15 % B, 126.9 min/15 % B. The normalized gradient slope was 0.3 %, the flow-rate
142 0.2 mL/min, the injection volume 40 µL and the column temperature 50°C.

143 *2.3.3. LC × LC separations*

144 For ^{1}D RPLC optimization, three columns were screened: Atlantis T3 (150 mm x 3 mm, 3 µm,
145 Waters), Purospher STAR RP-18 endcapped (250 mm x 4 mm, 5 µm, Merck) and Kinetex PFP F5
146 (150 mm x 4.6 mm, 5 µm, Phenomenex). The final ^{1}D RP separation was conducted on a Kinetex
147 PFP F5 column. For ^{1}D HILIC analyses, three columns were screened: XBridge BEH Amide (150
148 mm x 4.6 mm, 5 µm, Waters), Polyhydroxyethyl A (200 mm x 4.6 mm, 5 µm, PolyLC) and TSKgel
149 amide (150 mm x 4.6 mm, 3 µm, Tosoh). The chromatographic conditions used after optimization
150 are summarized in Table 1.
151 The fractions collected from the first dimension were evaporated to dryness (30°C, 250 mbar for
152 4 h then 40°C, 150 mbar for 4 h) and recovered in water in the same volume as collected.
153 The second dimension was conducted on a Waters Acquity BEH C18 column (50 mm x 2.1 mm,
154 1.7 µm) using H_2O (A) and ACN (B) with 0.1 % FA in both phases. Gradient conditions are indicated
155 in Table 1.

156 ## 2.4. Theory supporting off-line LC × LC optimization

157 Predicted peak capacities n_{th} were calculated based on the hypothesis of Linear Solvent Strength
158 (LSS) theory [14], with the hypothesis of very large retention factors at initial composition :

$$n_{th} = 2.3\ S\ \Delta Ce \frac{1}{1+2.3sS} \frac{\sqrt{N_{col}}}{4} + 1 \quad \text{(Eq. 1)}$$

where ΔCe is defined as the range of composition at elution C_e between the first and last saponins. S is the absolute value of the slope of the relationship between the logarithm of the retention factor and the stronger solvent composition, s the normalized gradient slope (Eq. 2) and N_{col} is the column efficiency. The normalized gradient slope s is given by:

$$s = \Delta C \times \frac{t_0}{t_g} \quad \text{(Eq. 2)}$$

Where t_0 and t_g are respectively the column dead time and the gradient time and ΔC the gradient composition range.

In off-line two-dimensional separation, the overall chromatographic duration t_{total} is the sum of the ^{1}D cycle time $^1t_{total}$ and n_f times the ^{2}D cycle time $^2t_{total}$, where n_f is the number of fractions transferred from the first dimension to the second dimension (Eq. 3).

$$t_{total} = {}^1t_{total} + n_f \times {}^2t_{total} = \left({}^1t_{dwell} + {}^1t_0 + {}^1t_g\right) + n_f \times \left({}^2t_{dwell} + {}^2t_0 + {}^2t_g\right)$$

(Eq. 3)

with $^it_0$, ^it_g , $^it_{dwell}$ the column dead time, the gradient time and the system dwell time of the dimension *i*, respectively.

Considering the ^{1}D flow-rate ^{1}F, the ^{1}D gradient slope ^{1}s and the sampling rate τ as variables, for each triplet (^{1}F, ^{1}s, τ), it is hence possible to determine the ^{1}D cycle time ${}^1t_{total}$ (Eq. 4) and the number of fractions n_f (Eq. 5).

$${}^1t_{total} = \left({}^1V_{dwell} + {}^1V_0\left(1 + \frac{{}^1\Delta C}{{}^1s}\right) \times {}^1F\right) \quad \text{(Eq. 4)}$$

where $^1V_{dwell}$ and $^1V_0$ stand for the dwell volume and column void volume of the first dimension, respectively.

$$n_f = \frac{1}{6} \frac{{}^1t_g}{{}^1\sigma_{exp}} \times \tau \quad \text{(Eq. 5)}$$

where ${}^1\sigma_{exp}$ reflects the ^{1}D experimental peak standard deviation. The external dispersion is mainly due to the injection process. One of the advantages of off-line LC × LC is that the injection volume can be tuned in each dimension so that the ratio $\beta = \frac{\sigma_{col}}{\sigma_{exp}}$ (Eq. 6) remains above a value of 0.9, corresponding to a plate loss of less than 20 %.

Combining Eq. 5 with Eq. 6, the number of fractions can be expressed as followed:

$$n_f = \frac{1}{6}\frac{{}^1t_g}{{}^1\sigma_{col}} \times \frac{{}^1\sigma_{col}}{{}^1\sigma_{exp}} \times \tau = \frac{1}{6}\frac{{}^1\Delta C}{{}^1s}\frac{\sqrt{{}^1N}}{(1+{}^1ke)} \times {}^1\beta \times \tau \quad \text{(Eq. 7)}$$

By extension, the ^{2}D cycle time ${}^2t_{total}$ is deduced, hence the ^{2}D gradient time ${}^2t_g$ (Eq. 3). Working at maximal flow rate in second dimension, the associated normalized gradient slope is thus imposed.

The effective theoretical peak capacities in off-line LC × LC can hence be predicted for each triplet (^{1}F, ^{1}s, τ). The calculation has to take into account corrections for the injection band broadening effect in both dimensions as well as for undersampling and occupation space rate.

$$n_{2D,predicted} = {}^1\beta \times {}^1n_{th} \times \alpha \times {}^2\beta \times {}^2n_{th} \times f_{coverage} \text{ (Eq. 8)}$$

$^1n_{th}$ and $^2n_{th}$ are the theoretical peak capacities (Eq. 2) in ^{1}D and ^{2}D respectively. α is a correction factor, taking into account under-sampling in ^{1}D and calculated according to [15] (Eq. 9) and $f_{coverage}$ is a correction factor which takes into account the occupation space rate of the 2D separation. For prediction purposes, it was set at 0.5 for RPLC × RPLC and 0.75 for HILIC × RPLC.

$$\alpha = \frac{1}{\sqrt{1+0.21(\frac{6}{\tau})^2}} \quad \text{(Eq. 9)}$$

Experimental peak capacities for each dimension in LC × LC were measured according to:

$$n_{exp} = 1 + \frac{t_n - t_1}{w_{4\sigma}} \quad \text{(Eq. 10)}$$

t_n and t_1 are the retention times of the last and first eluted compound respectively, and $w_{4\sigma}$ is the average peak width at 4σ over about 10 peaks.

The effective experimental peak capacities in LC × LC was estimated using (Eq. 11), where $f_{coverage}$ was measured on the UV contour-plot using Convex-Hull approach [16]. The external dispersion was already taken into account in the measurement of the experimental peak widths.

$$n_{2D,effective} = f_{coverage} \times \alpha \times {}^{1}n_{exp} \times {}^{2}n_{exp} \quad \text{(Eq. 11)}$$

The dilution factor D_f is defined as the quotient of the peak volume V_{exp} over the injected volume V_{inj}. Since the column variance can be predicted according to chromatographic operating conditions, it is possible to predict the dilution factor:

$$D_f = \frac{V_{exp}}{V_{inj}} = \frac{\sigma_{exp} \times \sqrt{2\pi}}{V_{inj}} = \frac{\sigma_{col} \times \sqrt{2\pi}}{\beta\, V_{inj}} \quad \text{(Eq. 12)}$$

3. Results and discussion

The first objective of this study was to develop a reference chromatogram of Quil-A. The total chromatographic duration of 120 minutes was selected in order to generate sufficient peak capacity while limiting the instrumental occupancy rate in the QC laboratory, the latest being usually the bottleneck of routine analysis. This duration is compatible with the fact that the analysis of the stock plant material has to be performed in a regular basis, but on a limited number of samples. The duration of sample handling (evaporation) is not taken into account as it does not involve costly operations nor human supervision.

3.1. One-dimensional separation

The separations of saponins reported in the literature are usually based on reversed phase mechanism using C18 stationary phases. A BEH C18 column was selected as suggested in literature [13]. The addition of a small quantity of formic acid increased the retention without inhibiting the MS signal. Figure 1 shows the resulting optimized one-dimensional RPLC analysis of Quil-A. Saponins peaks were identified in the extract as exhibiting simultaneously UV

absorbance at 205 nm and m/z over 1 000. The first saponin eluted at 12.5 min (C_e 17 %) and the last one at 116 min (C_e 41 %). A total of 122 peaks were detected but a cluster was present between 90 and 115 min with many co-eluting peaks including the saponins QS18 and QS21 that are commonly used as quality markers. The experimental peak capacity was found to be 260 (Eq. 10) and the dilution factor 1.3 (Eq. 12). This example highlights the fact that the 1D-LC separation does not provide sufficient resolution power to analyze such a complex sample without high-resolution mass spectrometer.

3.2. Development of reversed-phase LC × reversed-phase LC and hydrophilic interaction chromatography × reversed-phase LC

The second dimension was selected to provide high efficiency, speed and compatibility with both UV and MS detection. The RPLC method described previously was selected in the second dimension, with adjustment of the column length to 50 mm to reduce the ^{2}D gradient time. The selection of the first dimension mode was based on the search for the highest degree of orthogonality relative to the RPLC used in the second dimension.

3.2.1. Screening for the first dimension

Saponins present an ionizable glucuronic acid group with a pKa of around 3.2 [17]. Using a basic mobile phase with 10 mM ammonium acetate (pH around 7) was suggested to induce a different selectivity compared to the second dimension [18]. Three different RP columns were screened using a generic gradient: Atlantis T3, Purospher STAR RP-18 endcapped and Kinetex PFP F5. The two C18 columns provided identical ΔCe of 20 %. The PFP stationary phase offered a larger elution range (44 %) with a different selectivity (Figure S2a). Saponins were found to elute between 37 % and 81 % MeOH. The injection process was optimized both in concentration and in volume (Figure S3). Thanks to the high solubility of saponins in water and to the on-column

250 focusing effect, it was possible to inject 87 µL at 50 mg/mL, while limiting the loss of plates down
251 to 20 % (β above 0.9 in Eq. 6).
252 For HILIC, a generic gradient was set up from 5 to 40 % of water + 10 mM ammonium acetate, to
253 screen the three pre-selected columns: XBridge BEH Amide, Polyhydroxyethyl A and TSKgel
254 amide. The elution window was found the same on the three columns, in the range of 15 % to 35
255 % water. The TSKgel amide column was selected for the ^{1}D HILIC separation due to its higher
256 kinetic performances (Figure S2b). Injection in HILIC is much more challenging as water is the
257 strong eluent. At the HILIC initial composition (15 % water) the solubility was below 0.5 mg/mL,
258 compromising detection. Using butanol instead of acetonitrile significantly improved Quil-A
259 solubility [19], reaching 12.5 mg/mL in a 75/25 2-BuOH/water. This high-strength solvent allowed
260 the injection of only 20 µL but was found the best compromise between solubility and injection
261 variance.
262 Repeatability is a critical parameter in quality control. For RPLC, an average standard deviation
263 of 0.009 min was calculated on the retention times of five compounds (n = 5), which is about 15
264 times lower than the average base peak width (0.14 min). For HILIC, an average standard
265 deviation of 0.002 min was calculated based on the retention times of ten compounds and an
266 average base peak width of 0.21 min. Therefore, both methods ensured that compounds
267 remained in the same fraction during repeated sampling. Worth mentioning, HILIC separation
268 was repeatable with an equilibration volume of only 5 void volumes between analyses.
269 First dimension effective peak capacity was 99 for RPLC and 39 for HILIC (Eq. 10). The lower peak
270 capacity in HILIC mode was due to the much lower elution composition range. However this low
271 separation space may be counterbalance in LC × LC by a large orthogonality related to the second
272 dimension.
273

3.2.2. Off-line LC × LC optimization

The off-line chromatographic conditions were optimized for a 120-min duration, this overall time including the first dimension duration and the cumulative second dimension durations to analyze the n_f fractions (Eq. 3). It is to be noted that the dwell volume is twice the column volume in the second dimension so a repeatable equilibrium was easily reached [20] without additional equilibrium time.

To determine kinetically optimised conditions, analyte constants were determined. The viscosity was calculated using empirical equations derived by Cabooter et al. [21] and appropriate molecular diffusion references were from Li and Carr [22, 23] in 100 % water at 20 °C. Considering an average molecular weight of 1500, the diffusion coefficient of saponins was estimated according to Wilke-Chang equation [24] at 2.76. 10^{-11} m²/s for ^{1}D RPLC (30 % - 85 % MeOH, 30 °C), 6.7.10^{-11} m²/s for ^{1}D HILIC (40 °C, 65 % - 85 % ACN) and 1.13. 10^{-12} m²/s for ^{2}D RPLC (15 % - 65 % ACN, 50 °C). The LSS constants were derived from linear retention models for RPLC. For HILIC, linear plots of log k vs. % water were considered in the very narrow composition of interest. LSS equations were applied on HILIC dimension [25], aware of the errors this model can generate on amide columns [26]. Of course, due to the diversity of the molecules engaged in this extract, all the constants are approximated as average. Nonetheless, the aim of the study was to find optimal conditions, as opposed to predict accurately the performances.

The variables for this optimization were the ^{1}D flow-rate and normalized gradient slope, as well as the sampling rate. The ^{2}D flow-rate was set to a maximal value of 1 mL/min, limited by the system pressure. The ^{2}D gradient time was dictated by the three variables through Eq. 3.

The plot of the dilution factor vs the predicted peak capacity (Figure 2) were illustrated for RPLC × RPLC (Figure 2a) and HILIC × RPLC (Figure 2b). Each open mark represents the results for a variable triplet (^{1}F, ^{1}s, τ). The mark shape relates to the gradient slope ^{1}s, its color to the sampling

298 rate τ and the various data points to the scanned ^{1}D flow-rates. Technical constraints due to the
299 collector were also taken into account, namely the maximal number of fractions (96 vials) and
300 the minimal sampling time required by the collecting needle (in our case 0.2 min). Conditions
301 that meet those criteria are reported as full marks. Improvements on collector designs may allow
302 access to better off-line 2D performances.
303
304 As a general trend, the conditions leading to the largest peak capacities in RPLC × RPLC were
305 found for a low sampling rate and a short ^{1}D separation (around 10-15 min) as shown in Figure
306 S4a. This trend is opposite to what is observed in online LC × LC. For HILIC × RPLC, the trend was
307 similar but the gradient steepness rate was found to have more influence on the peak capacity
308 than on RPLC × RPLC, as clearly visible on Figure 2b and Figure S4b, with better results at low
309 gradient slope. This is partly due to the poor kinetic performances of the selected HILIC column
310 at large flow rates.
311 The best conditions providing both a large peak capacity and a low dilution factor are indicated
312 with a black arrow. For RPLC × RPLC, working with ^{1}F = 1.4 mL/min; s = 3 %; τ = 1 generated a
313 predicted peak capacity of about 750 alongside with an acceptable dilution factor around 0.3,
314 whereas in HILIC x RPLC, the best conditions were ^{1}F = 0.6 mL/min; τ = 2; s = 4 % and gave a
315 predicted peak capacity of 600 and a dilution factor around 0.7. Table 1 summarizes these
316 kinetically optimized conditions.
317
318 *3.2.3. Quil-A reference chromatograms*
319 The experimental optimized off-line RPLC × RPLC (a) and HILIC × RPLC (b) separations are shown
320 as UV contour plots (Figure 3). Because of the necessary zoom to illustrate the presence of minor

compounds, major compounds spots may not be all distinguishable. Nonetheless, these UV contour plots are helpful in a QC environment to rapidly spot non-conform samples.

The observed occupation space was quite low in RPLC × RPLC. Diagonalized clusters were visible (Figure 3a). The ConvexHull method provided an occupation rate of 46 %. This measurement overestimated the occupation space since clusters were present. Unfortunately, the orthogonality is still difficult to apprehend when molecules are not evenly distributed over the 2D separation space [27]. An effective peak capacity 831 was calculated, which is close to the predicted value of 748. Qiao et al. [28] obtained comparable results with a peak capacity of 524 for the online RPLC × RPLC separation on triterpenoid saponins in licorice.

The HILIC × RPLC separation (Figure 3b) was characterized by a large occupation space rate (73 %), which is close to the values found in the literature. For instance, Qiu et al. calculated a value of 79 % on *Panax ginseng* HILIC x LC separation [29]. The diversity of oligosaccharides (nature, number and positions) explains the better spread along the separation space using HILIC as first dimension. The peak capacity for this 2D system was 673, close to the predicted value (594) and slightly under the RPLC × RPLC result. The HILIC dimension resulted in a dilution factor of 4. But the evaporation of ACN-rich fractions and their recovery in 100 % water allowed on-column focusing in the second dimension, resulting in an overall dilution factor of 0.5 (Eq. 12).

Table 2 summarizes the performances of the three proposed methods. At first sight, it is clear that the performances of the 1D RPLC are well below those of LC × LC separations, in terms of peak capacity, dilution factor or signal-to-noise ratio for a selected target. Regarding the two LC × LC separations, the ^{1}D RPLC exhibited a higher peak capacity than ^{1}D HILIC, but the larger orthogonality of HILIC × RPLC led to a comparable effective peak capacity (831 vs 673).

From the sensitivity point of view, HILIC mode is unfavorable. The quantity of water-soluble samples that can be injected without inducing large band broadening is limited. In this

application, only 0.25 mg of saponin sample could be injected in HILIC, which is more than 17 times lower than in RPLC. Despite on-column focusing in the second dimension, the HILIC × RPLC method provided a lower signal-to-noise ratio for QS18.

3.2.4. Application for degraded samples

While RPLC × RPLC seems to offer better performances, the potential of each LC × LC configuration was highlighted through their application on degraded samples. Quil-A was chemically and enzymatically hydrolyzed in order to obtain two different altered samples. Contour plots are displayed as Figure 4 and the pattern of the original Quil-A sample was overlapped (in black dotted lines) to easily observe the differences.

The contour plot of the RPLC × RPLC analysis (Figure 4a) showed a general decrease of the retention along the main diagonal. This was in accordance with the assumption that the hydrolysis under mild basic conditions cleaved the ester bonds and cut the fatty-acid chain-arabinose carried by the sugar at the C-28 position arabinose (noted X on Figure S1). The overall hydrophobicity decreased, hence the retention in RPLC. MS data confirmed the disappearance of the two saponins QS18 and QS21 (noted 1 and 2, respectively in Figure 4). The corresponding degradation products, namely QS18-X and QS21-X (m/z=1674 noted 1a and m/z=1512 noted 2a, respectively) were identified in the degraded sample, exhibiting as expected, a lower retention in both dimensions. On the contrary, saponins located in the upper diagonal cluster (i.e. more retained in first dimension) do not have the fatty-acid chain-arabinose and remained consistent with the original fingerprint despite hydrolysis. Discrete spots with lower masses appeared, exhibiting poor retention in both RPLC dimensions. They were identified as sugar moieties resulting from the degradation.

For the illustration of enzymatic degradation, the commercial mixture “Rapidase Revelation Aroma” was selected as it shows, among others, an apiofuranosidase activity [30]. The loss of sugars led to a shift of compounds towards lower retention in HILIC and larger retention in RPLC, as expected (Figure 4b). QS18 and QS21 were degraded as they present an apiofuranose at R_3 position (Figure S1). MS also confirmed the presence of corresponding degradation products QS18-Api (identified as 1b on Figure 4b; m/z=2019) and QS21-Api (identified as 2b on the same figure; m/z=1857). Interestingly, the selected enzyme mixture did not degrade all carbohydrate moieties as illustrated by the still large coverage of the HILIC separation space.

4. Conclusion

The objective of this study was to develop a separation method appropriate to an industrial quality control context, for the comparison of samples of Quillaja Saponaria to a reference extract. A 1D-RPLC separation was not providing enough separation power to be used without high resolution mass spectrometry detection, so LC × LC methods were investigated. RPLC × RPLC and HILIC × RPLC were optimized with a total separation time of 120 min and then compared in terms of peak capacity. The RPLC × RPLC separation was based on a pH and solvent change. However, only a limited space coverage of 46 % was reached. Nonetheless, due to high efficiency in both dimensions, an effective peak capacity of 831 was obtained. On the contrary, the HILIC × RPLC led to a high space coverage of 73 % but the effective peak capacity was only 673 because of the limited peak capacity of the HILIC first dimension. Sensitivity wise, the dilution factor Df was below 1 for both configurations, thanks to the recovery of fractions in water between the two dimensions. Despite the best performances of RPLC × RPLC for Quil-A separation, both approaches are of interest for the quality control of saponins samples, as illustrated by the analysis of degraded samples. When RPLC × RPLC is ideal to spot modifications of the carbon

skeleton, HILIC × RPLC is more suitable to detect cleavages of sugar moieties. For both, a simple UV contour plot is sufficient to rapidly spot non-conform samples in a quality control environment, while being fully compatible with MS detection if further identification is required.

Acknowledgements: The authors wish to thank Noé Biedermann for preliminary experiments and express their gratitude to Agilent Technologies for the gracious loan of the UHPLC 1290 Infinity.

Conflict of interest statement

The authors declared they have no conflict of interest.

5. References

1. David B., Wolfender J.-L., Dias D.A., The pharmaceutical industry and natural products: historical status and new trends. *Phytochem. Rev.* 2015, *14*, 299-315.
2. European Medicines Agency. Guidelines on quality of herbal medicinal products 2006; Available from: www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003370.pdf. Accessed on: February 2021.
3. Food and Agricultural Organization of the United Nation. Quillaia Extracts Type 1 and Type 2 - 65th Joint FAO/WHO Expert Committee on Food Additives 2005; Available from: http://www.fao.org/documents/card/en/c/6f3de017-d850-45d0-8c56-5ed0d062e069/. Accessed on: February 2021.
4. Makkar H.P.S., Becker K., in Oleszek W. and Marston A. (Eds.), Saponins in Food, Feedstuffs and Medicinal Plants, Springer Netherlands 2000 pp. 281-286.
5. Gautam, M., B. Patwardhan, S. Gairola, and S. Jadhav, in Patwardhan B. and Chaguturu R. (Eds.),Innovative Approaches in Drug Discovery, Academic Press, Boston 2017, pp. 315-341.

6. Fleck D.J., Betti H.A., Da Silva P. F., Troian A.E., Olivaro C., Ferreira F., Verza G.S., Saponins from Quillaja saponaria and Quillaja brasiliensis: Particular Chemical Characteristics and Biological Activities. *Molecules* 2019, *24,* 171 - 196.

7. Nord, L.I. , Kenne L., Separation and structural analysis of saponins in a bark extract from Quillaja saponaria Molina. *Carbohydr. Res.* 1999, *320,* 70-81.

8. Wang, Y., Lu X. , Xu G., Development of a comprehensive two-dimensional hydrophilic interaction chromatography/quadrupole time-of-flight mass spectrometry system and its application in separation and identification of saponins from Quillaja saponaria. *J. Chromatogr. A* 2008, *1181*, 51-59.

9. Thalhamer B., Himmelsbach M., Characterization of quillaja bark extracts and evaluation of their purity using liquid chromatography–high resolution mass spectrometry. *Phytochem. Lett.* 2014, *8*, 97-100.

10. Kirk D.D., Rempel R., Pinkhasov J., Walmsley A.M., Application of Quillaja saponaria extracts as oral adjuvants for plant-made vaccines. *Expert Opin. Biol. Ther.* 2004, *4,* 947-958.

11. Kalili K.M., de Villiers A., Systematic optimisation and evaluation of on-line, off-line and stop-flow comprehensive hydrophilic interaction chromatography×reversed phase liquid chromatographic analysis of procyanidins, Part I: Theoretical considerations. *J. Chromatogr. A* 2013, *1289*, 58-68.

12. Bankefors J., Nord L.I., Kenne L., Multidimensional profiling of components in complex mixtures of natural products for metabolic analysis, proof of concept: Application to Quillaja saponins. *J. Chromatogr. B* 2010, *878,* 471-476.

13. Xing Q.Q., Liang T., Shen G.B., Wang X.L., Jin Y., Liang X.M., Comprehensive HILIC x RPLC with mass spectrometry detection for the analysis of saponins in Panax notoginseng. *Analyst* 2012, *137*, 2239-2249.

14. Snyder, L.R., Dolan J. W., High-Performance Gradient Elution: The Practical Application of the Linear-Solvent-Strength Model, First Edition, Wiley. New York, 2006.

15. Davis J.M., Stoll D.R., Carr P.W., Effect of First-Dimension Undersampling on Effective Peak Capacity in Comprehensive Two-Dimensional Separations. *Anal. Chem.* 2008, *80,* 461-473.
16. Nowik W., Héron S., Bonose M., Nowik M., Tchapla A., Assessment of Two-Dimensional Separative Systems Using Nearest-Neighbor Distances Approach. Part 1: Orthogonality Aspects. *Anal. Chem.* 2013, *85,* 9449-9458.
17. Mitra S., Dungan S.R., Micellar Properties of Quillaja Saponin. 1. Effects of Temperature, Salt, and pH on Solution Properties. *J. Agric. Food. Chem.* 1997, *45,* 1587-1595.
18. Stoll D.R., O'Neill K., Harmes D.C., Effects of pH mismatch between the two dimensions of reversed-phase×reversed-phase two-dimensional separations on second dimension separation quality for ionogenic compounds—I. Carboxylic acids. *J. Chromatogr. A* 2015, *1383*, 25-34.
19. Güçlü-Üstündağ, Ö., Mazza G., Saponins: Properties, Applications and Processing. *Crit. Rev. Food Sci. Nutr.* 2007, *47*, 231-258.
20. Schellinger A.P., Stoll D.R., Carr P.W., High-speed gradient elution reversed-phase liquid chromatography of bases in buffered eluents: Part I. Retention repeatability and column re-equilibration. *J. Chromatogr. A* 2008, *1192,* 41-53.
21. Cabooter D., Heinisch S., Rocca J.L., Clicq D., Desmet G., Use of the kinetic plot method to analyze commercial high-temperature liquid chromatography systems I: Intrinsic performance comparison. J*. Chromatogr. A* 2007, *1143,* 121-133.
22. Li J., Carr P.W., Accuracy of Empirical Correlations for Estimating Diffusion Coefficients in Aqueous Organic Mixtures. *Anal. Chem.* 1997, *69,* 2530-2536.
23. Li J., Carr P.W., Estimating Diffusion Coefficients for Alkylbenzenes and Alkylphenones in Aqueous Mixtures with Acetonitrile and Methanol. *Anal. Chem.* 1997, *69,* 2550-2553.
24. Wilke C.R., Chang P., Correlation of diffusion coefficients in dilute solutions. *AlChE J.* 1955, *1,* 264-270.
25. D'Attoma, A., C. Grivel, and S. Heinisch, On-line comprehensive two-dimensional separations of charged compounds using reversed-phase high performance liquid chromatography and

hydrophilic interaction chromatography. Part I: Orthogonality and practical peak capacity considerations. *J. Chromatogr. A* 2012, *1262,* 148-159.

26. Pirok B.W.J., Molenaar S.R.A., van Outersterp R.E., Schoenmakers P.J., Applicability of retention modelling in hydrophilic-interaction liquid chromatography for algorithmic optimization programs with gradient-scanning techniques. *J. Chromatogr. A* 2017, *1530,* 104-111.

27. Gilar M., Olivova P., Daly A.E., Gebler J.C., Orthogonality of Separation in Two-Dimensional Liquid Chromatography. *Anal. Chem.* 2005,*77,* 6426-6434.

28. Qiao X., Song W., Ji S., Wang Q., Guo D., Ye M., Separation and characterization of phenolic compounds and triterpenoid saponins in licorice (Glycyrrhiza uralensis) using mobile phase-dependent reversed-phase×reversed-phase comprehensive two-dimensional liquid chromatography coupled with mass spectrometry. *J. Chromatogr. A* 2015, *1402*, 36-45.

29. Qiu S., Yang W.-z., Shi X.-j., Yao C.-l., Yang M., Liu X., Jiang B.-h., Wu W.-y., Guo D.-a., A green protocol for efficient discovery of novel natural compounds: Characterization of new ginsenosides from the stems and leaves of Panax ginseng as a case study. *Anal. Chim. Acta* 2015, *893*, 65-76.

30. Kis P., Potocká E., Mastihuba V., Mastihubová M., Efficient chemoenzymatic synthesis of 4-nitrophenyl β-d-apiofuranoside and its use in screening of β-d-apiofuranosidases. *Carbohydr. Res.* 2016, *430*, 48-53.

Figures

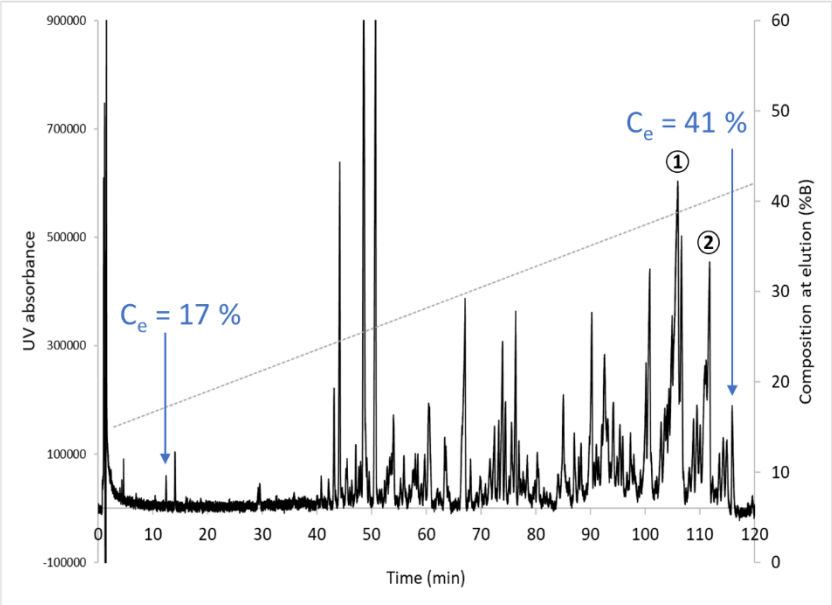


Figure 1. UV chromatogram of the 120-min RPLC separation of Quil-A. BEH C18 column (100 mm x 2.1 mm, 1.7 µm). Mobile phase: H_2O (A) and ACN (B) with 0.1 % FA in both phases. Gradient 15 % to 42 % (B). Flow-rate 200 µL/min. Injected volume 40 µL (50 mg/mL in water). UV detection: 205 nm. Secondary axis represents the composition at elution (% B) along the analysis. Arrows point the first and the last saponin peak with the corresponding composition at elution and numbers (1) and (2) indicate saponins QS18 and QS21 respectively.

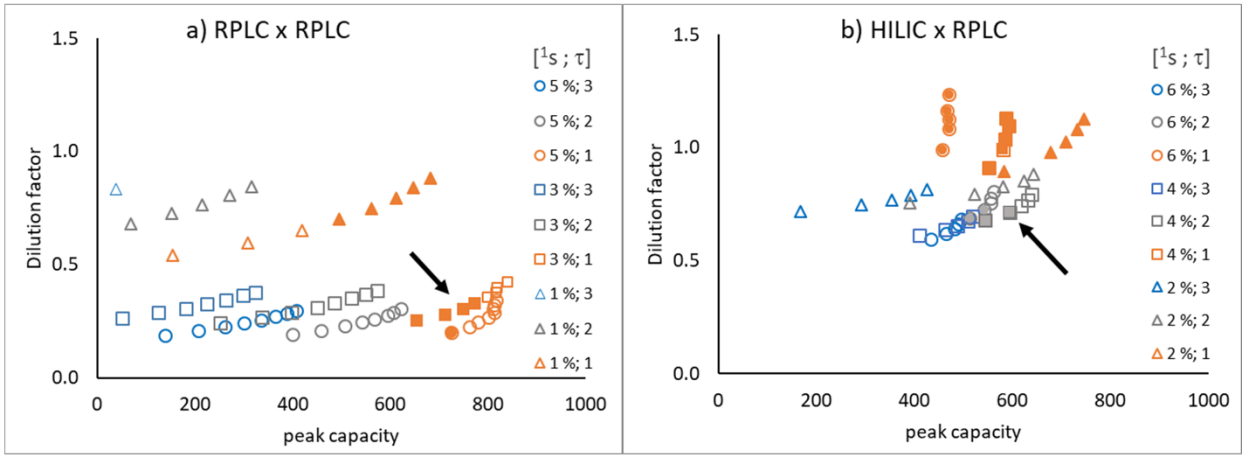


Figure 2. Plot of the predicted dilution factor vs predicted peak capacity for a chromatographic duration of 120 min. a) RPLC × RPLC and b) HILIC × RPLC. Calculations were performed using three selected variables: ^{1}F [1- 3 mL/min] for RPLC × RPLC and [0.2 – 1.2 mL/min] for HILIC × RPLC ; ^{1}s [1 %; 3 %; 5 %] for RPLC × RPLC and [2 %; 4 %; 6 %] for HILIC × RPLC; sampling rate τ [1; 2; 3] for both methods. Open marks represent theoretical results and full marks represents results that are compatible with technical constraints: number of transferred fractions below 96 and a sampling time over 0.2 min. The arrow indicates the selected operating conditions.

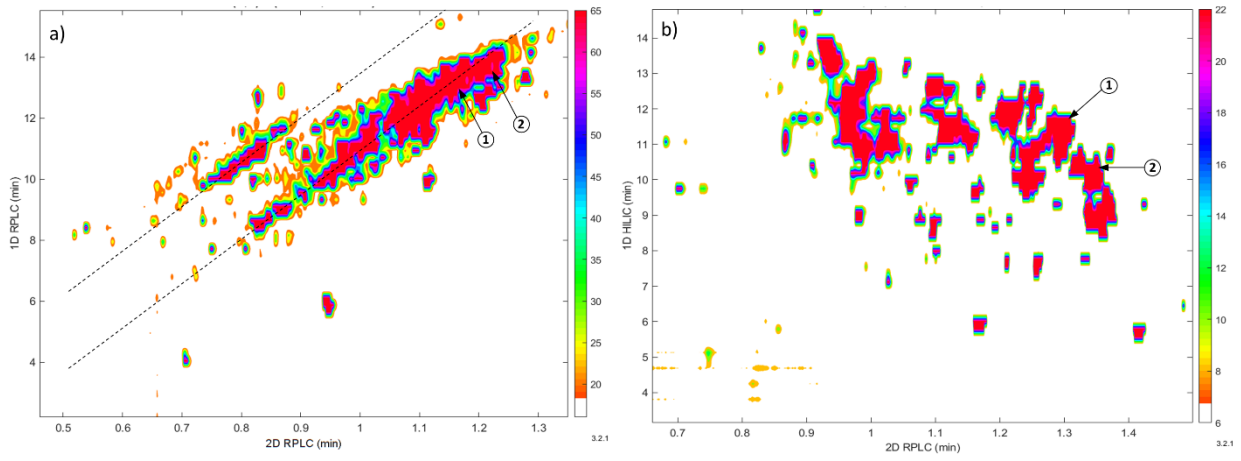


Figure 3. UV contour plots of off-line separation of Quil-A using a) RPLC × RPLC and b) HILIC × RPLC. Optimized conditions displayed in Table 1. Saponins QS18 (1) and QS21 (2) are indicated.

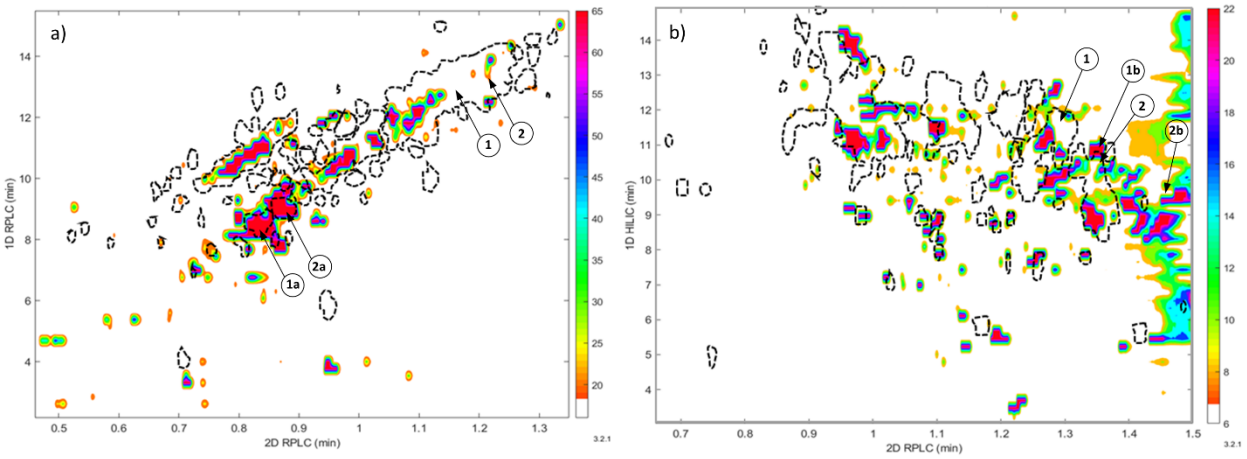


Figure 4. 2D-contour plots of off-line separations of degraded Quil-A. a) RPLC × RPLC of Quil-A hydrolyzed; b) HILIC × RPLC of enzymatically hydrolyzed Quil-A. Original RPLC × RPLC and HILIC × RPLC contour plot patterns of Quil-A overlapped (black dotted line). Similar conditions as displayed in Table 1. UV signal at 205 nm. Numbers indicate saponins QS18 (1) and QS21 (2) and corresponding degraded species a) QS18-X (1a) and QS21-X (2a) or b) QS18-Api (1b) and QS21-Api (2b).

Supplementary material

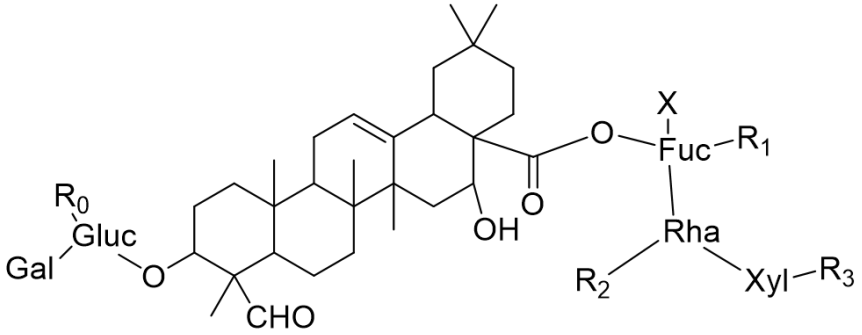
Figure S1. Structures of saponins from *Quillaja Saponaria*

Figure S2. UV chromatogram of ^{1}D separation of QuilA, a) RPLC Kinetex PFP F5 column.

Figure S3. Injection study in RPLC.

Figure S4. Plot of the predicted 2D peak capacity vs. the gradient time of the first dimension

Generic structure



Some known saponins

Name	R_0	R_1	R_2	R_3	X	Mw (g/mol)
QS21	Xyl	H	H	Api/Xyl	FA-Araf	1990.1
QS18	Xyl	H	Glc	Api	FA-Araf	2152.3

Structure of FA-Araf

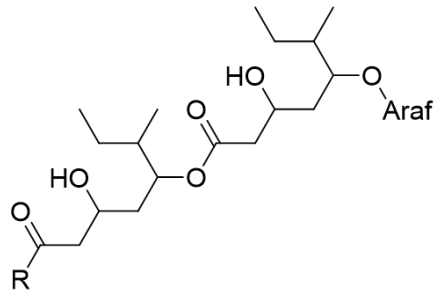


Figure S1. Structures of saponins from *Quillaja Saponaria*

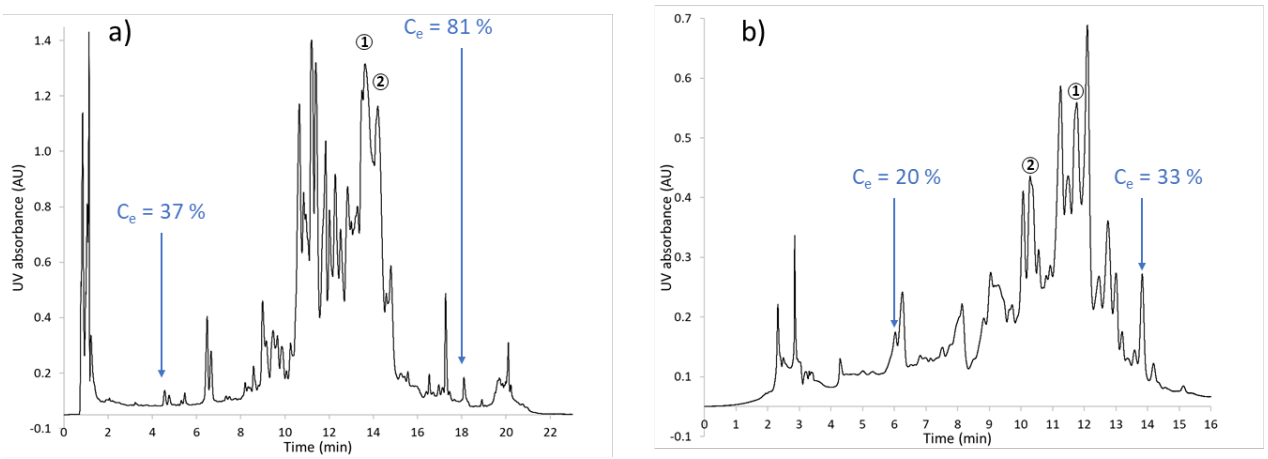


Figure S2. UV chromatogram of [1]D separation of QuilA, a) RPLC Kinetex PFP F5 column. Mobile phase: 10 mM AA in water (A) and in MeOH (B). Organic gradient started from 30 to 85 % (B) in 17.5 min at 1.4 mL/min. Injected volume of 50 µL (50 mg/mL in water); b) HILIC TSKgel amide column. Mobile phase: ACN (A) and 10 mM ammonium acetate in water (B). Aqueous gradient started from 15 % to 35 % (B) in 12.5 min at 0.6 mL/min. Injected volume of 20 µL (12.5 mg/mL in 2-BuOH/water75/25 v/v). UV detection at 205 nm. Arrows point the first and the last saponin peak with the corresponding composition at elution and numbers (1) and (2) indicate major saponins QS18 and QS21 respectively.

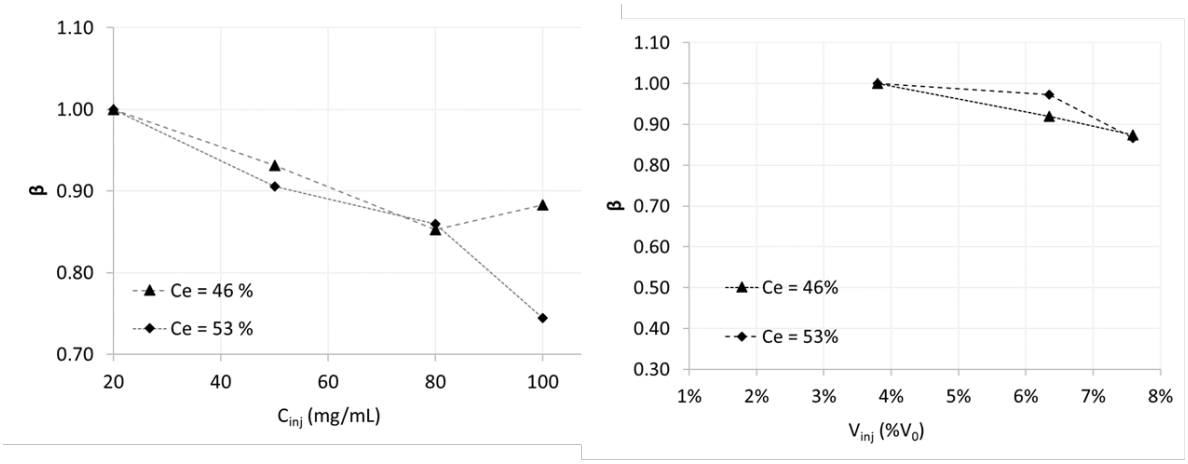


Figure S3. Injection study in RPLC. Plot of the fraction of remaining peak capacity β for two saponins, as a function of a) the concentration of Quil-A sample in water (Vinj 10 µL) and of b) the injection volumerelated to the column volume (50 mg/mL in water). Kinetex PFP F5 column. Mobile phase: 10 mM AA in water (A) and in MeOH (B). Gradient 30 % to 85 % (B) in 17.5 min. Flow-rate 1.4 mL/min.

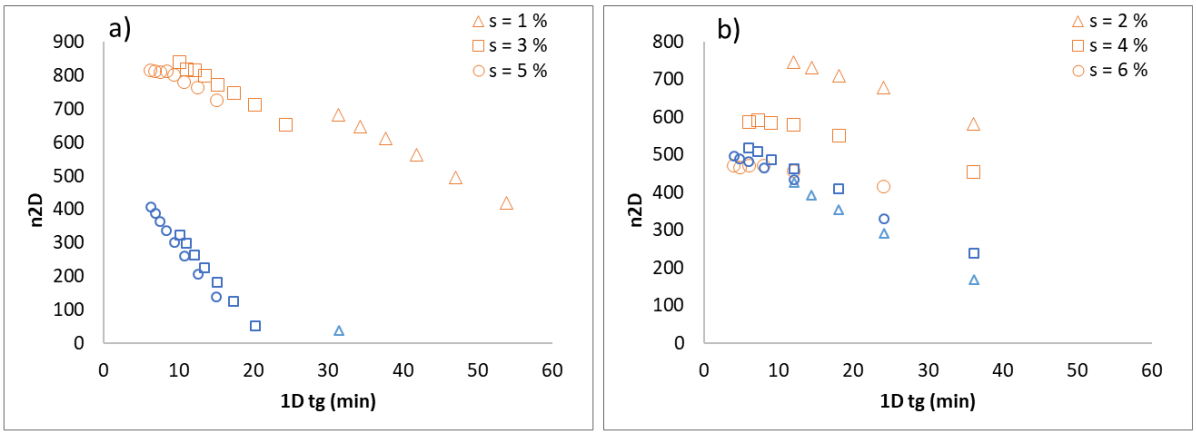


Figure S4. Plot of the predicted 2D peak capacity vs. the gradient time of the first dimension, for a chromatographic duration under 120 min. a) RPLC x RPLC and b) HILIC x RPLC. Calculations were performed using three variables : ^{1}F [1- 3 mL/min] for RPLC x RPLC and [0.2 – 1.2 mL/min] for HILIC x RPLC; ^{1}s (normalized gradient slope) [1 %; 3 %; 5 %] (marked as triangle, square and circle, respectively) for RPLC x RPLC and [2 %; 4 %; 6 %] for HILIC x RPLC; sampling rate τ of 1 (orange) and 3 (blue) for both methods.

563
564
565