



**HAL**  
open science

# In silico screening of comprehensive two-dimensional centrifugal partition chromatography x liquid chromatography for multiple compounds isolation

Léa Marlot, Magali Batteau, Dalene de Beer, Karine Faure

► **To cite this version:**

Léa Marlot, Magali Batteau, Dalene de Beer, Karine Faure. In silico screening of comprehensive two-dimensional centrifugal partition chromatography x liquid chromatography for multiple compounds isolation. *Analytical Chemistry*, 2018, 90 (24), pp.14279-14286. 10.1021/acs.analchem.8b03440 . hal-01937324

**HAL Id: hal-01937324**

**<https://hal.science/hal-01937324>**

Submitted on 23 Jul 2020

**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.

# In silico screening of comprehensive two-dimensional centrifugal partition chromatography x liquid chromatography for multiple compounds isolation

Léa Marlot<sup>1</sup>, Magali Batteau<sup>1</sup>, Dalene De Beer<sup>2,3</sup>, Karine Faure<sup>1,\*</sup>

<sup>1</sup> Université de Lyon, CNRS, Université Claude Bernard Lyon 1, ENS de Lyon, Institut des Sciences Analytiques, UMR 5280, 5 rue de la Doua, F-69100 VILLEURBANNE, France

<sup>2</sup> Plant Bioactives Group, Post-Harvest & Agro-Processing Technologies, Agricultural Research Council Infruitec-Nietvoorbij, Private Bag X5026, 7599 Stellenbosch, South Africa

<sup>3</sup> Department of Food Science, University of Stellenbosch, Private Bag X1, 7602 Matieland (Stellenbosch), South Africa

\*corresponding author karine.faure@isa-lyon.fr

## Abstract

The isolation of unknown compounds for structural identification and the collection of target molecules to generate unavailable standards remain a challenge when dealing with complex sample. While tedious multistep purification is commonly used, it is not appropriate for a limited amount of sample or when a full recovery of expensive molecules is required. Two-dimensional preparative chromatography in a comprehensive mode provides an effective mean to collect a large number of molecules in such a case. However, there is currently a lack of metrics to estimate preparative performances with a minimal number of experiments. An in silico comparative study of various pairs of chromatographic systems is proposed, focusing on the occupation rate and the homogeneity of peak spreading in the 2D separation space. Off-line combination of centrifugal partition chromatography (CPC) with liquid chromatography (LC) exhibits numerous advantages for 2D preparative separation. Our in silico approach was illustrated through the isolation of eight bioactives compounds with very similar structures from *Cyclopia genistoides* plant by CPCxLC. The column screening was performed considering predictive 2D plots in the light of the preparative performance descriptors and compared to real 2D preparative separations.

Since the early 2000s, two-dimensional preparative separations have emerged especially in the natural products field in order to isolate targeted compounds from highly complex samples. Simultaneous isolation of a large number of molecules from high value samples is able to generate small amounts of high purity material for NMR analysis, biological evaluations, mass quantification or as reference substances. Orthogonal separation mechanisms offer a large separation space, improving the ability to separate impurities from the compounds of interest. Moreover, as the distance between peaks is related to both purity and recovery in preparative chromatography, it was demonstrated that the amount of recovered molecules was significantly higher in 2D separations compared to a succession of 1D separations<sup>1</sup>. The first 2D preparative separation was realized in 1976 in order to recover the maximum amount of two compounds in a simple mixture<sup>2</sup>. This approach led to the development of 2D preparative separations with both the use of liquid chromatography (LC) and liquid-liquid chromatography (LLC). As described in a recent review from our group<sup>3</sup>, the first 2D preparative separations involving LLC were dedicated to the isolation of natural products using countercurrent chromatography (CCC) technology in heart-cut mode. The preparative 2D-LC separations started to emerge later, in the years 2010. The first comprehensive 2D-LC was implemented in off-line mode in 2012 for the isolation of thirteen compounds from a plant extract<sup>4</sup>. Since this study, off-line and on-line preparative 2D-LC separations were performed in comprehensive mode<sup>5-9</sup>, to obtain the required high purity on multiple molecule isolation.

To ensure an ordered spreading of peaks in the 2D space and hence prevent the risk of overlapping, it is of utmost importance to screen different 2D configurations using adequate selection criteria. While reviewing the two-dimensional preparative chromatography methods involving either liquid-liquid techniques (centrifugal partition chromatography (CPC) or CCC)<sup>3</sup> or LC<sup>4</sup>, it appears that most developments rely on one-dimensional selectivity improvement of a critical pair using visual estimation. This is especially obvious when the 2D method is applied in a heart-cut mode on specific areas of the chromatograms. For the isolation of multiple compounds from complex sample for which a 2D separation is fully justified, a comprehensive mode is recommended, but to the best of our knowledge, only three papers describe a quantitative assessment of the 2D space quality for preparative purposes. In 2016, the orthogonality of the 2D-LCxLC system has been evaluated with the bin box approach<sup>10</sup>. In 2017, the convex hull approach and the conditional entropy have been used for the evaluation of a 2D-CCCxLC system<sup>11</sup>. In 2018, a 2D-CPCxLC system has been evaluated through the measurement of the degree of orthogonality<sup>12</sup>. These criteria are derived from the evaluation of analytical 2D separations<sup>13</sup> and used to estimate the quality of the final 2D separation once it has experimentally been performed. None of the studies provides insights into the basis of column selection for each dimension, which is a critical stage in the optimization of 2D preparative separations.

The development strategy for 2D preparative separation currently relies on the optimization of separation selectivity in each dimension. Since combined selectivities can sometimes be correlated, this strategy does not guarantee an optimized 2D separation. In the present paper, a simultaneous screening of various chromatographic columns is demonstrated with a minimum number of experiments using *in silico* 2D-plots. A number of criteria specific to preparative purposes are discussed such as the occupation of the separation space, the distance between peak apexes and the peak spreading homogeneity.

Several of the 23 species in the genus *Cyclopia* are used for production of honeybush herbal tea, a South African herbal tea prized for its sweet-associated aroma<sup>14</sup>. *Cyclopia genistoides* contains high levels of highly polar molecules such as the xanthenes mangiferin and its regioisomer, isomangiferin, as well as a number of benzophenones and flavanones<sup>15</sup>, which are extracted during tea brewing. These compounds exhibit a bioactivity related to alleviation of the metabolic syndrome which position the plant as valuable source material<sup>16</sup>. Several qualitative studies have been performed for the comprehensive

characterization<sup>15-17</sup> of this sample, but the simultaneous isolation of the components that can be found in the hot water extract would facilitate the investigation on the plant bioactivity. The proposed screening methodology for 2D preparative separation was therefore illustrated by the simultaneous isolation of the eight major compounds present in *C. genistoides* water extract by off-line comprehensive 2D-CPCxLC. Since these compounds have very similar chemical structures (Fig. 1), their simultaneous isolation with a high degree of purity is not achievable in a single dimension approach.

Due to the liquid nature of the stationary phase in CPC, this technique provides many advantages such as high selectivity thanks to multiple choice of solvent systems, high loading capacity and total recovery of the sample. Therefore, it is of interest to combine the advantages of CPC with the high efficiency of HPLC for the isolation of the eight targeted compounds. The criteria that we propose to screen various pairs of chromatographic systems are designed for the combination of a first-dimension CPC separation with a variable amount of stationary phase under isocratic elution mode, and a second-dimension LC separation with a gradient elution mode. We suggest these criteria may also be applied to any kind of chromatographic column and elution mode.

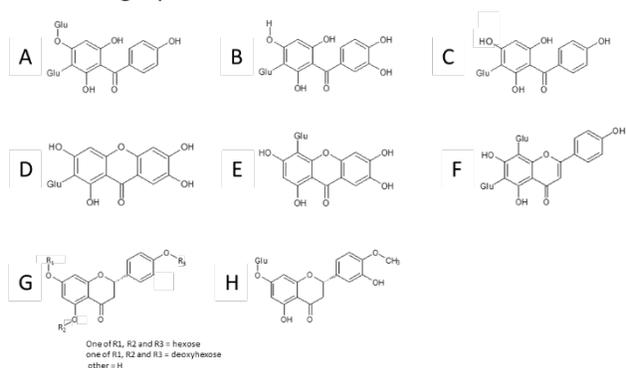


Figure 1. Chemical structures of the eight targeted compounds present in *Cyclopia genistoides* plant extract. A: 3- $\beta$ -D-Glucopyranosyl-4- $\beta$ -D-glucopyranosyloxyiriflophenone, B: 3- $\beta$ -D-Glucopyranosylmaclurin, C: 3- $\beta$ -D-Glucopyranosyliriflophenone, D: Mangiferin, E: Isomangiferin, F: Vicenin-2, G: Naringenin-*O*-hexose-*O*-deoxyhexose isomer B<sup>15</sup>, H: Hesperidin.

## DESCRIPTORS

In analytical separations, the quality of 2D separations are usually evaluated through two notions: the practical sample peak capacity and the degree of orthogonality<sup>18</sup>. Different methods based on the linear regression of retention data<sup>19</sup> and chemometric tools<sup>20</sup> have been performed for a qualitative assessment while methods based on geometrical approach<sup>18,21-23</sup> allow a quantitative measurement. While the aim in analytical separation is to separate all compounds in a maximized separation space in a minimum separation time, the aim in preparative separation is different. Thus, the criteria for evaluation of the 2D systems have to be different. We suggest that for the isolation of multiple compounds, the most important criterion is a homogeneous spreading of the peaks in a large occupation space in order to provide the largest distance between peaks in the 2D separation space for both compound purity (i.e. separation from impurities or overlapping) and recovery requirements (i.e. for loading capacity). Clustering is the most critical situation for preparative chromatography development, while time constraint is not so crucial.

Thereby, in this work, the 2D systems are evaluated through three criteria. The first criterion is the occupation rate of the peaks in the separation space denoted  $\gamma$ . It is defined as the ratio of the used retention space measured by the convex hull approach<sup>24</sup> using the Delaunay triangulation method on peak apexes, over the usable retention space, considered as the space between the void volume of the column and the retention volume of the last eluting compound in each dimension.

The second criterion is the homogeneity of the peak spreading in the used separation space. This criterion is evaluated through the ratio of the harmonic mean  $\bar{H}$  with the arithmetic mean  $\bar{A}$  of the nearest-neighbor distances<sup>25</sup>. The nearest-neighbor distance  $d_i$  is the shortest of all distances connecting a peak apex with every other peak apex. Thus, for  $n$  peaks, it exists  $n-1$  distances  $d_i$ . The harmonic mean is a descriptor sensitive to the shortest distances between the nearest neighbors while the arithmetic mean is useful for the assessment of the extension of peak spreading in the total space (full expression available in supporting information). Thereby, the peak spreading is homogeneous when the harmonic mean is equal to the arithmetic mean, i.e. when the ratio  $\bar{H}/\bar{A}$ , called system homogeneity<sup>26</sup>, is equal to 1. The third criterion is the loading capacity of the 2D system. For that, we suggest the use of the minimal nearest-neighbor distance, denoted  $d_{i_{min}}$ . The higher the minimal distance, the higher the loading can be. This minimal distance is of course relative to specific instrumentation and operating conditions.

## EXPERIMENTAL SECTION

### Reagents and materials

All solvents were of analytical grade. 1-Butanol, formic acid, sucrose and ammonium sulfate were purchased from Acros Organics (Fisher Scientific, Illkirch, France). Ethanol and acetonitrile were from Sigma-Aldrich (Saint-Quentin-Fallavier, France). *Cyclopia genistoides* plant material (leaves and fine stems) were harvested from a commercial plantation (Gouritzmond, Western Cape, South Africa). The plant material was dried and a hot water extract prepared as described by Beelders et al.<sup>15</sup>.

### Instrumentation

The HPLC system was an Alliance 2690 system from Waters (Saint-Quentin-en-Yvelines, France) using a 5.2 ml sample loop and a photodiode array detector Waters 996 set up at 288 nm. Data acquisition was performed by EmPower 3 software from Waters.

The CPC instrument was set with a SpotPrep II system from Gilson (Saint-Avé, France) connected to a FCPC-A frame equipped with a 33.1 ml rotor from Kromaton Rousselet-Robatel (Annonay, France) and a SPD6A detector (Shimadzu, Noisel, France) set up at 288 nm. Either a 1.1 ml or 2 ml sample loop was used. Data were collected with Azur software (Datalys, Le Touvet, France).

Predictive and experimental 2D plots were processed by home-made scripts on Matlab (Mathworks, Massachusetts, USA).

### Partition coefficient measurement

A suitable amount of plant extract was added to a test tube and 2 ml of each phase of the equilibrated solvent system was added at 30°C. After vigorous agitation and decantation, the lower phases were taken into vials for dilution and analysis by HPLC-UV (LC1 method, see below). The partition coefficient  $K_D$  value was calculated by comparing the peak area of the target compound in the plant extract and the one in the lower phase of the evaluated solvent system.

### CPC separation conditions

Four CPC methods were fully evaluated in this study and are noted as follows: CPC1 (1-butanol/water 50/50 (v/v) descending mode; CPC2 1-butanol/water 50/50 (v/v) ascending mode; CPC3 ethanol/ammonium sulfate/water 22/21/57 (w/w) descending mode; CPC4 acetonitrile/sucrose/water 50/13.3/36.7 (w/w) ascending mode. As a reminder, descending mode means that the upper phase is the stationary phase and inversely for ascending mode. Operating conditions are detailed on supporting information file (Table S1). A volume of 350  $\mu$ l plant extract (5 mg/ml in water) was injected. The fractions were evaporated to dryness then diluted in water before HPLC analysis.

### HPLC separation conditions

The conditions for HPLC methods were as follows: mobile phase water (A)/ acetonitrile (B) + 0.1% formic acid (A and B); elution composition range 3% to 25% B; normalized gradient slope 1%. Column temperature 30°C. The injection volumes were 20 µl for the plant extract analysis (3mg/ml in water) and from 620 to 1040 µl for analysis of CPC fractions. Two reversed-phase columns were evaluated as follows: LC1 method: XSelect CSH PhenylHexyl column, 150 mm x 4.6 mm i.d., 5 µm, from Waters; flow rate 2 mL/min, gradient time 17.6 min, total time 25 min. LC2 method: Gemini-NX C18 column, 150 mm x 4.6 mm i.d., 3 µm, from Phenomenex; flow rate 1.5 mL/min, gradient time 24.2 min, total time 34 min.

## RESULTS AND DISCUSSION

### In silico evaluation of 2D systems using predictive contour plots

Two HPLC columns were screened, namely an XSelect CSH Phenylhexyl (LC1), and a Gemini-NX C18 (LC2) both with a gradient of acetonitrile and water as mobile phase. In order to achieve high selectivity and thus high distances between peaks, HPLC methods were implemented with a minimum gradient normalized slope of 1%. The two chromatograms of the HPLC analysis of the crude extract are shown in Figure S1 (supporting information). There is no significant difference of selectivity between the two HPLC methods except for the positional isomers D and E where a higher selectivity is observed for the LC1 method.

Twenty potential CPC columns were screened by evaluating the partition coefficient, denoted  $K_D$ , of the targeted compounds in ten solvent systems. Four solvent systems leading to high selectivity in a  $K_D$ -value range from 0 to 10 were selected for discussion. The corresponding partition coefficients are available in Table S.2 (supporting information). As the partition coefficients are determined by HPLC analysis, coeluted peaks in HPLC could lead to incorrect  $K_D$ -values, hence it is of importance to base the screening on a large number of peaks. A full HPLC resolution was not necessary to determine  $K_D$ -values of selected compounds, as peak height can be used to provide a fair estimation.

In order to evaluate the eight selected 2D-CPCxLC systems (noted #1 to #4 with LC1 column and #5 to #8 with LC2 column), the in silico 2D-plots of the retention volumes in the two dimensions were generated. The relation between the retention volume ( $V_r$ ) of the compounds in <sup>1</sup>D-CPC and their  $K_D$ -value is as follows:  $V_r = [(K_D - 1) \cdot Sf + 1] \cdot V_c$ , with  $Sf$  the retention rate of the stationary phase in the column and  $V_c$  the column volume. These data depend on the retention rate of the CPC stationary phase which was initially set at 0.5 in the assessment. Estimation of the peak broadening was also plotted. In <sup>1</sup>D-CPC, a column efficiency of 300 was initially considered. Due to the isocratic elution mode, peak broadening was simulated according to retention extent. The theoretical efficiency of the HPLC columns was considered to estimate the peak width in the <sup>2</sup>D-HPLC. As no significant difference was observed between LC1 and LC2 methods, only the 2D-plots using LC1 method as second dimension are discussed further (Fig. 2). It is important to notice that only two HPLC injections are required to realize a full 2D-plot, one with the plant sample and one with the lower phase of the solvent system containing the sample. Moreover, peak identification is achieved only by their HPLC retention time and no peak tracking is required.

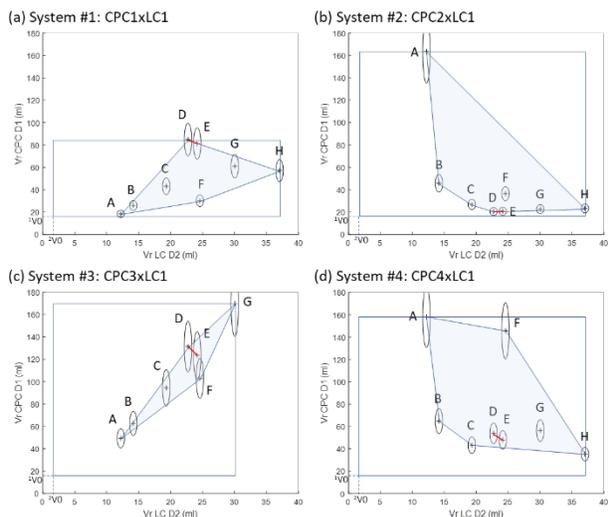


Figure 2. Predictive 2D-plots of (a) CPC1xLC1 separation, (b) CPC2xLC1 separation, (c) CPC3xLC1 separation and (d) CPC4xLC1 separation. The blue shape represents the occupation space in the usable separation space and the red line represents the minimal distance between nearest-neighbor peaks (peaks D and E) (see online version for color figure).

A simple visualization of the four 2D-plots illustrates that the <sup>1</sup>D-CPC can easily generate different selectivities depending on the nature of the solvent system. With the systems #1, #2 and #4, the retention mechanisms in the two dimensions are not correlated leading to a potential separation of the eight targeted compounds (Fig. 2a, 2b, 2d). However, this is not the case of system #3 where the compounds are spread around a regression line in the 2D separation space leading to a coelution of compound E with compound F on the corresponding 2D-plot (Fig. 2c). It can be noted that compound H, with  $K_D$ -value over 10, is not present in the 2D separation space. Thereby, a visual estimation of these predictive 2D-plots can very simply help the developer to exclude sets of conditions that are likely to fail (here system #3) with a minimal number of experiments. In order to refine the predictive 2D-plots, a prior knowledge of the plate number in CPC would be helpful. This data is currently very dependent on the solvent system, the operating conditions and the equipment supplier, for which preliminary tests could be realized. To illustrate a more generic approach, the *in silico* evaluation presented hereafter is only based on the thermodynamic data.

To discriminate pairs of chromatographic columns, performance descriptors have to be calculated. Purity is the number one concern for preparative development, especially when multiple compounds are targeted in a complex matrix where impurities may be numerous. Since impurities are not considered in the representation, it is important for an efficient preparative 2D system to exhibit a homogeneous spreading of the peaks in a large occupied separation space. This provides a better chance to separate a large number of targeted compounds from impurities. Thus, two criteria related to purity were evaluated: the occupation rate of the peaks in the usable separation space, denoted  $\gamma$ , and the system homogeneity, denoted  $\bar{H}/\bar{A}$ .

The occupation space of the peaks is represented in blue on the 2D-plots (Fig. 2). A difference in occupancy can be clearly noted between systems #1, #2, #4 and the system #3 where the blue shape is much larger for the former than for the latter. This lower occupied space due to correlated separation mechanisms between the dimensions leads to the coelution of compounds E and F. While 5 out of 8 targets appear to be separated in system #3 during the elution, the impurities have a high chance to coelute with other target compounds.

The homogeneity of the peak spreading is also of importance as clustering of peaks in the separation space may result in peak overlapping and purity loss as soon as sample loading is increased.

Hence purity capacity can be understood as a combination of occupation rate and system homogeneity. Chromatographic sets can be compared according to these two criteria (Fig. 3a). For a given <sup>1</sup>D-CPC, no significant difference can be seen when using LC1 or LC2. The CPC1 and CPC4 systems offer high occupation rate and high system homogeneity, expecting to lead to a high purity capacity. The number of pure compounds may be limited when using CPC2 and CPC3 systems, due to the low system homogeneity for CPC2 systems or the low occupation rate for CPC3 systems.

While the number of peaks to be resolved may be the only goal for most preparative challenges, the improvement of the sample load can be a requirement, especially for routine separations. In this case, the resolution is the key parameter and the separation time may be longer than analytical separations where a large number of samples must be analyzed in a given time. Thus, we propose to evaluate the sample loading capacity of the 2D system by measuring the minimal distance  $di_{min}$  between peaks (red line Fig. 2). However, lengthening distances means increasing time and solvent consumption and therefore the running cost, especially if the separation is supposed to be scaled up to some extent. Thus, for each 2D system, the minimal distance has to be considered regarding to the overall separation time, which depends on the separation time in each dimension and on the fractions number. We acknowledge that the relationship between  $di_{min}$  and load is much more complex than a simple proportionality, since (i) it may depend on sample solubility in CPC mobile phase and (ii) the distances are not equivalent in isocratic and gradient methods in terms of loading ability. Nevertheless, comparing distances provides a first estimation of loading that is actually lacking for column selection in 2D preparative method development. The purity capacity (considered as the occupation rate times the homogeneity) can be plotted with regard to the minimal distance  $di_{min}$  reflecting the loading capacity (Fig. 3b). Not surprisingly, CPC1 and CPC4 systems display high purity capacity and consequently interesting loading capacity. This graph enables discrimination between two combinations that exhibit different features. For example, system #2 offers a large occupation rate hence can handle impurities more easily, while the clustering makes sample loading challenging. On the opposite, system #3 can manage a higher load due to a more homogeneous spreading, but in a smaller separation space. This plot aids decision making, showing that system #3 is more interesting than the clustering system #2 in terms of loading capacity.

To choose the most suitable system between those presenting high purity and loading capacity, the three in silico criteria can be combined. This workflow could be fully automated so as to propose to the developer the best two combinations to test experimentally, discarding unsuitable combinations. In our example, the system #1 (CPC1xLC1 system) was found as the most suitable 2D system for the isolation of the eight targeted compounds. The system #4 (CPC4xLC1 system) comes in second position due to a lower  $di_{min}$  value. Of course, this combination can be adapted to the purification specifications by prioritizing one criterion over another.

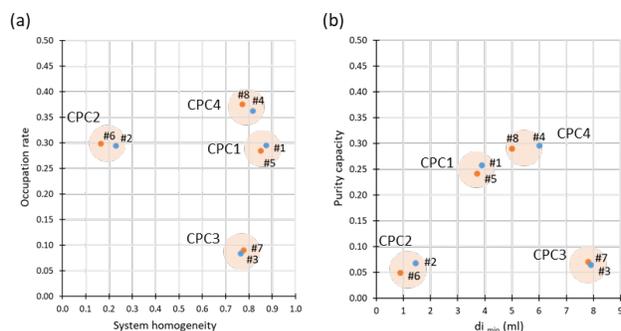


Figure 3. Predictive results for the eight 2D systems. (a) Occupation rate versus system homogeneity; (b) Purity capacity versus minimal distance between peaks. Combinations #1 to #4 are made with the phenyl hexyl column, combinations #5 to #8 with the C18 column.

### Experimental 2D-CPCxLC separations

To validate the predictive screening, the four 2D-CPCxLC systems using LC1 as second dimension were experimentally implemented in off-line comprehensive mode for a preparative purpose (Fig. 4). The CPC operating conditions were optimized for an effective stationary phase retention close to 50%. Since the *in silico* screening is based on retention data, the predictive 2D separation does not take into account the issue of transfer from the first to the second dimension. For the preparative separation to remain comprehensive, some conditions are required<sup>27</sup>. Each peak from the first dimension has to be cut at least in three fractions into their  $6\sigma$ -width to maintain <sup>1</sup>D-resolution. Depending on the efficiency of the <sup>1</sup>D-CPC separation and therefore the operating conditions, CPC1 was implemented with a sampling time of 0.5 min and CPC2, CPC3 and CPC4 with a sampling time of 1 min. With a preparative objective, the entire sample has to be subjected to the two dimensions without any split or loss, i.e. the full volume of the fraction coming from the first dimension has to be injected in the second dimension. For the development of the 2D-CPCxLC separations, the second dimension was implemented at a smaller scouting scale. In our lab, HPLC column was used to simulate the full transfer of the <sup>1</sup>D-fractions on larger LC column (10 mm i.d. x 150 mm), as previously described<sup>28</sup>. Thereby, an injection volume of 52% of the column void volume was considered for the transfer of CPC1 fractions, 65% for CPC2 and CPC3 fractions and 39% for CPC4 fractions. The peak band broadening due to injection effects being inversely related to the compression factor<sup>29</sup>, an off-line substitution of organic solvent by water in <sup>1</sup>D-fractions was considered to provide a <sup>2</sup>D-on-column refocusing<sup>30</sup>. Obviously, a preconcentration step could be set at this stage, but this was not the purpose of this study. Fraction concentration was kept constant during transfer and focusing was only occurring on the head of the <sup>2</sup>D column.

Comparing Figure 4 with Figure 2, the experimental 2D separations are obviously concordant with the predictive ones. Only some differences can be noted in <sup>1</sup>D-CPC retention times, due to the discrepancy between simulated and effective stationary phase retention and eventually between predicted and effective  $K_D$ -values. Hence, in system #1 (CPC1xLC1), the compounds D, E, G and H are more retained than expected ( $K_D > 1$ ; effective  $S_f$  75%). The phenomenon is even more visible on CPC3 where the effective  $S_f$  of 28% leads to a significant decrease in retention.

Concerning the separation quality, it is important to note that no peak distortion is observed in both dimensions for the targeted compounds. Our screening features (Fig. 3) indicated that system #1 (CPC1xLC1) and system #4 (CPC4xLC1) shall provide the best performances in terms of number of isolated compounds. This was confirmed by the experiments, as seven compounds of interest exhibit baseline return with these 2D systems, while this was the case for only five compounds with the CPC2xLC1 combination and five with the CPC3xLC1 combination.

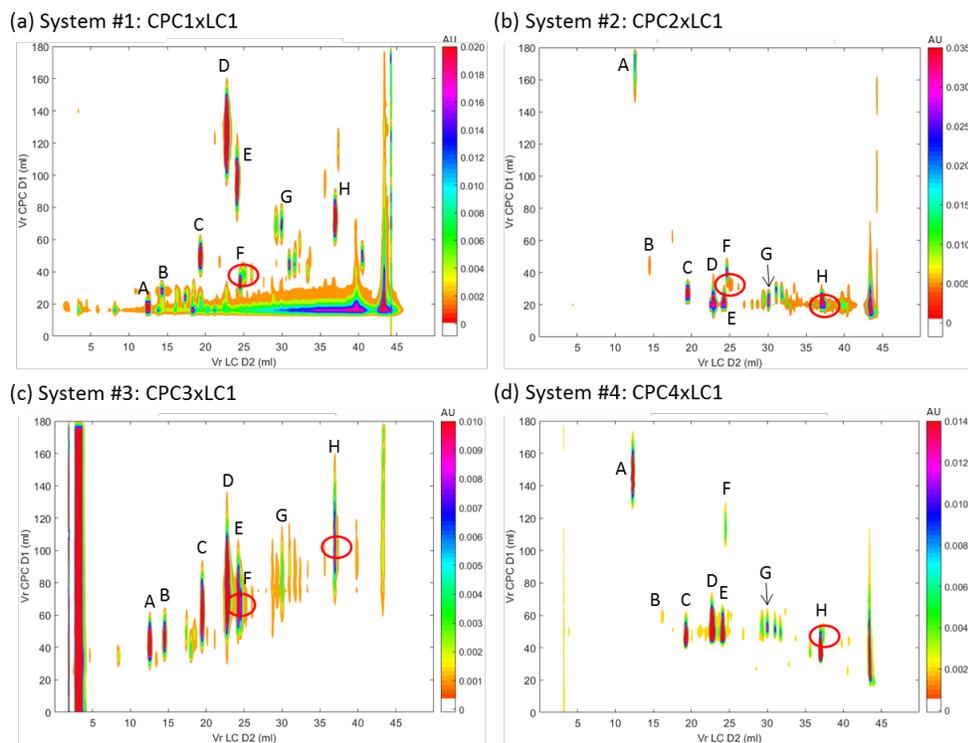


Figure 4. 2D-contour plots of off-line comprehensive (a) CPC1xLC1 separation, (b) CPC2xLC1 separation, (c) CPC3xLC1 separation and (d) CPC4xLC1 separation. Injection of 350  $\mu$ l of the *Cyclopia genistoides* sample (5 mg/ml in water). The red circles highlight peak overlapping.

In preparative separations, it is important to improve the sample load when possible in order to maximize the productivity for the isolation of major targeted compounds but also to recover minor compounds for structural elucidation. However, mono-dimensional HPLC methods are subjected to stationary phase overloading leading to peak distortion<sup>31</sup>, while CPC methods are only subjected to stationary phase loss<sup>32</sup>. While 2D LCxLC resolution can be strongly deteriorated with sample load, no peak distortions are noticed in 2D CPCxLC, thanks to the dilution effect of CPC<sup>28</sup>. When the sample load is increased (Fig. S2, supporting information), only CPC resolution is affected. Besides, the stationary phase loss can be avoided by studying the injection conditions in depth<sup>32</sup>.

Of course, the specificity of CPC, namely stationary phase content  $S_f$ , can be changed in our model, to simulate its influence on the performance quality of the 2D separation (Fig. S3, supporting information). For compounds with  $K_D$ -value over 1, the retention volume strongly increases with the retention rate of the stationary phase, while it decreases for compounds with low retention. Since the occupation rate is a ratio of areas which usually evolve in the same way, this criterion was found not to be affected by changes in stationary phase retention. The peak spreading is much more dependent on the  $K_D$ -values. When  $K_D$ -values are lower than 1, which is the case for system #2, users intuitively tend to increase the stationary phase amount to increase resolution of these poorly retained compounds. But this has the adverse effect of creating clusters, as the arithmetic mean is much higher than the harmonic mean leading to a substantial decrease in system homogeneity. In separations where compounds have  $K_D$ -values around 1 and few compounds with higher  $K_D$ -values, which is the case for systems #1 and #4, the arithmetic mean  $\bar{A}$  will be more important than the harmonic mean  $\bar{H}$  for small  $S_f$  leading to a decrease of the system homogeneity when the retention rate is also decreased. In that case, working at high  $S_f$  is more advantageous. Finally, where compounds with large  $K_D$ -values are present (as in system #3 with  $K_D$ -values between 2 and 10), the system homogeneity is not affected by a stationary phase change. The minimal

distance is dependent to the stationary phase changes, since the distance between peaks depend on their retention volumes thus the stationary phase retention. When the minimal distance comes from compounds with low  $K_D$ -values (system #2), this distance is not affected. Conversely, the minimal distance is even more affected that the  $K_D$ -values are higher than 1. Thus, developing methods with high  $K_D$ -values should be desirable to provide higher method robustness.

## CONCLUSION

While the interest in 2D preparative chromatography is growing, it is urgent to facilitate the selection of suitable combinations to isolate as many compounds as possible in a minimum number of experiments. The main purpose of this work is to provide a proof-of-concept for the 2D combination screening methodology, rather than precise optimization of the 2D preparative separation for all analytes in a particular application. Through the introduction of 2D-plots, it is possible to screen for a large set of 2D combinations using a minimal number of experiments to collect retention data and evaluate their preparative performances either visually or using in silico calculations based on the nearest-neighbor distances between peaks. We have selected performance metrics to be more adequate to preparative purposes than the conventional analytical 2D evaluation features. It appears from our example that while the separation mechanism orthogonality is a pre-requisite, clustering is the most important preparative challenge to overcome.

Though errors can occur for  $K_D$  estimation, and the equivalence of distances between each dimension, the proposed criteria have proven to be sufficient to screen and select adequate combinations for the efficient isolation of multiple compounds. Besides, in the case of a highly complex sample, it is possible to select a much larger number of target analytes to help with the orthogonality assessment.

In a similar manner, other 2D preparative setups such as LCxLC could be compared using relevant models. However, peak distortion during sample loading may be more difficult to predict.

Last but not least, with retention and kinetic data available, it is now also possible to compare various operating conditions ( $S_f$  and rotation speed for CPC, gradient slope and composition for LC) virtually and hence run only the optimized conditions.

## ASSOCIATED CONTENT

HPLC chromatograms of the plant extract (Fig. S1); 2D-contour plots of CPC1xLC1 separations at different sample loads (Fig. S2); Evolution of (a) system homogeneity, (b) minimal distance between peaks in function of stationary phase retention rate (Fig. S3); Operating conditions for CPC methods (Table S1);  $K_D$ -value measurements of compounds in the selected CPC systems (Table S2); Calculation of the harmonic mean  $\bar{H}$  and the arithmetic mean  $\bar{A}$  of the nearest-neighbor distances (Equations) (PDF)

## AUTHOR INFORMATION

### Corresponding author

Phone: +33 0437423686

E-mail: karine.faure@isa-lyon.fr

Address: Institut des Sciences Analytiques, UMR 5280, 5 rue de la Doua, F-69100 Villeurbanne, France  
ORCID Karine Faure 0000-0002-5416-4911

## REFERENCES

- (1) Corgier, A.; Sarrut, M.; Crétier, G.; Heinisch, S. *Chromatographia* **2016**, *79*, 255-260.
- (2) Wehrli, A.; Hermann, U.; Huber, J. F. K. *Journal of Chromatography A* **1976**, *125*, 59-70.
- (3) Marlot, L.; Faure, K. *Journal of Chromatography A* **2017**, *1494*, 1-17.
- (4) Fu, Q.; Guo, Z.; Zhang, X.; Liu, Y.; Liang, X. *Journal of separation science* **2012**, 1821-1827.

- (5) Li, K.; Zhu, W.; Fu, Q.; Ke, Y.; Jin, Y.; Liang, X. *Analyst* **2013**, *138*, 3313-3320.
- (6) Qiu, Y.-K.; Chen, F.-F.; Zhang, L.-L.; Yan, X.; Chen, L.; Fang, M.-J.; Wu, Z. *Analytica Chimica Acta* **2014**, *820*, 176-186.
- (7) Wang, X.-Y.; Li, J.-F.; Jian, Y.-M.; Wu, Z.; Fang, M.-J.; Qiu, Y.-K. *Journal of Chromatography A* **2015**, *1387*, 60-68.
- (8) Li, J.-F.; Fang, H.; Yan, X.; Chang, F.-R.; Wu, Z.; Wu, Y.-L.; Qiu, Y.-K. *Journal of Chromatography A* **2016**, *1456*, 169-175.
- (9) Yan, X.; Wang, L.-J.; Wu, Z.; Wu, Y.-L.; Liu, X.-X.; Chang, F.-R.; Fang, M.-J.; Qiu, Y.-K. *Journal of Chromatography B* **2016**, *1033-1034*, 1-8.
- (10) Yunpeng, F.; Yanhui, F.; Qing, F.; Jianfeng, C.; Huaxia, X.; Mei, D.; Yu, J. *Journal of Separation Science* **2016**, *39*, 2710-2719.
- (11) Walters, N. A.; de Villiers, A.; Joubert, E.; de Beer, D. *Journal of Chromatography A* **2017**, *1490*, 102-114.
- (12) Le Masle, A.; Santin, S.; Marlot, L.; Chahen, L.; Charon, N. *Analytica Chimica Acta* **2018**, *1029*, 116-124.
- (13) Stoll, D. R.; Li, X.; Wang, X.; Carr, P. W.; Porter, S. E. G.; Rutan, S. C. *Journal of Chromatography A* **2007**, *1168*, 3-43.
- (14) Joubert, E.; Joubert, M. E.; Bester, C.; de Beer, D.; De Lange, J. H. *South African Journal of Botany* **2011**, *77*, 887-907.
- (15) Beelders, T.; De Beer, D.; Stander, M.; Joubert, E. *Molecules* **2014**, 11760-11790.
- (16) Kokotkiewicz, A.; Luczkiewicz, M.; Pawlowska, J.; Luczkiewicz, P.; Sowinski, P.; Witkowski, J.; Bryl, E.; Bucinski, A. *Fitoterapia* **2013**, *90*, 199-208.
- (17) Ntlhokwe, G.; Muller, M.; Joubert, E.; G.J. Tredoux, A.; Villiers, A. *Journal of Chromatography A* **2017**.
- (18) D'Attoma, A.; Grivel, C.; Heinisch, S. *Journal of Chromatography A* **2012**, *1262*, 148-159.
- (19) Neue, U. D.; O'Gara, J. E.; Méndez, A. *Journal of Chromatography A* **2006**, *1127*, 161-174.
- (20) Van Gysegem, E.; Van Hemelryck, S.; Daszykowski, M.; Questier, F.; Massart, D. L.; Vander Heyden, Y. *Journal of Chromatography A* **2003**, *988*, 77-93.
- (21) Liu, Z.; Patterson, D. G.; Lee, M. L. *Analytical Chemistry* **1995**, *67*, 3840-3845.
- (22) Slonecker, P. J.; Li, X.; Ridgway, T. H.; Dorsey, J. G. *Analytical Chemistry* **1996**, *68*, 682-689.
- (23) Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. *Analytical Chemistry* **2005**, *77*, 6426-6434.
- (24) Semard, G.; Peulon-Agasse, V.; Bruchet, A.; Bouillon, J.-P.; Cardinaël, P. *Journal of Chromatography A* **2010**, *1217*, 5449-5454.
- (25) Nowik, W.; Héron, S.; Bonose, M.; Nowik, M.; Tchaplá, A. *Analytical Chemistry* **2013**, *85*, 9449-9458.
- (26) Nowik, W.; Bonose, M.; Héron, S.; Nowik, M.; Tchaplá, A. *Analytical Chemistry* **2013**, *85*, 9459-9468.
- (27) Marriott, P. J.; Wu, Z.-y.; Schoenmakers, P. *LCGC Europe* **2012**, *25*.
- (28) Marlot, L.; Batteau, M.; Faure, K. *Electrophoresis* **2018**, *39*, 2011-2019.
- (29) Bedani, F.; Schoenmakers, P. J.; Janssen, H.-G. *Journal of Separation Science* **2012**, *35*, 1697-1711.
- (30) Jandera, P.; Česla, P.; Hájek, T.; Vohralík, G.; Vyňuchalová, K.; Fischer, J. *Journal of Chromatography A* **2008**, *1189*, 207-220.
- (31) Seidel-Morgenstern, A.; Schulte, M.; Epping, A. In *Preparative Chromatography*; Wiley-VCH Verlag GmbH & Co. KGaA, **2012**, pp 7-46.
- (32) Marchal, L.; Intes, O.; Foucault, A.; Legrand, J.; Nuzillard, J.-M.; Renault, J.-H. *Journal of Chromatography A* **2003**, *1005*, 51-62.

for TOC only

