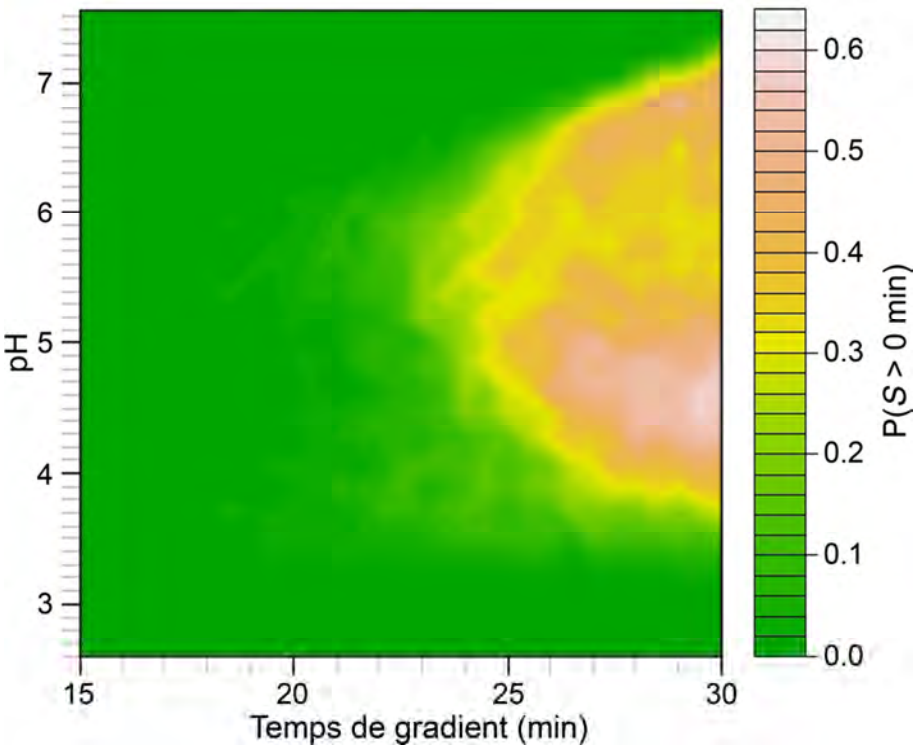
UNIVERSITE DE LIEGE
Faculté de Médecine
Laboratoire de Chimie Analytique
Professeur Philippe HUBERT

Nouvelle méthodologie pour le développement de méthodes chromatographiques



Benjamin DEBRUS
Licencié en Sciences Chimiques
Diplômé d'études approfondies en Sciences
Diplômé d'études approfondies en Sciences Pharmaceutiques

Thèse présentée en vue de l'obtention du grade de Docteur en Sciences Biomédicales et Pharmaceutiques
Année académique 2010-2011

"*La pierre n'a point d'espoir d'être autre chose qu'une pierre.*
Mais, de collaborer, elle s'assemble et devient temple."
Alexandre de Saint-Exupéry

Ce projet de recherche n'aurait pu aboutir sans la riche collaboration que j'ai pu avoir avec de nombreuses personnes ainsi que sans l'aide précieuse et le soutien infaillible qu'elles ont pu me procurer. Je désire leur témoigner ma gratitude via ces quelques lignes.

Je tiens tout d'abord à remercier Monsieur le Professeur Philippe Hubert, promoteur de cette thèse, pour m'avoir accueilli au sein de son laboratoire et m'avoir donné l'opportunité de réaliser un travail de recherche passionnant. Je désire lui exprimer ma profonde gratitude pour son aide constante, ses conseils toujours judicieux et ses préceptes qui se sont constamment révélés être d'une profonde sagesse.

Je désire également exprimer mes plus chaleureux remerciements à Monsieur Attilio Ceccato, Docteur en Sciences Pharmaceutiques et co-promoteur de cette thèse, pour m'avoir fait bénéficier de son expérience et de ses compétences. Outre ses nombreuses connaissances dans le domaine pharmaceutique qu'il a tenté de m'inculquer, je tiens à le remercier pour son aide inestimable.

Je souhaite aussi remercier Monsieur le Professeur Bruno Boulanger d'avoir consacré du temps afin de me faire profiter de ses connaissances et de son expérience considérable en statistiques. Je le remercie pour avoir toujours réussi à éveiller ma curiosité scientifique par sa vision de la statistique appliquée à la chimie analytique.

Un remerciement très particulier s'adresse à Monsieur Pierre Lebrun sans qui ce travail n'aurait probablement pas pu atteindre de telles dimensions. L'étroite collaboration que nous avons entretenue, et que j'espère nous entretiendrons encore longtemps, m'a permis d'acquérir de vastes connaissances aussi bien en statistiques qu'en informatique et en programmation. Ces quelques lignes ne suffiront pas pour lui témoigner ma profonde reconnaissance.

Mes sincères remerciements s'adressent aussi à tous les membres du laboratoire de chimie analytique, pour leur bonne humeur, leur amitié et les captivantes discussions partagées. Un remerciement particulier est destiné à Monsieur Frédéric Lecomte pour m'avoir inculqué les bonnes pratiques en matière d'analyse par chromatographie liquide dès mon arrivée au sein du laboratoire. Un autre remerciement particulier s'adresse à Monsieur Eric Rozet, Docteur en Sciences Pharmaceutiques, pour son aide très utile, ses remarques et conseils avisés.

Je remercie également Madame le Professeur Coral Barbas, de l'Université San Pablo CEU de Madrid, et Messieurs les Professeurs Jean-Luc Veuthey et Serge Rudaz, de l'Université de Genève, pour m'avoir accueilli au sein de leur laboratoire lors des stages que j'y ai réalisés.

Je remercie la Région Wallonne, la société Eli Lilly and Company, la société Arlenda, le Fonds Léon Frédéricq, le Fonds national de la Recherche Scientifique et l'Université de Liège, pour leur soutien financier.

Mes sincères remerciements s'adressent à Messieurs les Membres du Comité de thèse pour leur disponibilité et leurs précieux conseils prodigués tout au long de ce travail ainsi qu'à Messieurs les Membres du Jury qui me font l'honneur de juger ce travail.

Je tiens également à remercier mes amis, mes proches ainsi que les personnes qui, de près ou de loin, m'ont soutenu par leurs paroles et leurs encouragements.

También me gustaría agradecer a la Srta. Carmen Fernández quien siempre ha estado a mi lado.

Enfin, je tiens tout particulièrement à remercier du fond du cœur mes parents et mon frère pour leur soutien, leur patience et leurs encouragements durant toutes ces années d'études.

Table des matières

Chapitre I

Introduction

I.1. La chromatographie liquide

C'est en 1901 que Mikhail Semyonovich Tswett (1872-1919), désirant filtrer un broyât de feuilles vertes dans un tube de verre rempli de carbonate de calcium, découvrit la technique analytique qui donna naissance à la chromatographie liquide (CL). Cette méthode fut décrite le 30 décembre 1901 au 11ème congrès des naturalistes et des médecins à Saint–Pétersbourg et publiée dans la revue *Proceedings of the Warsaw Society of Naturalists* en 1903. Tswett utilisa, pour la première fois, le terme chromatographie dans le journal de la société Allemande de botanique [1]. Au cours du XX^{e} siècle, de nombreux chercheurs firent évoluer la chromatographie découverte par Tswett. Le moment le plus remarquable lors de l'essor de la chromatographie fut probablement le prix Nobel de chimie attribué à Archer J.P. Martin et Richard L.M. Synge en 1952 pour leur invention de la chromatographie de partage [2].

De nos jours, la chromatographie liquide, également appelée chromatographie liquide à haute performance (CLHP), peut être définie comme étant un procédé physico-chimique de séparation des constituants d'un mélange entre deux phases non miscibles dont l'une, dite stationnaire (solide), est emprisonnée dans une colonne ou fixée sur un support et l'autre, dite mobile (liquide), se déplace au contact de la première.

I.1.1. Utilisation de la chromatographie liquide

Actuellement, la CLHP est probablement la méthode d'analyse la plus répandue aussi bien dans les secteurs académique que privé. Comme illustré à la figure I.1, la majorité des travaux de recherches présentés au congrès *Drug Analysis* 2010, tant en recherche fondamentale qu'en recherche appliquée, se base et/ou utilise la CL comme outil analytique. Il ressort également de cette figure que les techniques de détection par spectrométrie de masse (SM) sont de plus en plus utilisées, notamment grâce à la simplification de l'interfaçage CL-SM ainsi qu'à la relative démocratisation des prix des appareillages de SM (simple quadrupôle) [3]. On remarque également que la CL de nouvelle génération (la chromatographie liquide à ultra–haute performance : CLUHP) semble être en bonne voie pour détrôner la CLHP en tant que technique analytique de référence.

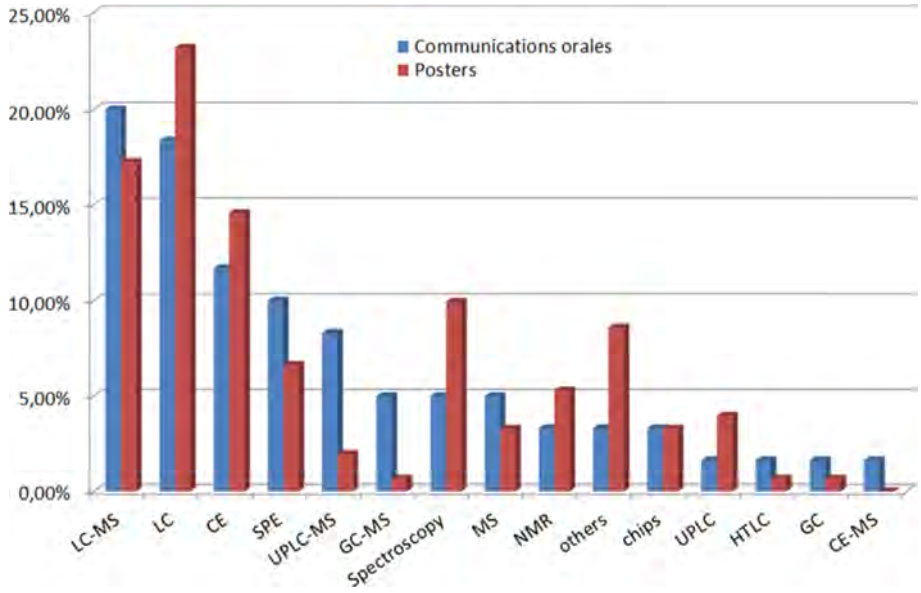


Fig. I.1 Distribution des techniques analytiques utilisées lors des travaux présentés au congrès *Drug Analysis 2010* à Anvers.

Cette tendance s'observe également si on analyse le nombre de publications parues dans le domaine au cours de ces dernières années (figure I.2). Notons toutefois que, sur la figure I.2, le nombre de publications en CLUHP a été multiplié par 10 afin de faciliter la visualisation de la croissance apparemment exponentielle du nombre de publications faisant référence à cette nouvelle technique chromatographique.

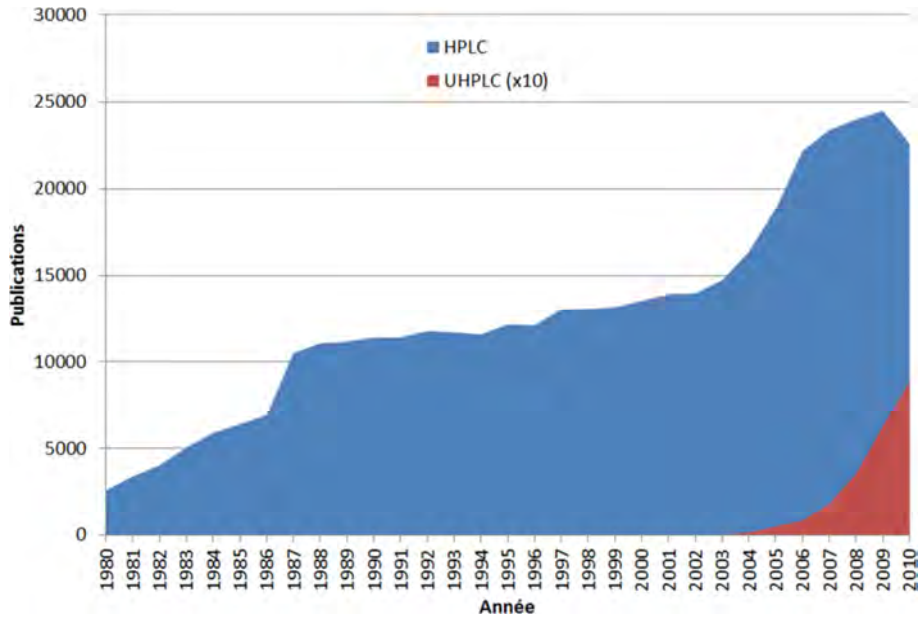


Fig. I.2 Nombre de publications se basant sur la CLHP et la CLUHP entre 1980 et 2010 (recherche bibliographique effectuée avec SciFinder®).

En 2009, la société *Strategic Directions International Inc.* a réalisé une étude portant sur le marché de la CLHP. Elle estime que le marché total de la CLHP atteindra le montant de 4,7 milliards de dollars en 2012 et que la croissance de ce secteur ne montre aucun signe de

fléchissement [4]. En octobre 2010, la société *Global Industry Analysts* a révisé ce chiffre en estimant que ce marché atteindra la valeur de 3,7 milliards de dollars en 2015 [5]. Cette tendance d'accroissement provient en partie de l'augmentation de la demande d'analyse et d'instrumentation dans le domaine des sciences de la vie générée par l'industrie pharmaceutique dont le marché mondial s'élève à plus de 875 milliards de dollars pour une croissance estimée de 4 à 6% en 2011. Ces observations permettent d'affirmer que la recherche fondamentale et appliquée dans le domaine de la CL s'insère parfaitement dans le contexte économique et stratégique actuel. En outre, la CLHP est la technique analytique de référence en ce qui concerne les dossiers d'enregistrement de médicaments. Ainsi, la grande majorité des méthodes de recherche d'impuretés reprises par exemple dans les pharmacopées européenne ou américaine font usage de la CLHP. Il en est de même en ce qui concerne le suivi des études de stabilité des spécialités pharmaceutiques.

I.1.2. Principes de bases

La CLHP fait notamment partie de la chromatographie d'adsorption. Un composé placé en solution est capable de s'adsorber sur un support solide. Cette adsorption est régie par un équilibre dynamique caractérisé par une constante d'équilibre ou coefficient de partage, *K*, qui se défini comme le rapport entre la concentration en phase stationnaire solide (C_S) et la concentration en phase mobile liquide (C_M) comme décrit par l'équation 1.

$$K = \frac{C_S}{C_M} \quad (1)$$

Si les interactions entre le soluté et la phase mobile sont fortes (molécule hautement solvatée), C_M sera relativement plus grand que C_S et la molécule aura une vitesse linéaire inférieur mais proche de celle de la phase mobile. A l'inverse, si le soluté interagit fortement avec la phase stationnaire, sa vitesse linéaire sera nettement réduite.

En CLHP, la phase stationnaire est majoritairement composée de particules sphériques de silice greffées ou non et placées dans une colonne d'acier inoxydable. La phase mobile se compose d'un mélange hydro-organique. Si la polarité de la phase stationnaire est plus élevée que celle de la phase mobile, on parle de CLHP en phase normale. Le cas contraire, où la polarité de la phase mobile est plus élevée que celle de la phase stationnaire, correspond à la CLHP à polarités de phase inversées. Dans ce cas, des chaines hydrocarbonées, par exemple n-$C_{18}H_{37}$, sont greffées sur la silice afin d'en diminuer la polarité et la phase mobile se compose d'un mélange d'eau ou de solutions tampons et de solvants organiques, tels que le méthanol (MeOH), l'acétonitrile (ACN) ou le tétrahydrofurane (THF).

L'instrumentation en CLHP se compose d'une pompe haute pression permettant de faire percoler, à un débit stable et déterminé, une phase mobile au travers d'une colonne remplie de particules de silice modifiée ou non. Un système d'injection est placé entre le système de pompage et la colonne et permet l'injection des analytes dans le système chromatographique. Finalement, la colonne, le plus souvent thermostatée, est connectée au

détecteur. L'ensemble est relié à un ordinateur permettant le pilotage du système et l'enregistrement des signaux générés par le détecteur. En CLHP, les particules contenues dans la colonne ont généralement des tailles comprises entre 3.5 et 10 µm.

I.1.3. Grandeurs fondamentales

A un débit *D* donné, la vitesse linéaire *u* de la phase mobile dépendra du volume mort V_0 d'une colonne de longueur *L* (Eq. 2). Ce volume mort provient principalement du volume interstitiel entre les particules de silice.

$$u = \frac{L \times D}{V_0} \tag{2}$$

Un composé non retenu est dès lors détecté au temps de rétention nul $t_0 = L/u$. Un composé retenu est quant à lui détecté au temps de rétention, t_R, qui est obligatoirement plus grand ou égal à t_0. Le temps de rétention d'un composé est défini comme le temps observé au sommet du pic chromatographique qui lui correspond. Afin de s'affranchir des paramètres géométriques du système chromatographique, le facteur de rétention peut être calculé grâce à l'équation 3. Ce dernier est particulièrement pratique car il est indépendant du débit et des dimensions de la colonne.

$$k = \frac{t_R - t_0}{t_0} \tag{3}$$

Dans le but de quantifier la séparation entre deux pics, le paramètre le plus répandu est la résolution (cf. Eq. 4).

$$R_S = 2 \times \frac{t_{R,2} - t_{R,1}}{\omega_1 + \omega_2} = 1.18 \times \frac{t_{R,2} - t_{R,1}}{\delta_1 + \delta_2} \tag{4}$$

$t_{R,1}$ et $t_{R,2}$ sont les temps de rétention des deux pics concernés et $t_{R,1} \leq t_{R,2}$. ω_1 et ω_2 sont les largeurs correspondantes à la ligne de base définie par les tangentes aux points d'inflexion. δ_1 et δ_2 sont les largeurs respectives à mi-hauteur.

Un autre paramètre permettant d'évaluer la séparation entre deux pics est la sélectivité (α). Ce paramètre sans dimension se définit comme le rapport entre les facteurs de rétention des deux pics concernés (cf. Eq. 5).

$$\alpha = \frac{k_2}{k_1}, \qquad k_2 \geq k_1 \tag{5}$$

Afin de pouvoir comparer l'efficacité d'un système chromatographique (et plus particulièrement d'une colonne) à fournir des bandes d'élutions fines, le nombre de plateau théorique *N* et la hauteur équivalente à un plateau théorique *H* peuvent être calculés. Ces deux valeurs sont reliées par l'équation 6 où *L* représente la longueur de la colonne.

$$H = \frac{L}{N} \tag{6}$$

Ces termes ont été introduits par Martin et Synge [2], en considérant qu'une colonne était constituée de nombreux plateaux discrets et contigus appelés plateaux théoriques. Pour un

pic d'allure gaussienne, on peut calculer le nombre de plateaux théoriques au moyen de l'équation 7.

$$N = 16 \times \left(\frac{t_R}{\omega}\right)^2 = 5.54 \times \left(\frac{t_R}{\delta}\right)^2 \quad (7)$$

Pour un pic asymétrique, on préfèrera la relation de Foley et Dorsey décrite à l'équation 8.

$$N = \frac{41.7 \times \left(\frac{t_R}{\omega_{0,1}}\right)^2}{\left(B/A\right) + 1,25} \quad (8)$$

A et *B* étant les demi-largeurs (respectivement à gauche et à droite) du pic à 10% de sa hauteur et $\omega_{0,1}$ la largeur à 10% de sa hauteur.

La loi de Darcy (Eq. 9) est à l'hydrodynamique ce que la loi d'Ohm est à l'électricité. Cette loi définit la pression *P* (ou perte de pression ΔP) fournie par un système chromatographique en fonction de la viscosité de la phase mobile η, de la taille de particules d_p, de la longueur de la colonne *L*, de la vitesse linéaire de la phase mobile *u* et d'un facteur de résistance à l'écoulement Φ dépendant de la forme des particules et de la qualité du remplissage de la colonne.

$$\Delta P = \Phi \times \frac{\eta \times L \times u}{d_p^2} \quad (9)$$

La vitesse linéaire de la phase mobile, qui est directement reliée au débit, est un paramètre de première importance lorsque qu'on désire optimiser l'efficacité d'un système chromatographique. L'équation de van Deemter relie cette dernière avec la hauteur équivalente des plateaux théoriques tel que décrit à l'équation 10.

$$H = A + \frac{B}{u} + C \times u \quad (10)$$

Dans cette expression, *A* symbolise le terme de diffusion turbulente qui est représentatif des différents chemins pouvant être empruntés entre les particules de silice par une molécule. Ce terme est proportionnel à la taille des particules et indépendant de la vitesse linéaire de la phase mobile. *B* est le terme de diffusion longitudinale qui correspond à l'élargissement de la bande d'élution dû au mouvement brownien des molécules dans la phase mobile. Cet effet est fortement moins marqué lorsque *u* augmente. *B* est proportionnel au coefficient de diffusion des molécules dans la phase mobile. *C* correspond à la résistance au transfert de masse. Ce terme introduit des considérations cinétiques dans l'équilibre dynamique décrit à l'équation 1. Il dépend du coefficient de diffusion des molécules en phase liquide (proportionnel à l'habilité d'un analyte en solution de s'adsorber sur la phase solide) et du coefficient de diffusion des molécules en phase solide (proportionnel à l'habilité d'un analyte adsorbé de retourner en phase liquide). Ce terme augmente de façon linéaire avec la vitesse linéaire de la phase mobile (*u*). La courbe de van Deemter (figure I.3) est la résultante de ces trois termes.

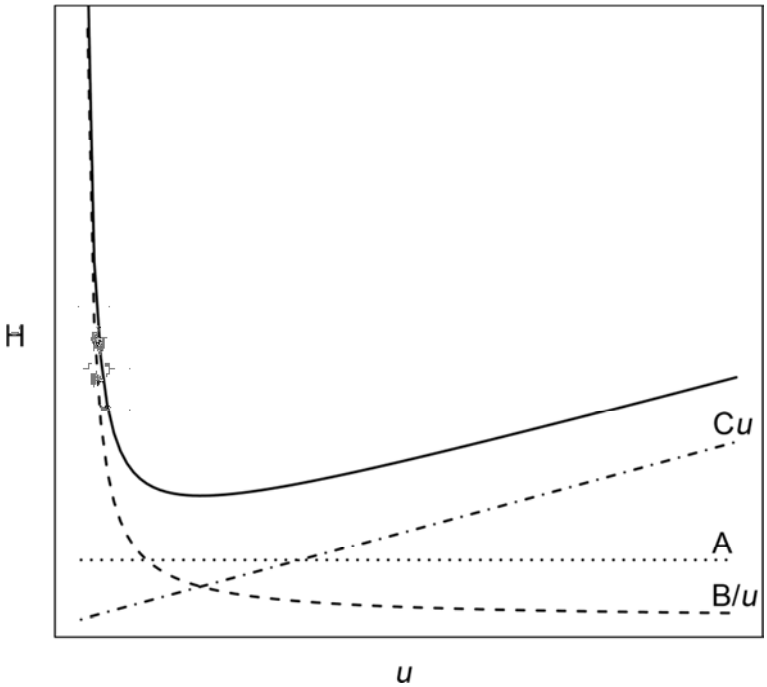


Fig. I.3 Exemple de courbe de van Deemter (courbe noire) où la hauteur équivalente aux plateaux théoriques (*H*) est affichée en fonction de la vitesse linéaire de la phase mobile (*u*). La valeur de *u* où *H* est minimal est appelée vitesse linéaire optimale ou minimum de van Deemter.

I.1.4. Evolution de la chromatographie liquide

Ces dernières années, de nouvelles avancées ont été réalisées dans le domaine de la CL afin d'augmenter l'efficacité et la rapidité des analyses. Ces avancées peuvent être réparties dans trois domaines de recherche : la chromatographie liquide sur supports monolithiques (CLM), la chromatographie à haute température (CLHT) et la chromatographie liquide à ultra-haute performance (CLUHP).

I.1.4.1. La chromatographie sur support monolithiques

La CLM se base sur l'utilisation de colonne remplie avec un support polymérique monolithique. J.H. Knox et P.A. Bristow de l'Université d'Edinburgh reconnurent les avantages potentiels de la CLM il y a plus de trente ans [6]. L'avantage principal de la CLM réside dans le fait que le support monolithique offre des structures macroporeuses (2 µm) et des structures mésoporeuses (10-15 nm). Les macropores offrent une faible résistance à l'écoulement et permettent l'augmentation du débit pour réduire le temps d'analyse tout en maintenant la pression à des valeurs acceptables. En comparaison à une colonne remplie des particules de 5 µm en CLHP, la CLM offre des pressions de 3 à 4 fois plus faibles pour un débit identique. Les mésopores, quant à eux, permettent d'obtenir une grande surface spécifique. En termes de performance chromatographique, l'utilisation de monolithes est comparable à l'utilisation de colonnes de silice remplies avec des particules de 3 à 3,5 µm [7]. La comparaison des courbes de van Deemter entre une colonne de silice remplie de particules de 5 µm et une colonne monolithique permet de conclure que les supports

monolithiques offrent de meilleures efficacités (le *H* minimum est plus faible) à un débit plus élevé (analyse plus rapide) [8].

I.1.4.2. La chromatographie à haute température

L'utilisation de haute température, typiquement de 80 à 200 °C, induit une réduction, de 5 à 10 fois, de la viscosité de la phase mobile permettant l'utilisation de débits plus élevés tout en maintenant la pression à des valeurs acceptables. De plus, la vitesse linéaire optimale est nettement améliorée sans impact sur l'efficacité car le coefficient de diffusion des composés augmente avec la température [9,10]. La CLHT offre également d'autres avantages tels que la diminution de la consommation en solvant grâce à la diminution de la polarité de la phase mobile [11], une amélioration de la symétrie des composés basiques [12] et des modifications de sélectivité. La température peut alors être sélectionnée dans le cadre de développements de méthodes chromatographiques. Néanmoins, travailler à de hautes températures nécessite de s'assurer de la stabilité des analytes et des phases stationnaires utilisées dans ces conditions.

I.1.4.3. La chromatographie liquide à ultra-haute performance

Afin d'augmenter l'efficacité des méthodes chromatographiques, une dernière stratégie consiste en la diminution de la taille des particules de silice. En CLHP, les particules s'échelonnent entre 3.5 et 10 µm alors qu'en CLUHP, la taille des particules est égale ou inférieure à 2 µm. L'impact de cette diminution de la taille des particules sur l'efficacité chromatographique est notable et il est possible de l'observer par la comparaison des courbes de van Deemter en CLHP et CLUHP. Premièrement, la hauteur équivalente de plateaux théoriques minimale diminue. La CLUHP est donc une technique chromatographique plus efficace. Deuxièmement, la vitesse linéaire optimale augmente. La CLUHP permet donc d'obtenir des séparations identiques en un temps d'analyse considérablement réduit [13]. Si l'on considère une méthode analytique utilisant une colonne de 150 mm de long remplie de particules de 5 µm, le passage à une colonne de 50 mm remplie de particules de 1,7 µm engendrera une réduction du temps d'analyse d'un facteur 9 pour une efficacité identique. D'un autre côté, l'utilisation d'une colonne de 150 mm remplie de particules de 1,7 µm permettra l'obtention d'une efficacité 3 fois supérieure tout en réduisant le temps d'analyse d'un facteur 3 [14]. Dans les deux cas, le prix à payer réside dans le fait que la pression sera drastiquement augmentée, nécessitant l'utilisation d'équipements pouvant supporter des pressions allant jusque 1200 bars. Bien évidemment, il faut noter qu'il est toujours possible de combiner différentes stratégies, comme la réduction de la taille des particules et l'utilisation de températures élevées, pour améliorer l'efficacité et la rapidité des analyses en chromatographie liquide [15].

I.1.5. Développement de méthodes en chromatographie liquide

I.1.5.1. Objectifs

Lors du développement d'une méthode chromatographique, la première étape consiste à identifier les critères à optimiser. Le critère principal et non facultatif est l'optimisation de la séparation des composés. Le critère chromatographique généralement utilisé dans ce but est la résolution (Eq. 4). Ensuite, d'autres critères comme par exemple, le temps d'analyse, la robustesse, la répétabilité des résultats peuvent être introduits.

L'objectif de la méthode chromatographique dépend principalement de son champ d'application.

- Dans le cadre de la mise au point de formulation pharmaceutique, il convient de développer une méthode robuste qui permet de séparer tous les composés en vue de leur dosage.
- On peut également envisager de développer des méthodes de *screening* (méthodes de criblage) qui se focalisent sur l'identification de principes actifs comme en toxicologie ou dans la lutte contre la contrefaçon.
- Dans le domaine de la bioanalyse ou de l'analyse d'extraits de plantes, la complexité des échantillons, en termes de nombre de composés, rend souvent difficile l'obtention d'une séparation complète de tous les composés. Il est alors généralement admis de maximiser le nombre de composés séparés et/ou de se concentrer sur la séparation d'un ou plusieurs composés d'intérêt.
- Enfin, les analyses de mélanges provenant de réactions chimiques (analyse du milieu réactionnel afin d'évaluer le suivi d'une réaction chimique) ainsi que les analyses effectuées dans le but d'évaluer la stabilité d'un produit requièrent des temps d'analyses suffisamment courts au vu du nombre important d'échantillons à analyser.

Dans tous les cas, dès que l'objectif de la méthode est identifié, il convient de choisir une stratégie d'optimisation. Deux grands types de stratégie s'offrent à l'utilisateur : les méthodes univariées et les méthodes multivariées.

Les méthodes univariées tentent l'optimisation de la séparation en ne faisant varier qu'un facteur à la fois. Actuellement, certains logiciels commerciaux permettent d'optimiser des méthodes chromatographiques. Citons par exemple Drylab [16], Chromsword [17], ACD/LC simulator [18] et Osiris [19]. Ces logiciels se basent sur la théorie solvophobique [20-25] et/ou sur la *linear solvent strength theory* [26-29]. La limitation majeure de ces logiciels d'optimisation réside dans le fait qu'ils ne permettent d'optimiser les méthodes qu'en fonction de deux facteurs préétablis par le fabricant. Les facteurs pouvant être sélectionnés sont le pourcentage de modificateur organique (MeOH, ACN ou THF) en mode isocratique, le temps ou la pente du gradient en élution graduée, la température et le pH de la phase mobile. Récemment, La dernière version de certains d'entre eux propose d'optimiser les méthodes selon trois de ces facteurs. Bien que les facteurs sélectionnés soient ceux qui ont

généralement le plus d'effet sur la séparation et la sélectivité, l'utilisateur se trouve dès lors limité car il ne pourra réaliser d'optimisations de méthodes chromatographiques sur d'autres facteurs tels que la concentration en agent d'appariement d'ions. De plus, ces programmes ne réalisent aucune gestion de l'erreur prédictive et donnent des temps de rétention et des résolutions moyennes. Il n'est dès lors pas rare d'observer une certaine dissimilarité entre prédictions et observations même si la majorité des cas semblent favorables [30-32]. Néanmoins, ces stratégies d'optimisation ont l'avantage d'être efficaces et rapides.

Les méthodes multivariées prônent, quant à elles, la réalisation d'expériences à des conditions opératoires où tous les facteurs ont été modifiés. Dans ce cas, les plans d'expériences, détaillés à la section I.3.1., représentent la technique de choix.

I.1.5.2. Robustesse et optimisation de méthodes

La robustesse est définie en détail à la section suivante (section I.2.2.). Néanmoins, pour brièvement introduire ce concept, la robustesse peut être définie comme la capacité d'un système ou d'un état à rester inchangé malgré de faibles modifications des paramètres le définissant. En CLHP, une des manières d'évaluer la robustesse d'une méthode est de délibérément faire varier certains facteurs et d'observer si la variation de certains critères prédéfinis reste dans des limites acceptables [33-35]. Ces méthodes d'évaluation de la robustesse sont basées sur l'utilisation de plans d'expériences qui sont décrits en détail à la section I.3.1.

De plus, la tendance actuelle veut que la robustesse soit estimée en même temps que l'optimisation de la méthode chromatographique. Par conséquent, si l'estimation de la robustesse se base sur la planification expérimentale et que cette dernière semble devoir être réalisée conjointement à l'optimisation de la méthode chromatographique, l'utilisation des plans d'expériences à cette fin semble être une évidence. Par ailleurs, les travaux fondateurs de W. Dewé et al. [36] prônent l'utilisation de la planification expérimentale pour modéliser les temps de rétention et optimiser les méthodes en CLHP.

I.2. Contexte règlementaire pharmaceutique

Dans l'industrie pharmaceutique, l'objectif principal est de fournir aux patients des produits efficaces et de qualité. Les processus de contrôle de la qualité se retrouvent donc tout au long du procédé de production. Afin de prescrire des garanties identiques pour toutes les industries pharmaceutiques, des organismes tel que la conférence internationale sur l'harmonisation des exigences techniques pour l'enregistrement des produits pharmaceutiques à usage humain (*International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use* ; ICH) ont été créés. La mission première de l'ICH est de fournir des recommandations en vue d'atteindre une meilleure harmonisation dans l'interprétation et l'application des directives et exigences techniques pour l'enregistrement des produits pharmaceutiques permettant notamment de réduire ou d'éviter la duplication des tests effectués au cours de la recherche et du développement de nouveaux médicaments à usage humain.

Dans le cadre du développement de méthodes chromatographiques, l'ICH publie certaines directives se basant sur de multiples concepts tels que le "*quality by design*", le "*design space*" et le "*quality risk management*" qui sont décrits ci-après.

I.2.1. Le concept de *Quality by Design*

Le concept de "*Quality by design*" (QbD) pourrait être traduit en français par "la qualité par la conception". Le document ICH Q8(R2) [37] définit le concept de QbD comme étant une approche systématique du développement qui commence avec des objectifs prédéfinis et mettant l'accent sur le produit, la compréhension des processus et le contrôle des processus, fondée sur des bases scientifiques solides et la gestion des risques de non-qualité.

Les notions fondatrices sont la conception du développement impliquant la qualité et la gestion des risques. Dans le contexte pharmaceutique, la qualité peut être définie comme la conformité ou le caractère approprié d'une substance médicamenteuse ou d'un produit médicamenteux pour l'usage prévu.

I.2.2. Le concept de *Design Space*

Le concept de "*Design Space*" (DS) pourrait être traduit en français par "l'espace de conception". Il faut être attentif au fait que le DS ne définit pas l'espace de travail ou le domaine expérimental dans lequel un procédé sera conçu mais plutôt l'espace dans lequel les objectifs prédéfinis seront atteints.

La directive Q8(R2) [37] de l'ICH définit le DS comme la combinaison multidimensionnelle où l'interaction des variables d'entrée et les paramètres du procédé ayant permis de donner l'assurance de la qualité. En d'autres termes, le DS est un sous-espace du domaine expérimental à l'intérieur duquel les objectifs ont été atteints en considérant l'incertitude des procédés concernés.

Cette directive poursuit en stipulant que travailler au sein du DS n'est pas considéré comme un changement. Par contre, sortir du DS est considéré comme un changement et doit normalement engager une procédure réglementaire d'approbation. Dès lors, si le fait de travailler dans le DS n'est pas considéré comme un changement, cela signifie que le DS correspond à une zone de robustesse. En effet, la robustesse d'une procédure ou d'une méthode analytique est définie dans la directive Q2(R1) de l'ICH [38] comme étant une mesure de sa capacité à rester inchangée face à des faibles variations délibérées des paramètres affectant la procédure ou la méthode analytique. La robustesse donne donc une indication de la fiabilité d'une méthode analytique dans les conditions normales d'utilisation. En conséquence, le DS, son étendue, sa forme et les paramètres le définissant, semblent bien être représentatifs de la robustesse d'une méthode analytique.

La directive Q8(R2) prône l'utilisation de la planification expérimentale pour identifier le DS. Les objectifs à atteindre sont définis avant le début du processus de développement. Ils peuvent être traduits en termes de critères ou d'attributs critiques de qualité pour quantifier l'accomplissement du processus de développement. Les attributs critiques de qualité peuvent être définis comme une propriété physique, chimique, biologique ou micro-biologique ou caractéristique devant être dans une limite, une gamme appropriée ou suivant une distribution donnée pour assurer la qualité désirée d'un produit. Ainsi, lors du développement de méthodes chromatographiques, un critère représentatif de la séparation des composés peut, par exemple, être la résolution entre 2 pics chromatographiques.

Il existe cependant une contradiction dans cette directive. En effet, la définition du DS fait intervenir la notion d'assurance de qualité qui repose sur la gestion des risques. Par contre, l'exemple mentionné en fin de document (exemple 1 de la section C de l'appendice 2 de la directive Q8(R2)) montre que le DS est l'espace multidimensionnel à l'intérieur duquel la réponse sélectionnée (un pourcentage de dissolution) atteint un seuil prédéfini (dissolution > 80%). Dans ce cas, la définition du DS va à l'encontre de l'exemple car aucune gestion du risque, ni aucune vérification de l'assurance de qualité n'ont été réalisées car seule une réponse moyenne est prédite. Il est dès lors très important de bien définir des méthodologies permettant de prendre en compte la gestion du risque, en termes d'assurance de qualité, et l'erreur de prédiction dans le processus de développement de méthodes chromatographiques.

I.2.3. Le concept de Quality Risk Management

Le concept de "*Quality Risk Management*" peut être traduit par "la gestion des risques de qualité". Dans la directive Q9 de l'ICH [39], ce concept est défini comme le processus systématique pour l'évaluation, le contrôle, la communication et l'examen des risques pour la qualité d'un principe actif ou d'un médicament à travers le cycle de vie du produit.

Dans le cas du développement de méthodes analytiques, la gestion des risques, parfois appelée assurance de la qualité, consiste en la prise en compte de l'erreur lors de la prédiction des attributs critiques de qualité. Le risque concerné correspond donc à celui de ne pas observer une valeur prédite ou de prédire la probabilité de ne pas atteindre un seuil sélectionné pour un attribut critique de qualité.

I.3. La chimiométrie appliquée au développement de méthodes chromatographiques

La chimiométrie peut être définie comme l'ensemble des méthodes mathématiques, essentiellement statistiques, permettant d'améliorer la compréhension de l'information obtenue à partir de données physiques, chimiques, pharmaceutiques, biologiques ou autres. D.L. Massart définit le chimiométrie comme une discipline chimique qui utilise les mathématiques, les statistiques et la logique formelle pour concevoir ou sélectionner des procédures expérimentales optimales, afin de fournir le maximum d'informations chimiques pertinentes par l'analyse des données chimiques et d'obtenir des connaissances sur les systèmes chimiques [40]. Cette science est actuellement en plein essor. Cela est notamment dû au fait que les techniques analytiques actuelles sont capables de générer des quantités considérables de données (techniques analytiques rapides, efficaces et multi-dimensionnelles) mais aussi de les traiter grâce à l'évolution de l'informatique. Par exemple, un chromatogramme de 5 min (ce qui est relativement court) enregistré en CL–SM dans une plage de rapport masse sur charge allant de 50 à 200 (intervalle de 0.1 entre les canaux) avec une fréquence d'acquisition de 20 Hz correspond à une matrice dont les dimensions sont 6000×1500. Il devient relativement compliqué pour un utilisateur même averti de traiter l'énorme quantité d'informations en un temps comparable à celui de l'enregistrement du chromatogramme, c'est-à-dire en 5 min. C'est pourquoi des techniques chimiométriques alliées à la puissance de l'informatique font actuellement l'objet de tant de recherches novatrices et prometteuses.

Par ailleurs, si les lois régissant un processus, en fonction de toute une série de variables, sont inconnues, il est intéressant d'identifier les équations mathématiques mimant le comportement de certaines réponses relatives à ce processus en fonction desdites variables. Dans cet objectif de modélisation du comportement d'une réponse en fonction de certains facteurs, la planification expérimentale est la seule stratégie raisonnablement possible.

I.3.1. Les plans d'expériences [41,42]

I.3.1.1. Objectifs de la planification expérimentale

La planification expérimentale revient à identifier un nombre minimal de conditions opératoires bien réparties dans un espace multidimensionnel afin d'extraire un maximum

d'informations concernant un processus défini par une ou plusieurs réponses. Grâce aux expériences réalisées, une réponse définissant le processus peut être modélisée par une équation mathématique. Ensuite, l'équation ainsi obtenue peut être utilisée pour prédire la réponse dans le domaine expérimental afin de maximiser ou minimiser cette dernière.

I.3.1.2. La stratégie de la planification expérimentale

En pratique, la stratégie de la planification expérimentale peut se décomposer en cinq étapes étroitement liées mais qui consistent à énoncer et formaliser le problème, à sélectionner des facteurs et un plan d'expériences, à réaliser les expériences, à modéliser la réponse sélectionnée et finalement à valider le modèle mathématique.

Premièrement, il convient d'énoncer et de formaliser le problème. Cela correspond à identifier la ou les réponses correspondant à un objectif initial. La sélection de la réponse qui est modélisée est intimement liée à la sélection des facteurs qui permettront de créer le plan d'expériences. Si le nombre de facteurs est élevé (> 4), il convient de réaliser un plan de criblage permettant la classification des facteurs en fonction de leur impact sur la réponse. Les facteurs ayant le plus d'effet peuvent dès lors être sélectionnés en vue d'une optimisation ultérieure du processus étudié. Au contraire, si le nombre de facteurs est égal ou inférieur à 4, un plan d'optimisation peut directement être réalisé. Les plans d'optimisation permettent de modéliser la réponse en fonction desdits facteurs afin de maximiser ou minimiser cette dernière ou des attributs critiques de qualité qui dépendent de cette réponse. C'est également à cette étape qu'il est nécessaire d'évaluer le degré de l'équation mathématique (linéaire, quadratique, cubique, etc.) pour connaitre le nombre de niveaux devant être utilisé pour chaque facteur. Si la réponse varie de manière quadratique en fonction d'un facteur, un nombre minimal de 3 niveaux doit alors être sélectionné. L'équation polynomiale servant à la modélisation peut ensuite être formalisée. Cette étape de sélection des facteurs et de leurs niveaux est étroitement liée avec l'étape suivante consistant à choisir un plan d'expériences.

Deuxièmement, un plan d'expériences doit être sélectionné. De nombreux plans sont à la disposition de l'utilisateur et ont chacun leurs avantages. Les plans d'expériences les plus faciles à appréhender sont les plans factoriels complets. Ils définissent les conditions opératoires en combinant tous les niveaux de chaque facteur. Un plan factoriel complet à 3 facteurs ayant chacun 3 niveaux définira donc un nombre de conditions opératoires égal à 3^3 = 27. Afin d'estimer, la variabilité du processus, le point central du plan doit être répété au minimum 2 fois, soit 3 mesures au point central. La sélection d'un plan d'expériences performant se base également sur le nombre de degré de liberté lors de la modélisation de la réponse. Les plans factoriels fractionnaires résultent d'un fractionnement de plans factoriels complets. Le sous-ensemble de conditions expérimentales est sélectionné sur la base d'une évaluation (ou d'une hypothèse) des facteurs et des interactions qui ont les effets les plus importants. Une fois cette sélection effectuée, la planification expérimentale

doit séparer ces effets. En particulier, des effets significatifs ne doivent pas être confondus. En d'autres termes, la mesure de l'un ne devrait pas dépendre de la mesure de l'autre. Les plans de Plackett-Burman représentent une sous-classe des plans factoriels fractionnaires. Ils sont généralement utilisés comme plan de criblage. Les plans centraux composite et de Box-Behnken sont quant à eux plutôt utiliser comme plan d'optimisation.
Certains plans peuvent être générés par ordinateur afin d'optimiser certains paramètres statistiques. Les plus connus sont les plans D-optimaux où l'on cherche à minimiser la covariance des paramètres estimés d'un modèle donné. Par exemple, pour un plan factoriel complet à deux facteurs (X_1 et X_2) comportant 3 niveaux chacun, une réponse *Y* peut être modélisée par l'équation 11.

$$Y = \beta_0 + \beta_1 \times X_1 + \beta_2 \times X_1^2 + \beta_3 \times X_2 + \beta_4 \times X_2^2 + \beta_5 \times X_1 X_2 + \varepsilon \qquad (11)$$

où β_0… β_5 sont les paramètres du modèles et ε représente l'erreur.

Troisièmement, les expériences sont réalisées et les réponses mesurées avec la plus grande précision possible. L'expérimentateur veille de la sorte à minimiser la variabilité expérimentale et à garantir la meilleure répétabilité possible des résultats.

Quatrièmement, les données obtenues sont modélisées par l'équation polynomiale précédemment choisie. Cette modélisation des données est réalisée par une régression linéaire multiple qui permet l'estimation des paramètres du modèle par la méthode des moindres carrés [43]. La modélisation peut également être réalisée par un processus de régression pas à pas (*stepwise regression*). Cela correspond à éliminer les termes du modèle n'améliorant pas un critère statistique choisi comme, par exemple, le coefficient de détermination ajusté ou le critère d'information d'Akaike [44]. Le modèle obtenu permet la prédiction de la réponse en tout point du domaine expérimental. Ces prédictions sont utilisées pour identifier la condition opératoire optimale permettant d'atteindre les objectifs préalablement fixés, c'est-à-dire la maximisation de la réponse. A cette étape, il est également important de vérifier si le modèle préalablement sélectionné ne sur-ajuste pas les données. Dans un cas de sur-ajustement (*overfitting*), le modèle s'adapte aux phénomènes aléatoires qui peuvent provenir du bruit et/ou de la variabilité de la méthode. Le modèle n'est alors plus représentatif de la variation de la réponse moyenne uniquement. Pour pallier ce problème, il convient d'augmenter le nombre d'expériences et/ou de répétitions ou de simplifier le modèle en éliminant des effets ou des interactions des plus hauts degrés. A l'inverse, il est nécessaire de vérifier si le modèle ne souffre pas d'un manque d'ajustement (*lack of fit*). Dans ce cas, certains effets sont manquants et le modèle ne s'ajuste pas suffisamment à la réponse observée. Des tests statistiques comparant notamment la variance prédite par le modèle et la variance observée expérimentalement permettent d'identifier les cas de manque d'ajustement [45]. L'ajout de termes ou d'interactions de plus hauts degrés permet généralement de résoudre ce problème.

Cinquièmement, il convient d'examiner l'aptitude à la prédiction du modèle en vérifiant l'accord entre observation et prédiction à la condition opératoire optimale. La

validité du modèle peut également être confirmée par la réalisation d'expériences supplémentaires à des niveaux intermédiaires du domaine expérimental.

I.3.1.3. Application à la chromatographie liquide

La première étape pour pouvoir appliquer la méthodologie des plans d'expériences en CL est de sélectionner une réponse modélisable par une équation linéaire multiple. Dès lors, la modélisation d'un critère variant de manière non linéaire, comme la résolution, n'est pas souhaitable. En effet, lors d'une inversion de l'ordre d'élution des pics, la résolution varie de manière discontinue et n'est pas donc pas correctement modélisée par un polynôme quel que soit son degré. Par contre, le temps de rétention, et plus particulièrement le logarithme du facteur de rétention suit une variation linéaire quadratique ou de degré supérieur en fonction des paramètres chromatographiques courants. Il peut donc être modélisé par des équations linéaires multiples [36]. Notons toutefois que la résolution peut servir de critère, pouvant être calculée à partir des logarithmes des facteurs de rétention et des logarithmes de demi-largeurs des pics de la paire critique pour représenter la qualité d'une séparation.

I.3.2. Analyse en composantes indépendantes

L'analyse en composantes indépendantes (ACI) fait partie des techniques d'analyse de données généralement utilisée en chimiométrie comme l'analyse en composantes principales (ACP), les méthodes de classification hiérarchique (ACH) ou l'analyse discriminante. Introduite par Herault et Jutten en 1986 [46], l'ACI a été formalisée par P. Comon en 1994 [47]. Elle se base sur l'hypothèse qu'un signal est la résultante d'un mélange de plusieurs signaux inconnus (également appelés sources).

L'ACI est utilisée pour estimer les sources initiales en maximisant la non-gaussianité (basée sur le kurtosis et la néguentropie [48,49]) ou en minimisant l'information mutuelle [48,49]. L'ACI est également appelée technique de séparation de sources en aveugle car les sources initiales ne sont pas connues. Un exemple d'analyse de signal par ACI (utilisant l'algorithme fastICA [50]) est présenté à la figure I.4.

I.3.2.1. Comparaison entre l'ACI et l'ACP

L'ACP est utilisée pour décomposer un ensemble de données multidimensionnelles en un ensemble de composantes orthogonales qui maximisent successivement la variance. Au contraire, les composantes obtenues par ACI ne sont pas nécessairement orthogonales mais maximisent l'indépendance des signaux. La figure I.5 présente un exemple d'application d'ACI et d'ACP à un jeu de données bidimensionnel.

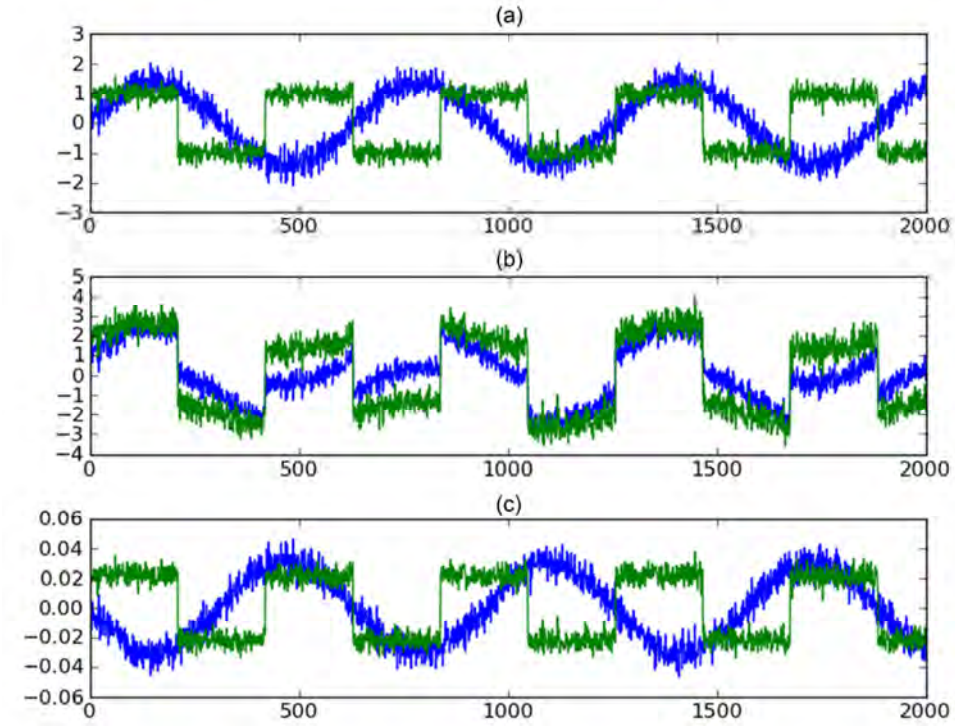


Fig. I.4 Exemple d'analyse de signal par ICA. (a) signaux/sources originaux inconnus, (b) signaux enregistrés (correspondant à un mélange des signaux en (a)) et utilisés par appliquer l'algorithme d'ACI et (c) signaux estimés par ACI.

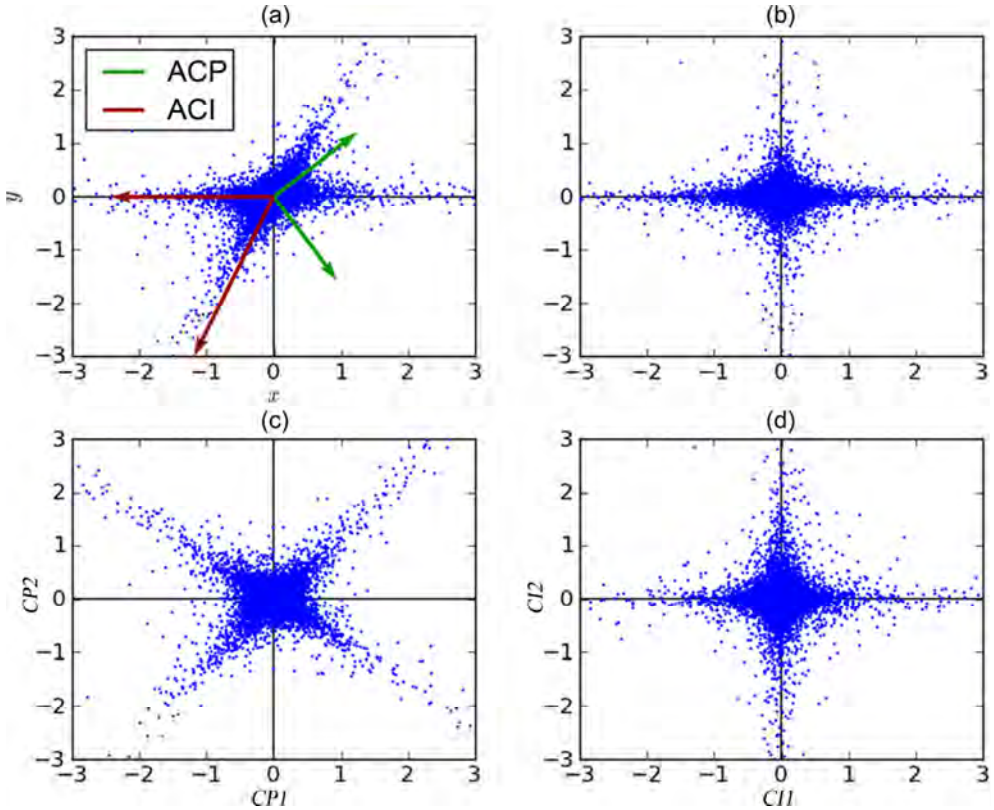


Fig. I.5 Exemple de comparaison entre l'ACI et l'ACP pour un jeu de données bidimensionnel. (a) jeu de données original. Les composantes principales sont représentées en vert et les composantes indépendantes en rouge foncé. (b) vraies sources indépendantes (inconnues), (c) distribution des données dans l'espace des composantes principales (CP1 et CP2), (d) distribution des données dans l'espace des composantes indépendantes estimées (CI1 et CI2).

I.3.3. *Design Space* et prédiction

Lorsqu'un modèle mathématique a été établi, il est très facile de prédire une valeur de la réponse décrite par le modèle. Cependant, lors de cette prédiction, l'erreur de prédiction n'est pas prise en compte. Il est donc nécessaire de se baser sur l'erreur estimée par le modèle, c'est-à-dire l'écart entre l'observation et la prédiction, pour pouvoir estimer la distribution de la valeur prédite. Certains intervalles statistiques peuvent être utilisés comme intervalles de prédiction [51].

I.3.3.1. Intervalles de confiance

L'intervalle qui est probablement le plus connu est l'intervalle de confiance de la moyenne d'une distribution normale. Cet intervalle, à un certain degré de confiance, se base uniquement sur l'estimateur de la moyenne arithmétique (Eq. 12) et de l'écart type (Eq. 13) d'un échantillon aléatoire de données. L'intervalle de confiance bilatéral à 95% est défini à l'équation 14. Il établit que dans 95% des cas, la fourchette définie par cet intervalle recouvre μ, la moyenne (inconnue) de la population.

$$\bar{x} = \frac{1}{n} \times \sum_{i=1}^{n} x_i \tag{12}$$

$$s = \sqrt{\frac{1}{n-1} \times \sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{13}$$

$$IC\ 95\% : \bar{x} - 1{,}96\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + 1{,}96\frac{s}{\sqrt{n}} \tag{14}$$

Où $\bar{x}$ est l'estimation de la moyenne d'un échantillon aléatoire d'effectif *n* dont les éléments sont x_i avec i=1,…, n. s est l'estimation de l'écart type.

Il est donc manifeste que l'intervalle de confiance n'est pas approprié pour estimer dans quelle fourchette se trouveront les futurs résultats et ne peut en aucun cas être utilisé comme intervalle de prédiction.

I.3.3.2. Intervalles de tolérance

L'intervalle de tolérance représente une fourchette à l'intérieur de laquelle une certaine proportion des résultats est attendue (β-*expectation tolerance interval* [51-53] également appelé intervalle de prédiction) ou observée avec un certain niveau de confiance (β-*content,* γ-*confidence tolerance interval* [51,52]). Lorsque *n* augmente, l'intervalle de confiance tend vers 0 alors que l'intervalle de tolérance tend vers une valeur dépendante des résultats. Pour un intervalle de prédiction bilatéral, les limites inférieure et supérieure (respectivement *L* et *U*) sont définies telles que l'équation 15 soit respectée. β représente la

proportion moyenne de valeur observée dans cet intervalle. En d'autres termes, si β = 0,95, cela signifie qu'en moyenne, 95% des résultats seront inclus dans l'intervalle [*L*;*U*] [51].

$$E(P[L \leq x_i \leq U]) = \beta \quad (15)$$

Où E et P représentent respectivement les opérateurs d'estimation de l'espérance mathématique et de la probabilité.

I.3.3.3. Propagation de l'erreur prédictive

Lorsqu'une prédiction est réalisée par un modèle multivarié, l'erreur de prédiction (ou incertitude de prédiction) doit être représentative de la variance des résultats utilisés pour créer ce modèle. Une manière d'obtenir une estimation de l'erreur de prédiction est de se baser sur une distribution multivariée de Student créée à partir de la matrice de variance-covariance des paramètres du modèle. Cette distribution multivariée est représentative de l'erreur d'estimation des paramètres du modèle. En utilisant des simulations de Monte Carlo [54] pour tirer un échantillon de cette distribution, il est finalement possible d'estimer la distribution de la valeur prédite. De manière simplifiée, l'erreur provenant de l'incertitude de mesure et de l'ajustement non parfait du modèle mathématique aux données se reflète dans la distribution des résidus qui représentent l'écart entre observations et prédictions. En d'autres termes, l'écart type des résidus est proportionnel à l'erreur de prédiction. Dès lors, il est possible d'estimer la distribution de la valeur prédite à partir d'une loi normale dont l'écart type est égal à celui mesuré sur les résidus après avoir vérifié la distribution normale de ces derniers par un test de Shapiro-Wilk [55,56]. La loi normale résultante permet donc d'estimer la distribution prédictive des réponses. Ensuite, elle peut être utilisée pour déterminer la distribution prédictive des critères d'intérêt. En effet, ce n'est pas nécessairement la réponse du modèle qui est l'objectif final mais un critère obtenu à partir de ses réponses.

En CL, c'est un critère comme la résolution qui est optimisé. Celui-ci est différent de la réponse directement modélisée. Dans ces conditions, il est fort important de connaitre l'erreur associée afin de donner de la crédibilité à une prédiction telle que $R_{S,\text{prédit}} > 1{,}5$. La méthode de Monte Carlo utilise un nombre fini de calculs (simulations) pour estimer une valeur ainsi que sa distribution (dans notre cas la distribution du critère en n'importe quel point du domaine expérimental). Un échantillon fini (un vecteur de 2500 points) est tiré de la distribution gaussienne créée grâce aux résidus et le calcul de prédiction est réalisé pour chaque point du vecteur. Finalement, l'erreur de prédiction affectant la réponse est propagée jusqu'au critère afin d'obtenir l'erreur de prédiction se rapportant au critère (la distribution du critère prédit) et permet de donner de la crédibilité à une prédiction donnée. Identiquement à la philosophie appliquée pour obtenir les intervalles de prédiction (cf. section I.3.3.2.), la prédiction ne sera plus formalisée comme telle : $R_{S,prédit} \geq 1{,}5$ mais plutôt sous la forme $E\left[P\left(R_{S,prédit} \geq 1{,}5\right)\right] \geq \pi$ où π représente le niveau de qualité, par exemple égal à 95%. A l'intérieur du domaine expérimental, la région vérifiant cette relation

définit une zone dans laquelle le risque de ne pas atteindre le seuil du critère (1,5) est étudié et minimal (par exemple = 5% si π = 95%) et peut en conséquence correspondre au DS tel que défini précédemment (cf. section I.2.2.).

Bibliographie

[1] M. Tswett, M. Ber. Dtsch. Chem. Ges. 24 (1906) 316.
[2] A.J.P. Martin, R.L.M. Synge, Biochem. J. 35 (1941) 1358.
[3] Gerra L. Bosco, TRAC–Trend. Anal. Chem. 29 (2010) 781.
[4] SDi Global Assessment Report, Edition 10.5 : http://www.strategic-directions.com/apps/reports/index.cfm?action=detail&sdi=1203697703
[5] HPLC systems & accessories, A Global Strategic Business Report, http://www.strategyr.com/HPLC_Systems_and_Accessories_Market_Report.asp
[6] J.H. Knox, P.A. Bristow, Chromatographia 10 (1977) 279.
[7] D.Lubda, K. Cabrera, W. Kraas, C. Schaefer, D. Cunningham, LCGC 19 (2001) 1186.
[8] D. Guillarme, D.T.T. Nguyen, S. Rudaz, J.L. Veuthey, J. Chromatogr. A 1149 (2007) 20.
[9] H. Chen, C. Horvath, J. Chromatogr. A 705 (1995) 3.
[10] J. Li, Y. Hu, P.W. Carr, Anal. Chem. 69 (1997) 3884.
[11] I.D. Wilson, Chromatographia 52 (2000) S28.
[12] S. Heinisch, G. Puy, M.P. Barrioulet, J.L. Rocca, J. Chromatogr. A 1118 (2006) 234.
[13] D.T.T. Nguyen, D. Guillarme, S. Rudaz, J.L. Veuthey, J. Chromatogr. A 1128 (2006) 105.
[14] D.T.T. Nguyen, D. Guillarme, S. Rudaz, J.L. Veuthey, J. Sep. Sci. 29 (2006) 1836.
[15] D.T.T. Nguyen, D. Guillarme, S. Heinish, M.P. Barrioulet, J.L. Rocca, S. Rudaz, J.L. Veuthey, J. Chromatogr. A 1167 (2007) 76.
[16] http://www.molnar-institute.com/
[17] http://www.chromsword.de/
[18] http://acdlabs.eu/products/com_iden/meth_dev/lc_sim/
[19] http://www.datalys.net/index.htm
[20] Cs. Horváth, W. Melander, I. Molnár, J. Chromatogr. 125 (1976) 129.
[21] Cs. Horváth, W. Melander, I. Molnár, Anal. Chem. 49 (1977) 142.
[22] Cs. Horváth, W. Melander, I. Molnár, P. Molnár, Anal.Chem. 49 (1977) 2295.
[23] S. Heron, A. Tchapla, J. Chromatogr. A 556 (1991) 219.
[24] P.W. Carr, J. Li, A.J. Dallas, D.I. Eikens and L. Choo Tan, J. Chromatogr. A 656 (1993) 113.
[25] S.V. Galushko, I.P. Shishkina, J. Chromatogr. A 445 (1988) 59.
[26] L.R. Snyder, High-performance liquid chromatography. Advances and perspectives, Vol. 1, Cs. Horvath Ed., Academic press, New York, 1980.
[27] P. Jandera, J. Churacek, Gradient elution in column liquid chromatography, Elsevier, Amsterdam, 1985.
[28] L.R. Snyder, J.W. Dolan, Adv. Chromatogr. 38 (1998) 115.
[29] P. Jandera, Adv. Chromatogr. 43 (2004) 1.

[30] S.V. Galushko, A.A. Kamenchuk, G.L. Pit, Am. Lab. 27 (1995) 33G
[31] V. Concha-Herrera, G. Vivo-Truyols, J.R. Torres-Lapiaso, M.C. Garcia-Alvarez-Coque, J. Chromatogr. A 1063 (2005) 79.
[32] J.W. Dolan, L.R. Snyder, N.M. Djordjevic, D.W. Hill, L. Van Heukelem, T.J. Waeghe, J. Chromatogr. A 857 (1999) 21.
[33] B. Dejaegher, Y. Vander Heyden, J. Chromatogr. A 1158 (2007) 138.
[34] R. Ragonese, M. Mulholland, J. Kalman, J. Chromatogr. A 870 (2000) 45.
[35] E. Hund, Y. Vander Heyden, M. Haustein, D. L. Massart, J. Smeyers-Verbeke, Anal. Chim. Acta 404 (2000) 257.
[36] W. Dewé, R.D. Marini, P. Chiap, Ph. Hubert, J. Crommen, B. Boulanger, Chemom. Intell. Lab. Syst. 74 (2004) 263. http://hdl.handle.net/2268/6036
[37] Guidelines of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), Pharmaceutical development, Q8(R2) Step 4, Geneva, 2009.
[38] Guidelines of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), Pharmaceutical development, Q2(R1) Step 4, Geneva, 2005.
[39] Guidelines of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), Quality Risk Management, Q9 Step 4, Geneva, 2005.
[40] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P. Lewi and J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics: Part A. Volume 20A in Data Handling in Science and Technology, Elsevier, Amsterdam, 1997.
[41] G. Cox, W. Cochran, Experimental Designs, 2nd Ed., McGraw-Hill, New York, 1962.
[42] H. Martens, M. Martens, Experimental Designs, 2nd Ed., Wiley, New York, 2001.
[43] N.R. Draper, H. Smith, Applied regression analysis, 3rd Ed., Wiley, New York, 1998.
[44] H. Akaike, IEEE Trans. Automat. Control 19 (1974) 716.
[45] J. Neter, M.H. Kutner, C.J. Nachstheim, W. Wasserman, Applied Linear Statistical Models, 4th Ed., Irwin, New York, 1996.
[46] J. Herault, C. Jutten, Space or time adaptive signal processing by neural models, in Proceedings AIP Conference: Neural Networks for Computing, J.S. Denker, Ed. American Institute for Physics, 151 (1986) 206.
[47] P. Comon, Signal Process. 36 (1994) 287.
[48] A. Hyvärinen, IEEE Trans. Neural Networks 10 (1999) 626.
[49] A. Hyvärinen, E. Oja, Neural Comput. 9 (1999) 1483.
[50] A. Hyvärinen and E. Oja, Neural Networks, 13 (2000) 411.
[51] G.J. Hahn, W.Q. Meeker, Statistical Intervals — a guide for practitioners, Wiley, New York, 1991.
[52] I. Guttman, Statistical Tolerance Regions: Classical and Bayesian, Hafner, Darian, 1969.
[53] S.S. Wilks, Ann. Math. Statist. 12 (1941) 91.
[54] M.A. Herrador, A.G. Asuero, A.G. Gonzalez, Chemom. Intell. Lab. Syst. 79 (2005) 115.

[55] P. Royston, J. R. Stat. Soc. C Appl. 31 (1982) 115.
[56] P. Royston, J. R. Stat. Soc. C Appl. 31 (1982) 176.

Chapitre II

Objectifs

La chromatographie liquide est à l'heure actuelle la technique analytique la plus répandue dans l'industrie pharmaceutique et dans bien d'autres domaines. Les méthodologies modernes de développement de méthodes chromatographiques font désormais appel à des logiciels informatiques commerciaux basés sur la théorie de la chromatographie liquide. Ces méthodologies permettent de prédire et d'optimiser des séparations chromatographiques. Cependant, dans les logiciels actuellement disponibles, aucune estimation de l'erreur de prédiction n'est réalisée.

Par ailleurs, la tendance actuelle en sciences analytiques est d'estimer le risque de ne pas atteindre les objectifs initialement fixés. C'est ainsi que, par exemple, la validation des méthodes analytiques évolue de plus en plus vers la possibilité d'évaluer la proportion de futurs résultats se trouvant dans les limites d'acceptation.

Dans le cadre du développement pharmaceutique et en particulier dans le domaine du développement de méthodes chromatographiques, le pendant de la validation est formalisé par le concept de "*quality by design*" que l'on pourrait traduire par "qualité par la conception".

Dans ce travail, nous nous attacherons tout d'abord à identifier et à établir des méthodologies qui permettront de développer les méthodes chromatographiques. Il conviendra de pouvoir optimiser les séparations chromatographiques tout en estimant la qualité du processus de développement et des prédictions réalisées par ce dernier. Les plans d'expériences seront utilisés dans ce but.

Parallèlement à cette étape, nous tenterons de développer une seconde méthodologie basée sur l'analyse en composantes indépendantes, qui permettra la séparation numérique de composés coélués afin de déterminer les temps au début et à la fin de pics chromatographiques correspondants. Cette méthodologie devrait donc permettre d'accélérer le processus de lecture automatique des chromatogrammes. Nous évaluerons ainsi sa capacité à estimer le plus précisément possible les limites d'intégrations des pics chromatographiques correspondant aux composés coéluants.

Ensuite, nous combinerons ces deux approches pour disposer d'une méthodologie globale permettant le développement rapide de méthodes chromatographiques. Nous évaluerons alors la capacité de cette méthodologie globale à fournir des prédictions précises et à permettre le développement de méthodes chromatographiques robustes.

Au terme de l'étape précédente, nous envisagerons l'interprétation des paramètres du modèle en fonction des facteurs définissant les conditions expérimentales (pH, température, etc.) afin de mettre en évidence la relation existant entre les comportements prédits et ceux décrits par la théorie de la chromatographie.

Enfin, dans la dernière partie de ce travail, nous essayerons d'éprouver la méthodologie développée en l'appliquant au développement de méthodes chromatographiques de mélanges complexes. Nous tenterons ainsi d'identifier les conditions chromatographiques optimales adaptées à la séparation des composés présents dans des extraits de plantes, au dosage de substances d'intérêt dans une formulation pharmaceutique et à la détermination d'un principe actif et de ses substances apparentées.

Chapitre III

Résultats et discussions

Partie III.1

Développement de la méthodologie

Section III.1.1

Fondement de la méthodologie DoE–DS

P. Lebrun, B. Govaerts, B. Debrus, A. Ceccato, G. Caliaro, Ph. Hubert, B. Boulanger, Development of a new modelling technique to find with confidence equivalence zone and design space of chromatographic analytic methods, Chemom. Intell. Lab. Syst. 91 (2008) 4.

Résumé

Une nouvelle méthode pour la modélisation de réponses chromatographiques est présentée en tant que pièce maitresse pour la réalisation du développement automatisé de méthodes analytiques. Cette méthodologie est basée sur quatre parties. Premièrement, nous proposons d'utiliser un petit ensemble d'équations statistiques pour créer les modèles prédictifs pour modéliser des réponses basées sur le temps de rétention, la largeur et l'asymétrie des pics. Deuxièmement, un plan d'expériences est créé pour réaliser les expériences. Troisièmement, en utilisant la méthode de "quadrillage" (*grid search*) sur le domaine expérimental, des décisions multicritères peuvent être prises en combinant des critères locaux ou globaux à l'aide de fonctions de désirabilité. Cela permet de trouver des chromatogrammes optimaux. Quatrièmement, il est primordial d'investiguer comment l'erreur de prédiction des modèles affecte la réponse et se propage à la condition optimale. Cette étape permet de donner de la confiance, de la crédibilité, à la condition optimale obtenue, en identifiant la région qui donne des solutions acceptables. "L'espace de conception" (*the design space*) peut être obtenu de cette manière. Cette approche est appliquée à un cas réel. Les prédictions des modèles, à la condition expérimentale optimale, ont été validées grâce à la réalisation d'expériences supplémentaires. Finalement, la méthodologie proposée présente l'avantage considérable d'être flexible et de pouvoir s'ajuster aux demandes des analystes.

Summary

A new method for modelling chromatographic responses is presented as a critical piece for the achievement of automated development of analytical methods. This methodology is based on four parts. First, we propose to use a very little set of statistical equations to create predictive models for retention time-based responses as the apex, the width and the asymmetry of peaks. Second, an experimental design is set up to realize experiments. Third, using grid search over the domain, multi-criteria decision is taken with respect to different local or global optimization criteria, used as desirability functions. This allows finding an optimal chromatogram. Fourth, we advise to investigate how the predictive error of the models propagates around optimal solution. This allows giving confidence in the optimal solution, in finding a set of zones that presumably will give an acceptable solution. Design spaces can be derived with a similar technique. The approach is exemplified with a real case and predictions of models at optimal analytical conditions are validated through new experiments. Flexibility is left over all the presented methodology.

1. Introduction

In analytical chemistry, the chromatographic techniques are widely used in different fields of activity such as chemical, pharmaceutical, biomedical, environmental and food analysis. Thus, the selection of the most appropriate experimental conditions allowing the separation of compounds of interest in various matrices is a matter of a very particular interest. Pharmaceutical industries are of course concerned by these problems and are more especially interested by all new approach allowing to separate their compounds properly and quickly in order to quantify them. Indeed, the analytical step is a crucial phase during the development of new drugs since the different decisions are taken based on results generated by one or more analytical methods. Amongst the chromatographic techniques, Liquid Chromatography (LC) is probably the most common technique to fulfil this objective.

Nowadays, the development of analytical methods in LC is still time consuming and not always under the perfect control of analysts. This is due to the fact that most of the parameters to manage to obtain acceptable separation conditions have complex effects on the chromatogram. The problem becomes even more complicated when the matrix is complex and contains many compounds with physico-chemical properties that are not necessarily known. A lot of contributions from various authors about optimization and methods development have been reported over the last 20 years. The use of design of experiments to find optimal conditions is now largely accepted in chromatography. However, it is often problematic to take into account the total complexity of the problems and it is probably the reason why actual softwares in chromatography still don't allow the optimization with many parameters.

Schoenmakers [1] reported formulations to compute global or limited optimization criteria used for optimization. Massart et al. [3] and Snyder [2] illustrated different tools for methods development. Vanbel [4] summarized the need of adequate and flexible optimization criteria, adapted to practical situations, and the need to provide various and flexible experimental designs and modelling equations. Dewé et al. [5] also proposed a methodology to optimize several analytical conditions. Both last references agree on the fact that modelling directly chromatographic criteria to find some optimum configuration is not good practice due to discontinuities in modelled responses. It is proposed to model retention times as characteristic of chromatograms instead of criteria. After, these criteria are derived from the retention times and a multi-criteria optimization can be envisaged. The main problem encountered in these methodologies is the assumed independence between responses. This leads to many equations, which is not recommendable, and errors can occur (see later). Finally, Vanbel and Dewé et al. completely ignore error of the models in their predicted optima. In this paper, these methodologies are generalized, allowing to optimize in an automated way as many parameters and criteria as needed by the analysis.

Changing parameters of the LC leads to different retention times for one analyte in the output chromatograms, as shown in Fig. 1. The purpose of this work is that, following the practice of design of experiments [6,7], it is possible to find some optimal parameter

configurations, considering different criteria computed on chromatograms, such as the minimal separation, the total retention time or the asymmetry of the peaks, for instance.

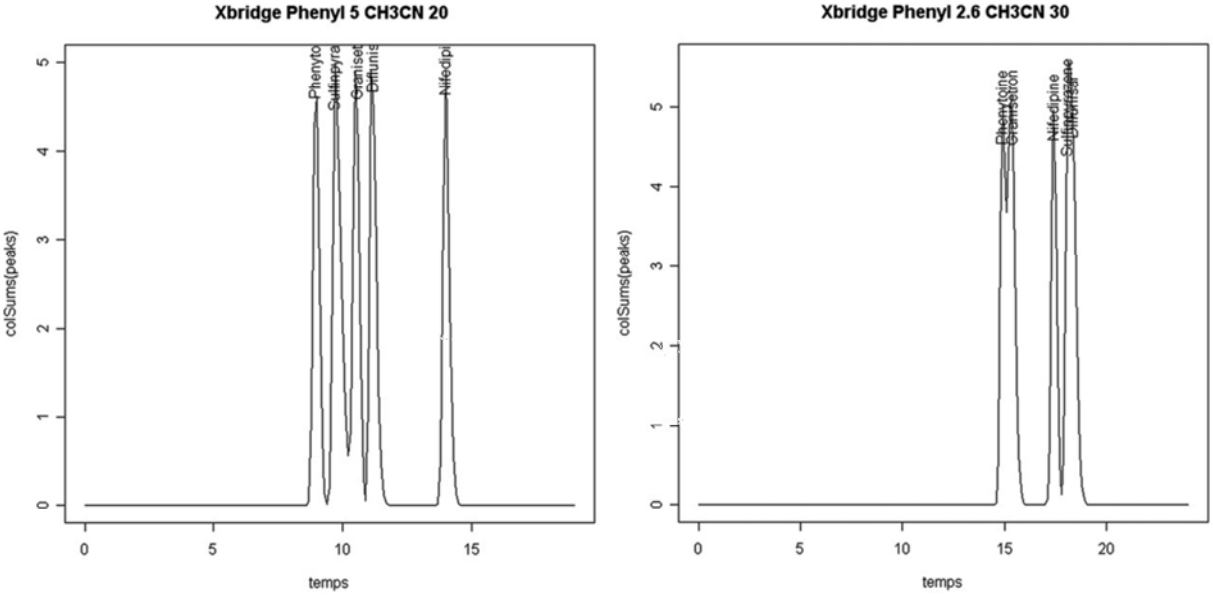


Fig. 1 Example: different analytical conditions lead to different retention times for the analytes (reconstructed chromatograms, method from Dewé et al. [5]). Chromatographic parameters are the nature of the stationary phase, the mobile phase pH, the nature of the organic modifiers and the gradient slopes (min) used in analysis. Left: column XBridge Phenyl with acetonitrile in the buffer, pH 5 and, gradient 5%–95% of organic modifier realized in 20 min. Right: same column and solvent, pH 2.6, and gradient in 30 min.

2. Objectives and main steps of the methodology

We present a methodology to predict, with a known level of confidence, the best tuning parameters of a HPLC in the range of possible analytical conditions in order to get the best chromatogram possible over the domain of potential mixtures. This global target can be subdivided into several smaller objectives.

The first objective aims at developing, from the results of experiments designed for this very purpose, predictive models of the peak retention times with respect to the HPLC tuning parameters. The second objective is to use these predictive models in order to find optimal chromatographic analytical conditions with respect to different criteria to characterize the quality of a chromatogram. This is a multi-criteria optimization problem. The third objective aims at investigating the confidence of the predicted optimal conditions through a Monte-Carlo study of the propagation of the models predictive errors into the optimized criteria. Finally, this methodology allows identifying a design space if existing in the explored domain [8].

Fig. 2 presents the main steps of our methodology. First, a designed experiment is set up in order to explore as best as possible the ranges of possible HPLC parameters. These experimental conditions are applied to a given complex mixture and supply related chromatograms. Second, for each chromatogram, the retention time of each compound is

identified from the observed peaks. Third, statistical models are developed to predict the retention times of each mixture from the analytical conditions. Models have to be precise enough in order to have good predictive capabilities. This includes an adequate choice of the responses to be modelled, a variable selection step and the application of adequate statistical techniques to quantify the quality of the fit and validate the models. Then, criteria are defined to characterize what is a “good” chromatogram and predictive models are used to predict expected chromatograms and related criteria in the domain of explored analytical conditions. Derringer desirability functions and index are proposed to summarize the criteria to be optimized. Finally, optimal conditions are searched out of these predictions and their accuracy are analysed by Monte-Carlo propagation of the model prediction errors into the calculated criteria. The same technique can be applied on each point of the experimental domain to find a design space, if it exists.

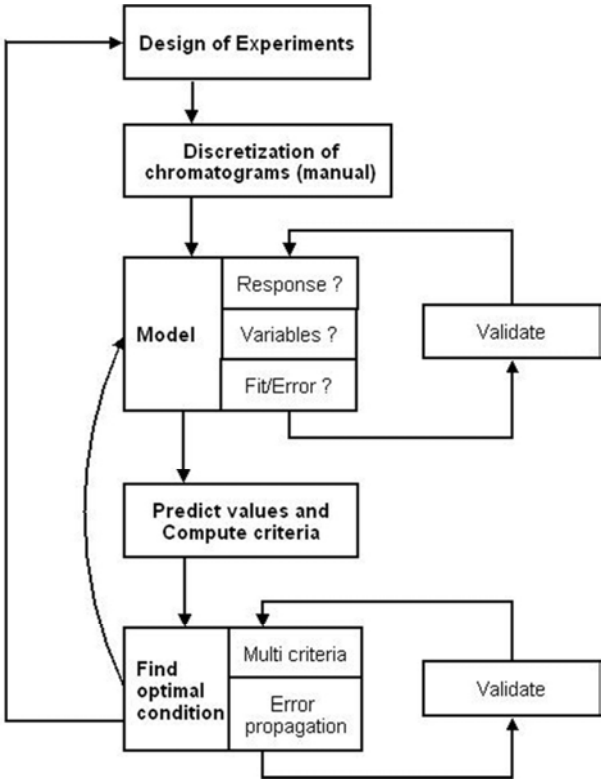


Fig. 2 Main steps of the presented methodology

Note that height of peaks of chromatograms has a limited interest for optimizing separation and is not concerned in this study. Dewé et al. [5] expose some formulations to reconstruct the height of peaks from their estimated area. It can easily be included in our methodology.

3. Details of the methodology

3.1. Design of experiments

Let's define F continuous or discrete tuning parameters of interest ($\boldsymbol{x}_1$,…, $\boldsymbol{x}_f$,…, $\boldsymbol{x}_F$) for the HPLC device under study. Each continuous factor (e.g. pH) is defined over a domain of interest [L_f, U_f] and each discrete factor (e.g. column) by a set of n_f levels. Let's note χ the experimental domain of $\boldsymbol{x}$'s.

Design of experiments and response surface methodology provide several methods to explore such domain χ according to the anticipated complexity of the factor effects on the responses (the peaks positions) [9]. Full factorial designs may be appropriate when the experiments are robotized. Central composite or D-Optimal designs are also well adapted in this context. Let's note $\boldsymbol{X}$ the resulting (N×F) design matrix.

The N experiments consist of applying each design factor setting $\boldsymbol{x}_i$=($\boldsymbol{x}_{i1}$,…, $\boldsymbol{x}_{iF}$) to a chosen (complex) mixture of M compounds $\boldsymbol{c}$=(c_1,…, c_j,…, c_M) and provide N chromatograms with, for each of them, M more or less separated peaks. Factors and levels to be used have to be identified by the analyst depending on the nature of the analytical method and the compounds to be separated.

3.2. Chromatogram discretization

Fig. 3 illustrates the discretization process of the chromatograms. This is the main manual part of the methodology although current studies show that an automation of this process is possible [10,11]. From each chromatogram, three retention times can be extracted: the retention times at the beginning, at the apex and at the end of the peaks at baseline-height (B, A and E, respectively). T_0 denotes the dead time of the system, associated to the analytical column. The retention times for the N chromatograms and the M components can be stored in three vectors B, A and E of size (N·M). These retention times are ordered as N blocks of M components in these vectors.

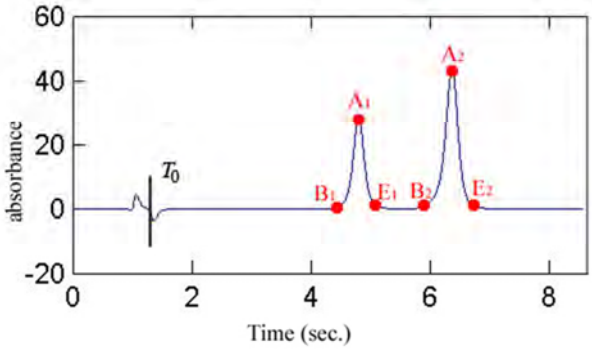


Fig. 3 Raw chromatogram with the positions of discretized points

Currently, it is still the task of the analyst to identify manually peaks in chromatograms. Patterns can be found in the DAD chromatogram, looking at the absorbance values for a specific time (1-D spectrum). For one peak corresponds one pattern. Difficulties of

identifications can arise, e.g. when there are co-eluted peaks or when there are impurities, which possess nearly the same absorbance spectrum than relevant analytes.

3.3. Models responses definition

The first step in the development of predictive models of the retention times with respect to HPLC analytical conditions, is to decide which responses to model. Dewé et al. [5] propose to create separate models for each retention time *B*, *A* and *E* and for each compound. Unfortunately, with this approach, uncertainty of the prediction of the models can lead to inversion in the predicted positions within a peak. For instance, it is undesirable that the predicted apex of a peak has a smaller retention time than its beginning, which is physically impossible. Moreover, there is little interest to try to make one model for each retention time vector because pairwise correlations between them are very close to 1. Modelling *B*, *A* and *E* leads to model three times non–independent information. Deriving separate models for each retention time vector and each compound can also be heavy to handle in the subsequent steps of the methodology.

This paper proposes two enhancements to this classical approach. First, it suggests to build, for each response of interest, one single global model involving all *F* experimental factors and all *M* compounds. In this framework, it is suited to transform retention times *B*, *A* and *E* in retention factors *k'*. They are computed as follows:

$$k'_A = \frac{A - T_0}{T_0}, k'_B = \frac{B - T_0}{T_0}, k'_E = \frac{E - T_0}{T_0}$$

Notice that it is common to work with the logarithmic form of the *k*'.

Second, it suggests to transform the three original responses k'_B, k'_A and k'_E to three new responses which represent three independent parts of the information included in the original responses. Possible responses are the position of the apex, the width of the peak and its asymmetry. We define f_p as the transformation of the original responses to the *p*th new response $Y^{(p)}$ (Eq. (1)). Examples will be given below. Note that, for simplicity, we consider that the transformations of retention times to retention factors are included in the functions f_p's.

$$Y^{(p)} = f_p(B, A, E), (1 \leq p \leq P) \quad (1)$$

Then, $Y^{(p)}$ is the vector of size $N \cdot M$ containing the observed values of $Y^{(p)}$ for the *N* experiments and *M* compounds. The responses are ordered as *N* blocks of *M* components in this vector: $Y^{(p)} = \left(y_{11}^{(p)}, y_{12}^{(p)}, \dots, y_{1M}^{(p)}, y_{21}^{(p)}, \dots, y_{NM}^{(p)}\right)$.

The inverse transformation function f^{-1} must exist in order to reconstruct the (predicted) retention times from transformed responses. This will be useful to compute retention time-based criteria to assess the quality of chromatograms.

$$(B, A, E) = f^{-1}\left(Y^{(1)}, \dots, Y^{(P)}\right) \quad (2)$$

Such approach will prevent peak inversion in model prediction and clearly highlight, in the statistical modelling, the part of information available in each response.

3.4. Predictive models building

The ultimate goal of the model building phase is to be able to predict, for given settings $\boldsymbol{x}_i$ of the HPLC tuning parameters, the retention times *B*, *A* and *E* for each compound c_j of the mixture of interest. This will be achieved through the development of one model for each transformed response $Y^{(p)}$, $p = 1,\ldots, P$.

For given experimental conditions $\boldsymbol{x}_i$ and a compound c_j, a very general theoretical model for $y_{ij}^{(p)}$ can be written as:

$$y_{ij}^{(p)} = g_p\left(x_i, c_j; \beta_p\right) + \epsilon_{ij}^{(p)} \tag{3}$$

where the experimental errors $\epsilon_{ij}^{(p)}$ are supposed to be independent, identically distributed with mean 0 and common variance σ_p^2. The independence hypothesis is taken at three levels: between responses, between experiments (chromatograms) and between compounds within chromatograms. In design of experiments applications, the function g_p is often defined as a polynomial model and estimated by classical least squares.

The following steps are recommended to develop these predictive models.

3.4.1. Full model matrix definition

For the (*N*×*F*) design matrix ***X*** defined in Section 3.1, let's denote ***Z*** as the (*N*×*R*) model matrix containing the higher polynomial model terms that may be estimated with ***X***. ***Z*** will typically contain a constant term, main (qualitative and quantitative) factor effects, quadratic or higher order terms for quantitative factors and interactions. Quantitative factors should be centred and scaled in the [–1, 1] interval before being included in ***Z***. Qualitative factors should be coded into dummy variables.

In order to build predictive models for x_i and c_j, matrix ***Z*** should be expanded with all the interactions between *c* and polynomial terms in ***Z***. This full model matrix is built as $\boldsymbol{S} = \boldsymbol{Z} \otimes \boldsymbol{I}_M$ where $\otimes$ is the Kronecker product and $\boldsymbol{I}_M$ the $M \times M$ identity matrix. ***S*** is thus a $((N \cdot M)\times(R \cdot M))$ matrix containing all the terms which may potentially explain $Y^{(p)}$, the $(N \cdot M)$ vector of observed values for the *p*th response. $(R \cdot M)$ is potentially high and variables selection techniques will be necessary to achieve good predictive models.

3.4.2. Model estimation and dimension reduction

Multiple linear regression (MLR) or other statistical techniques such as partial least squares (PLS) are natural model estimation techniques in this framework. Good introductions can be found in Martens and Martens [12]. If MLR is applied, a variable selection technique as forward or stepwise will be necessary to select, for each response $Y^{(p)}$, the terms that are the most informative and avoid model overfitting. Let's note by $\boldsymbol{S}^{(p)}$ the $(N \cdot M)\times q^p$ submatrix of ***S*** coming out of the variables selection step for response $Y^{(p)}$ ($q^p \leq R \cdot M$). The least squares parameters estimators are then given by:

$$\hat{\beta}_p = \left(S^{(p)\prime}\, S^{(p)}\right)^{-1} S^{(p)}\, Y^{(p)}$$

For MLR, the number of terms kept in the model will typically be chosen simply by optimizing some criterion on the training set like the AIC or the Adjusted R^2 or, more heavily, through cross-validation and optimization of a criterion like the RMSE.

3.4.3. Model validation

Models must be validated before being used for prediction. A residual analysis will be appropriate to check model adequacy, detect outliers or heteroscedasticity. Appropriate *X*–*Y* scatter plots will allow to check visually the independence assumptions. If available, lack of fit tests will allow verifying if the residual variance is close to the experimental variance.

3.4.4. Prediction

For given factor setting x_0, the estimated models supply predictions $\hat{y}_{ij}^{(p)}$'s for each response p (p = 1,…, P) and each compound c_j. In MLR, one must first build the vector of model terms $s_{0j}^{(p)}$ of size q^p associated to x_0, c_j and response $Y^{(p)}$. The predicted responses are given by:

$$\begin{aligned} \hat{y}_{0j}^{(p)} &= s_{0j}^{(p)}\, \hat{\beta}_p = s_{0j}^{(p)}\, \left(S^{(p)\prime}\, S^{(p)}\right)^{-1} S^{(p)}\, Y^{(p)} \\ &= \hat{E}\left[y_{0j}^{(p)} | x_0, c_j\right], p = (1, \dots, P). \end{aligned} \tag{4}$$

One can then use the f^{-1} function to get original predicted responses $\hat{B}_{0j}$, $\hat{A}_{0j}$ and $\hat{E}_{0j}$ from the $y_{0j}^{(p)}$ (p=1,…, P) and thus, predict the complete chromatogram retention times for the condition x_0.

3.5. Criteria to assess the quality of a chromatogram

Several criteria can be defined to express quantitatively the quality of a chromatogram. Schoenmakers [1] summarized some definitions of various useful criteria. In the framework of the methodology presented in this paper, we favour criteria which assess globally (for the whole chromatogram) a given characteristic of the chromatogram. Criteria are based on retention times only. Possible characteristics of interest are the resolution, separation, peak width, asymmetry or maximum elution time.

More formally, we let B_j, A_j and E_j, j=1,…, M denote the retention times of the M peaks of a given chromatogram and $B_{(j)}$, $A_{(j)}$ and $E_{(j)}$, j=1,…, M the ordered ones (with respect to the retention time of the apex). Each criterion cr_z can then be defined as a specific function t_z of these retention times:

$$cr_z = t_z\left(B_j, A_j, E_j; j = 1, \dots, M\right)$$

Under these notations, the following interesting criteria may be defined: cr_1, the longer elution time which should be minimum; cr_2, the minimum separation between two subsequent peaks which should be maximum; cr_3, the maximum peak width which should be minimum; cr_4, the minimum peak resolution which should be maximum and cr_5, the maximum peak asymmetry which should be minimum. A way to express them formally is as follows:

$$cr_1 = max\left(A_j\right), j = 1, \dots, M$$
$$cr_2 = min\left(B_{(j+1)} - E_{(j)}\right), j = 1, \dots, M-1$$
$$cr_3 = max\left(E_j - B_j\right), j = 1, \dots, M$$
$$cr_4 = \min\left(\frac{2 \times \left(A_{(j+1)} - A_{(j)}\right)}{\left(E_{(j+1)} - B_{(j+1)}\right) + \left(E_{(j)} - B_{(j)}\right)}\right), j = 1, \dots, M-1$$
$$cr_5 = \max\left(\frac{\left|E_j + B_j - 2A_j\right|}{E_j + B_j}\right), j = 1, \dots, M$$

Thus, each global criterion is defined as the worst value of a calculated characteristic in a given chromatogram. This ensures that all other computed values, for other peaks or between other pairs of peaks, are at least better.

The methodology described in this paper is, of course, applicable to other criteria. For example, Vanbel [4] and Dewé et al. [5] show the use of limited optimization criteria (e.g. the separation of only 2 peaks of interest) and robustness criteria.

3.6. Definition of a global optimization criterion

Finding optimal chromatographic analytical conditions according to several criteria as defined in Section 3.5 is a multi-criteria optimization problem. A common methodology to approach such question has been introduced by Harrington [13] and Derringer and Suich [14]. They propose to aggregate the criteria of interest in one global optimization criterion in two steps.

First, each original criterion cr_z (z=1,…, Z) is transformed into a desirability value $d_z(cr_z)$ through a desirability function d_z. d_z takes its values between 0 and 1 where 1 corresponds to a highly "desirable" value for cr_z and 0 to a non-acceptable value. Values increasing between 0 and 1 express an increase of the "desirability" of the criterion. Second, all desirability values $d_1(cr_1)$,…, $d_Z(cr_Z)$ are aggregated in one global desirability index $D(d_1(cr_1),\dots,d_Z(cr_Z))$ to be optimized. This desirability index is also restricted to the [0,1] interval.

Different types of desirability functions are proposed in the literature. Harrington introduced the first desirability functions, using exponential functions. Derringer and Suich based their desirability functions on a power of a linear transformation of the responses (criteria). Recently, le Bailly de Tilleghem and Govaerts [15,16] proposed functions based on

the Normal cumulative distribution function (see Fig. 4). These functions present no discontinuities, keeping the strict order of criteria. More formally, they define $d_z(\ )$ as:

$$d_z(cr_z) = \Phi\left(\frac{cr_z - a_z}{b_z}\right) \quad \text{if } cr_z \text{ has to be maximized,}$$
$$d_z(cr_z) = 1 - \Phi\left(\frac{cr_z - a_z}{b_z}\right) \quad \text{if } cr_z \text{ has to be minimzed,} \tag{5}$$

where Φ is the cumulative distribution function (CDF) of the standard Normal variable defined as:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(\frac{-t^2}{2}\right) dt \tag{6}$$

a_z and b_z are respectively localisation and dispersion parameters to be fixed by the analyst according to the context and the criterion. Changes in the parameter a_z imply left or right shifts of the curve. Increasing the parameter b_z will make the curve less stiff.

Several global desirability indices are possible. The most popular is the weighted geometric mean of all individual desirability values:

$$D(cr) = \prod_{z=1}^{Z} \left(d_z(cr_z)\right)^{w_z} \text{ with } \sum_{z=1}^{Z} w_z = 1 \tag{7}$$

w_z values are fixed by the analyst according to the importance he wants to give to each chromatogram quality criterion in the global desirability index. Geometric mean is particularly adapted because one non-fulfilled criterion will lead to a very bad global desirability. A good global desirability ensures to have all the (weighted) criteria as good as needed.

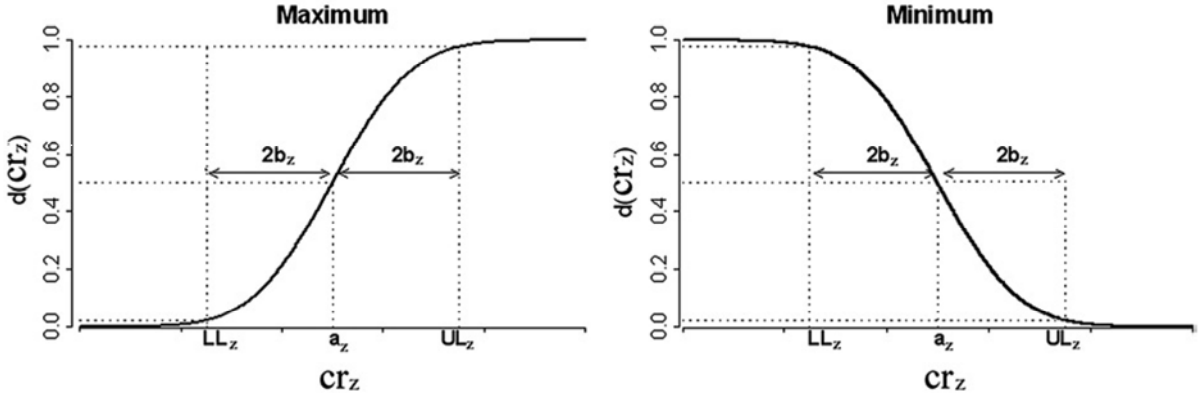


Fig. 4 Desirability functions based on the standard Normal distribution function. The first graph illustrates a criterion to be maximized while the second stands for a criterion to be minimized.

3.7. Search for optimal chromatographic conditions

The experimental design, the predictive models, the quality criteria and the desirability index described in the previous sections provide the necessary elements to reach the goal of this paper: optimize chromatographic analytical conditions. The models allow first to predict, for any values of the HPLC tuning parameters $\boldsymbol{x}=(x_1,\ldots, x_F)$ in the experimental domain χ, the

retention times B_j, A_j and E_j for the M compounds of the reference mixture of interest. These predictions can then be transformed to (estimated) chromatographic quality criteria $(\widehat{cr}_1, \widehat{cr}_2, \dots, \widehat{cr}_Z)$ and the global quality of the chromatogram summarized in an (estimated) desirability index $\widehat{D}\left(d_1(\widehat{cr}_1), \dots, d_Z(\widehat{cr}_Z)\right)$. Estimated optimal chromatographic conditions can then be obtained by searching a value $\boldsymbol{x}^*$ such that D is maximized:

$$\begin{aligned} \boldsymbol{x}^* &= \max_{x_0 \in \chi} D(\widehat{cr}) = \max_{x_0 \in \chi} \prod_{z=1}^{Z} \left(d_z(\widehat{cr}_z)\right)^{w_z} \\ &= \max_{x_0 \in \chi} \prod_{z=1}^{Z} \left(d_z\left(t_z\left(f^{-1}\left(g_1(x_0, c, \hat{\beta}_1), \dots, g_P(x_0, c, \hat{\beta}_P) \right) \right) \right) \right)^{w_z} \end{aligned} \tag{8}$$

Optimal solution can be found using grid search. Other methods exist, such as gradient descend or simplex algorithm but grid search is the most appropriate tool in our context because the dimensionality (number of factors to simultaneously optimize) is generally limited. Furthermore, it avoids falling into local optima and gives a global map of the evolution of the desirability index over the experimental domain χ.

3.8. Model prediction and error propagation

When optimal conditions are derived from statistical model predictions, it is crucial to study the impact of the model prediction error on the reliability of the solution found. This evidence is however rarely highlighted in the multi-response optimization design of experiments literature, but has been discussed recently by Trautmann and Weihs [17] and le Bailly de Tilleghem and Govaerts [16].

These authors propose different approaches to quantify the incertitude of the desirability functions values and of the desirability index. They introduce also the notion of equivalence zone around the optimum. For simple cases, exact or approximate analytical solutions exist to implement these concepts but for more complex situations, Monte-Carlo simulations are recommended. The situation of this paper must be considered as complex due to the introduction of the chromatogram quality measures in the calculation of the desirability index.

This section proposes a Monte-Carlo approach to establish the distributions of the desirability functions and of the global desirability index. It shows then how these results allow to derive an equivalence zone for this optimum. Let's take $\boldsymbol{x}_0$, a given value for the HPLC tuning parameters, and c_j, a given compound of interest in the chromatogram. The prediction models described in Section 3.4.4 provide, for each response $Y^{(p)}$ and each compound c_j, an estimate of the expected response: $\hat{\mu}_{Y_{0j}}^{(p)} = \hat{E}\left[y_{0j}^{(p)} | \boldsymbol{x}_0, c_j\right]$. Let's then suppose that an estimate $\hat{\sigma}^2\left(\hat{\mu}_{Y_{0j}}^{(p)}\right)$ is available for the variance of this estimator. In the MLR framework, it is given by (see Eq. (4) for notations):

$$\widehat{\boldsymbol{\sigma}}^2\left(\hat{\mu}_{Y_{0j}}^{(p)}\right) = \widehat{\boldsymbol{\sigma}}_{\boldsymbol{\epsilon}_P}^2 \boldsymbol{s}_{0j}^{(p)} \left(\boldsymbol{S}^{(p)\prime} \boldsymbol{S}^{(p)}\right)^{-1} \boldsymbol{s}_{0j}^{(p)\prime} \tag{9}$$

where the residual variance estimator $\widehat{\boldsymbol{\sigma}}^2_{\boldsymbol{\epsilon}_P}$ can be estimated from the training data as:

$$\widehat{\boldsymbol{\sigma}}^2_{\boldsymbol{\epsilon}_P} = \frac{1}{NM - q^p} \sum_{i=1}^{N} \sum_{j=1}^{M} \left(y_{ij}^{(p)} - \hat{y}_{ij}^{(p)} \right)^2 \quad (10)$$

One can then generate, for each response $Y^{(p)}$ and each compound c_j, a large set ($i^* = 1,\ldots, N_{sim}$) of simulated "predicted" responses in $\boldsymbol{x}_0$ according to:

$$Y_{0i^*j}^{(p)} = \hat{\mu}_{Y_{0j}}^{(p)} + \epsilon_{0i^*j}^{(p)}$$

where the $\epsilon_{0i^*j}^{(p)}$ are random Normal variables with mean 0 and variance $\hat{\sigma}^2 \left(\hat{\mu}_{Y_{0j}}^{(p)} \right)$. In the MLR context, this normality assumption is common.

For each simulation i, ($M \times P$) response values are thus generated. Then, original responses (B_j, A_j and E_j), quality criteria (cr_z) and the global desirability index D are derived. This allows establishing Monte-Carlo distributions for the quality criteria, and for the global desirability index, which give an idea of the impact of the models prediction error on the uncertainty of the resulting estimations.

3.9. Optimum equivalence zone

The Monte-Carlo distribution of the global desirability index gives also the necessary information to define a zone of the experimental domain which cannot be stated to give significantly worse results than the optimum $\boldsymbol{x}^*$ found. The equivalence zone will simply be defined as the set of $\boldsymbol{x}$'s of the experimental domain χ such that the (estimated) desirability index $D(\boldsymbol{x})$ is greater than the 5th percentile of the Monte-Carlo distribution of the desirability index at the optimum $\boldsymbol{x}^*$. This definition is not perfectly correct because there is also an uncertainty on $D(\boldsymbol{x})$, but it is easy to implement and is sufficiently informative in the context of this paper. The size of the equivalent zone will give an idea of the real interest of the optimization process. A large zone indicates that the optimization process could not really differentiate between the quality of different analytical conditions. This may be due to the fact that there is effectively no real difference between the experimental conditions over the explored domain, or that the uncertainty in the predictive models is too large to be able to highlight the potential differences.

From a practical point of view, note that, when the number of quantitative factors of interest is small, the equivalence zone can elegantly be represented on graphs of contour plots of the predicted desirability index over the experimental domain for fixed values of the qualitative factors. This will be illustrated in the next section.

3.10. Design space

The last section showed how to propagate the error of the statistical models, in the optimal factor configuration, to find an equivalence zone. A similar approach can be used to find, with a known confidence, a design space [8]. Design space is defined as the established

range of process parameters and formulation attributes that have been demonstrated to provide assurance of quality. A design space is particularly useful. It can be seen as a zone of robustness defined in the experimental domain because working within is not considered as a change in the analytical method. At opposite, working out of the design space is considered to be a change and normally initiates new steps of validation.

The desired minimum quality must be fixed a priori. For instance, it is desirable to know if there is a zone, in the experimental design, where it is likely to have a separation of at least m_1 minutes and a total processing time not exceeding m_2 minutes. So, one will look, for each point of the experimental domain, under the propagated error, how many hits of sufficiently good chromatograms are achieved. More formally, we have (for criteria to be maximized):

$$[\forall x_0 \in \chi | P(cr_z > \lambda_z) > \gamma\%] \; \forall z \text{ in } 1, \dots, Z \quad (11)$$

Similar formulation can be written for criteria to be minimized. A unified equation for maximized and minimized criteria can be written in the desirability space:

$$\left[\forall x_0 \in \chi | P\left(d_z(cr_z) > d_z(\lambda_z)\right) > \gamma\%\right] \; \forall z \text{ in } 1, \dots, Z \quad (12)$$

Fig. 5 illustrates this formulation, for a point $x_0 \in \chi$, and represents simulated distribution of a desirability function under the propagated error of the models. A limit value for the criteria $d_z(\lambda_z)$ is first chosen. The interest is to observe if γ% of the distribution is higher than $d_z(\lambda_z)$.

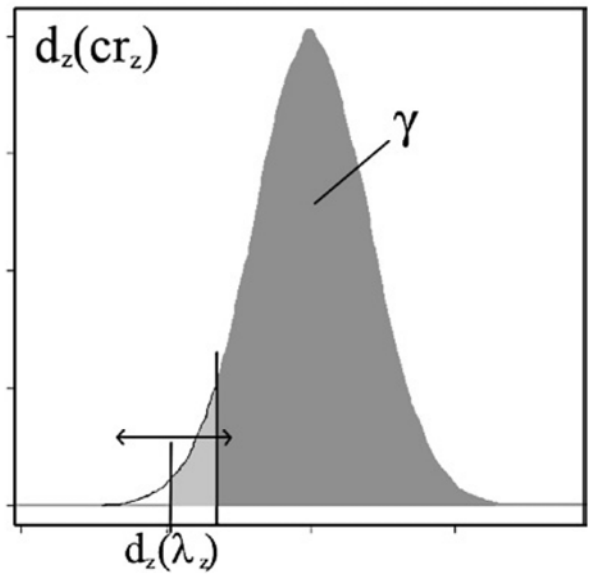


Fig. 5 Simulated distribution of a criteria (desirability function) under propagated error of the models. The interest is to look if γ% of the distribution of $d_z(cr_z)$ is higher than a chosen limit $d_z(\lambda_z)$.

Thus, a point x_0 belongs to the design space if all the criteria of interest are fulfilled with a certain level a confidence γ%. A wide design space, defined with strong constraints on the criteria (high $d_z(\lambda_z)$), and with a high level of confidence γ, is the most enviable situation in the experimental domain. As for the equivalence zones, if the number of quantitative factors of interest is small, the design space can be represented on graphs of contour plots.

4. Application

4.1. Experimental

The objective of the separation is the evaluation of the methodology. Consequently, the compounds under investigation were not selected on the basis of an analytical problem nor for a practical application.

4.1.1. Chemical and reagents

A mixture of five commercially available compounds (Diflunisal, Granisetron, Nifedipine, Phenytoine, Sulfinpyrazone) was analysed using the design described here after (Section 4.2). These five compounds were obtained from the Eli Lilly pharmaceutical company (Indianapolis, USA). Methanol and acetonitrile of HPLC grade were purchased from Sigma–Aldrich (Steinheim, Germany). Ultra-pure water was obtained from an Academic A10 Milli-Q system (Millipore, Eschborn, Germany). Acetic acid (N98%) was purchased from Fluka (Steinheim, Germany), ammonium formate (99%) was purchased from Sigma-Aldrich.

4.1.2. Appartus

All analysis were performed on a Waters 2695 separation module coupled to a Waters selector valve 7678 and a Waters 996 Photodiode array detector (Waters, Eschborn, Germany). The Empower 1.0 software was used to manage chromatographic data. Five analytical columns were used: C18, C8, RP18, Phenyl XBridge columns (Waters) (100×2.1 mm i.d.; particle size 3.5 µm) and a C18 Cogent Bidentate column (100×2.1 mm i.d.; particle size 4.0 µm) (Microsolv, Villecresnes, France).

4.1.3. Chromatographic conditions

The elution gradients were performed in 10, 20 or 30 min from 5% to 95% of organic modifier (methanol or acetonitrile) in the adequate buffer at a constant flow rate of 0.25 ml/min at 30 °C. The buffer solutions were adjusted at the desired level of pH:

• pH 2.6: 0.1% concentrated Formic acid (99%) in water,
• pH 4.0, 5.0 and 7.0: Ammonium Formate 10 mM in water adjusted with concentrated formic acid and/or ammonia aqueous solution (35%).

The analytes were monitored photometrically at 240 nm although chromatographic data were recorded from 210 to 400 nm for all the analytical conditions investigated. The analytes were dissolved in an acetonitrile/water mixture (50:50, v/v). The injection volume was 2.0 or 5.0 µl.

4.1.4. Software

We used the statistical language R 2.4.0 for Windows to implement the methodology presented in the last section.

4.2. Design of experiments

We applied a full factorial design on 4 factors: the pH of the mobile phase, the time used for the gradient, the solvent used in the composition of the buffer, and the analytical column. It is clear these factors highly affect the position of peaks in the chromatograms. We used gradient mode because the behaviour of peaks elutions is not as known as in isocratic conditions. The goal was to stress the methodology using less repeatable experiments with gradient mode with less known behaviour of peaks elutions.

Moreover, the analytes have been chosen to cover a large range of pKa and log P. Gradient mode gives more chance to obtain all the analytes eluted on the chromatogram in a constrained time. Temperature, column diameter, injection volume or other chromatographic parameters were not included in the design example. Their values were then fixed as explained in Section 4.1.

Factors and levels of factors were chosen to validate the methodology. Fig. 6 shows the values of the quantitative factors (pH and gradient time) over the experimental domain. 3 levels of pH and gradient were investigated in a full factorial design. Intermediate points were added to validate the methodology.

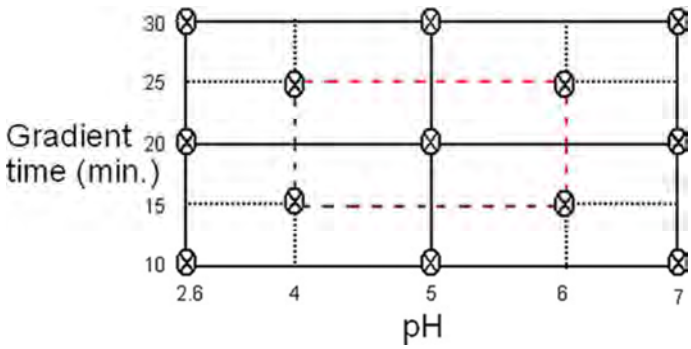


Fig. 6 Full factorial design applied on the quantitative factors (plain lines). Intermediate points (pH 4 and pH 6) have been added to the design (dashed lines) to validate the methodology.

5 analytical columns and 2 solvents have been selected as qualitative factors.

- Column: Bidentate C18, XBridge C18, XBridge C8, XBridge Phenyl and XBridge RP18
- Solvent: CH_3CN and MeOH

Analytical columns were chosen for their different chemical interactions with solutes and also their robustness against large range of pH. The number of columns was a good compromise between domain investigation and overloading of experimental design.

Full factorial design on the 4 used factors gives 3 levels of pH × 3 gradients × 5 columns × 2 solvents = 90 experiments that have been first realized. This is quite a lot of experiments, but it was useful to set up the models properly. The intermediate validation points were

realized on both CH_3CN and MeOH solvent, and on 2 columns: XBridge C18 and XBridge Phenyl. This gives 2×2×2×2=16 experiments to generate a test set. Validation and test data give 90+16 = 106 experiments. 5 compounds were analysed, giving $N \times M$ = 106×5 = 530 observations. Of course, final use of the methodology will lead to less experiment.

Fractional factorial design or D-optimal design are recommended when the number of factors to simultaneously optimize increases. Another improvement is to select less analytical columns before realizing experiments. A good practice is to choose columns which possess very different physico-chemical properties.

For instance, reducing this number of columns to 2 would lead to 36 experiments instead of 90 in full factorial design. Using D-optimal design, number of experiments can be decreased to 24 or even 12. This is quite reasonable to develop a method with 4 factors to be simultaneously optimized.

4.3. Responses selection

We first modelled the retention factors k'. We assumed that the peaks were symmetric, which was suitable for this example. Thus, we modelled the apex and the width of the peaks. The supposed symmetry allows to define easily f^{-1} and thus to reconstruct original retention times (see later). Formally,

$$\begin{aligned} Y^{(1)} &= \log(k'_A) \\ Y^{(2)} &= \log(k'_E - k'_B) \end{aligned} \tag{13}$$

Other responses definitions can also be used. We chose these responses because they possess the best modelling properties.

Table 1 Summary of the fits (MLR) for the selected responses

	$Y^{(1)} = \log(k'_A)$	$Y^{(2)} = \log(k'_E - k'_B)$
Adjusted R^2	0.96	0.51
# of parameters	180	68

4.4. Models

We used polynomial models with a forward variables selection of included terms maximizing the Adjusted R^2 for model $Y^{(1)}$ and $Y^{(2)}$. This method gave the best predictive results.

Models possess a high number of parameters such that it is not possible to describe them completely. Table 1 gives a brief summary of the models. The most impactful model, for the apex, seems good while the fit for the width is very limited. This can be seen as problematic, but the error on the widths estimation has less importance on the final estimation of the retention times (E and B). Moreover, this error is taken into account in the final steps of the methodology. Validations of the models are bypassed in this report.

4.5. Prediction

So far, the models have been set up and can be used to predict new values of responses on the complete experimental domain χ. Retention times vectors can be reconstructed as follows:

$$\hat{k}'_A = \exp\left(\hat{Y}^{(1)}\right); \hat{k}'_B = \hat{k}'_A - \frac{\exp\left(\hat{Y}^{(2)}\right)}{2}; \hat{k}'_E = \hat{k}'_A + \frac{\exp\left(\hat{Y}^{(2)}\right)}{2} \quad (14)$$
$$\hat{A} = (T_0 \times k'_A) + T_0; \text{ idem for } \hat{B} \text{ and } \hat{E}$$

This set of equations is the function f^{-1} presented in Eq. (2). Criteria can be computed from the retention times $\hat{A}$, $\hat{B}$ and $\hat{E}$. This allows computing the desirability functions and global desirability index on each points of χ.

4.6. Optimum finding

Then, a grid search was applied. The only parameter to tune is the value of the step of the quantitative variables. This defines a grid more or less dense. Computational explosion can arise if the step is too small. To avoid this, the search can be done in several steps, increasing the precision around the optima found. Our selected step is 0.1. Values of quantitative factors are still normalized in the interval [−1,1] in this process, giving 21 points to be computed for each quantitative factors. The full factorial combination of factors then led to 21 levels of pH × 21 gradients × 5 columns × 2 solvents = 4410 points to be predicted.

For each point, complete chromatogram, criteria, desirability values of criteria and global desirability index were easily computed, following Eq. (8).Different global desirability indexes can be found using different weights w_z (see Section 3.6).

The desired objective in our example was the best separation of all peaks in the minimum processing time. Only the criteria of minimal separation cr_2 and of maximum retention time cr_1 were then used. In this case, other criteria had little interest or were non-informative (e.g. asymmetry).

We first considered default weights for both criteria ($w_1 = w_2 = 1/2$). However, Fig. 7 (left) shows a non-achieved separation between the fourth and the fifth peaks. To manage this problem, we elaborated a second solution, that tends to give less weight to the maximum retention time ($w_1 = 1/6$, $w_2 = 5/6$) in order to achieve a better separation. The optimal points $\boldsymbol{x}^*$ for these two solutions are given in Table 2 (see Eq. (8).

Table 2 Values of optimal points $\boldsymbol{x}^*$

	(Column, pH, solvent, gradient time)
Solution 1	(XBridge Phenyl, 3.92, CH_3CN, 12)
Solution 2	(XBridge RP18, 7, CH_3CN, 18)

These points are the ones in χ that maximize the global desirability.

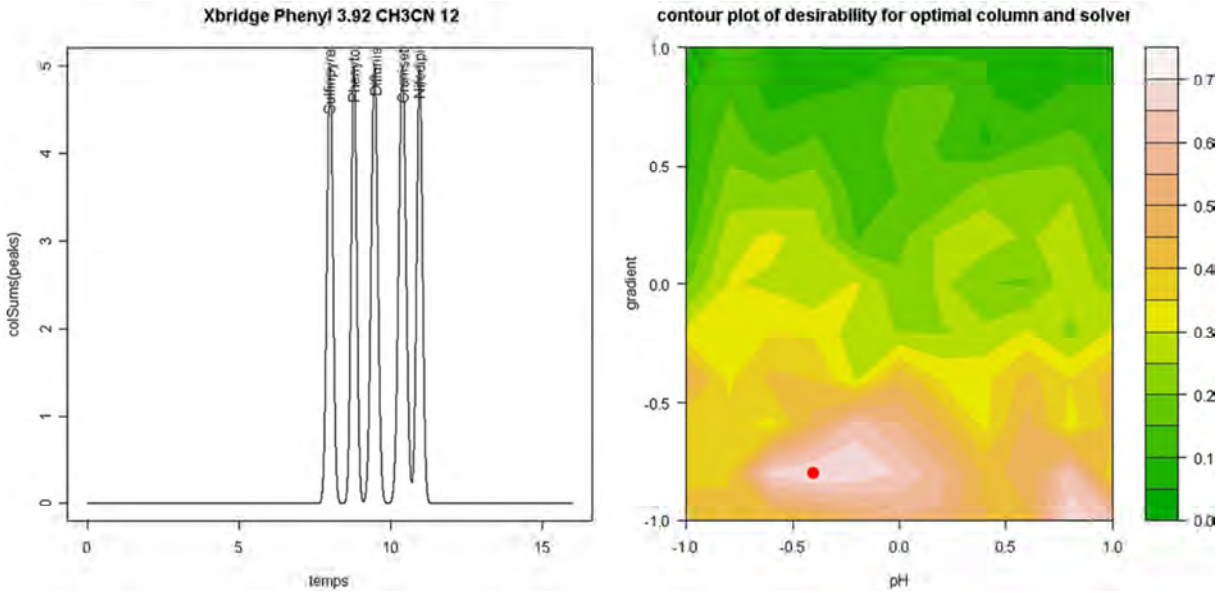


Fig. 7 Solution 1. Left: predicted optimized chromatogram with equal weights for criteria. Right: corresponding contour plot of the global desirability index across the quantitative normalized domain (qualitative factors fixed to optimum). The red (dark) point shows optimum.

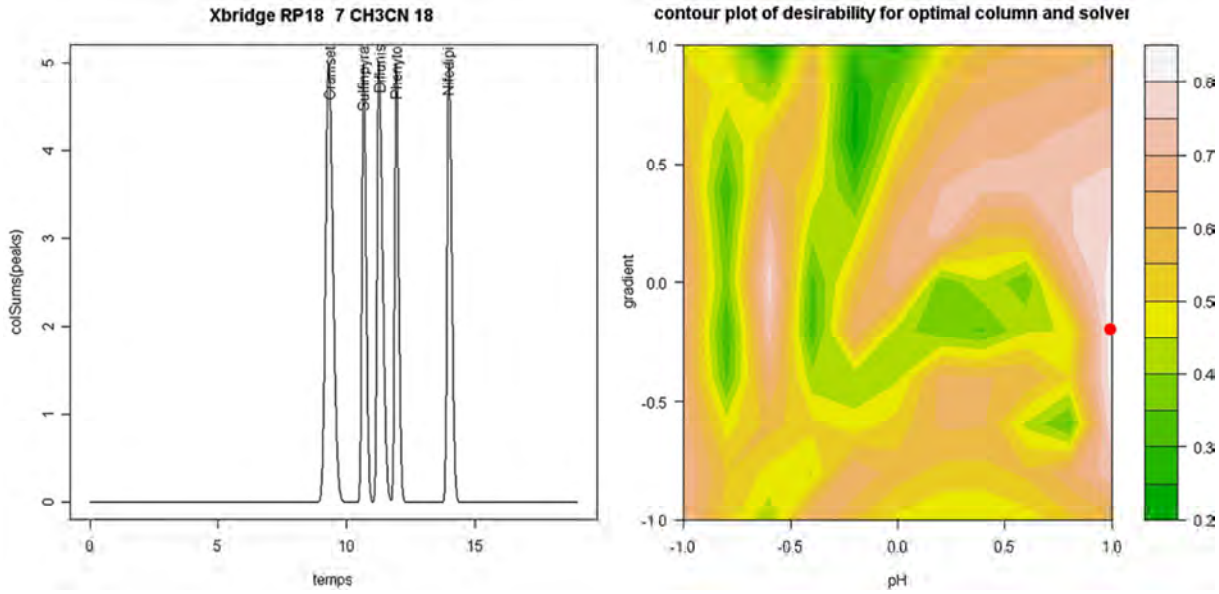


Fig. 8 Solution 2. Left: predicted optimized chromatogram with higher weight for separation. Right: corresponding contour plot of the global desirability index across the quantitative normalized domain (qualitative factors fixed to optimum). The red (dark) point shows optimum.

Figs. 7 and 8 show, for the two solutions envisaged, the optimal predicted chromatogram and the corresponding contour plot of the global desirability index

(qualitative factors are fixed to optimum). The second chromatogram has a better separation but takes more time to be processed. It is clearly shown that the weights allow finding an acceptable solution (without any more experiment). This gives a lot of flexibility to this approach.

However, the used gradient (from 5% to 95% of organic modifier in the buffer) leads to suboptimal solutions. Indeed, the peaks elutes in 5 min from 10 to 15 min. The optimality of a solution is only valid within the experimental domain and it is still the responsibility of the analyst to define the most interesting domain. Other gradient slopes or isocratic conditions would have been used and maybe, would have given better performance.

4.7. Error propagation

It is important to be able to give confidence in these optimal solutions. Uncertainty of the expected estimated responses for the predicted optimal points can be propagated. One can look at the values of the desirability of criteria, under propagated error, with histograms of their distribution. The main interest is to look if the error of the models allows having confidence in the values of desirability of criteria and also the global desirability index.

Fig. 9 shows the error propagated on the considered criteria and on the global desirability, for the two solutions envisaged. In the first solution (left column), we wanted to maximize separation and to minimize total retention time of the chromatogram. In the second solution, we gave less relative weight to the retention time criterion. This leads to a worse distribution of the error propagated for this criterion (first row of the array), while the distribution for the separation possess a certain asymmetry underlying the better achievement of this criterion (second row). The same observation can be done on the global desirability (last row). Notice that the criteria are shown non-weighted (rows 1 and 2 of the graph), but global desirability takes the weights into account (last row).

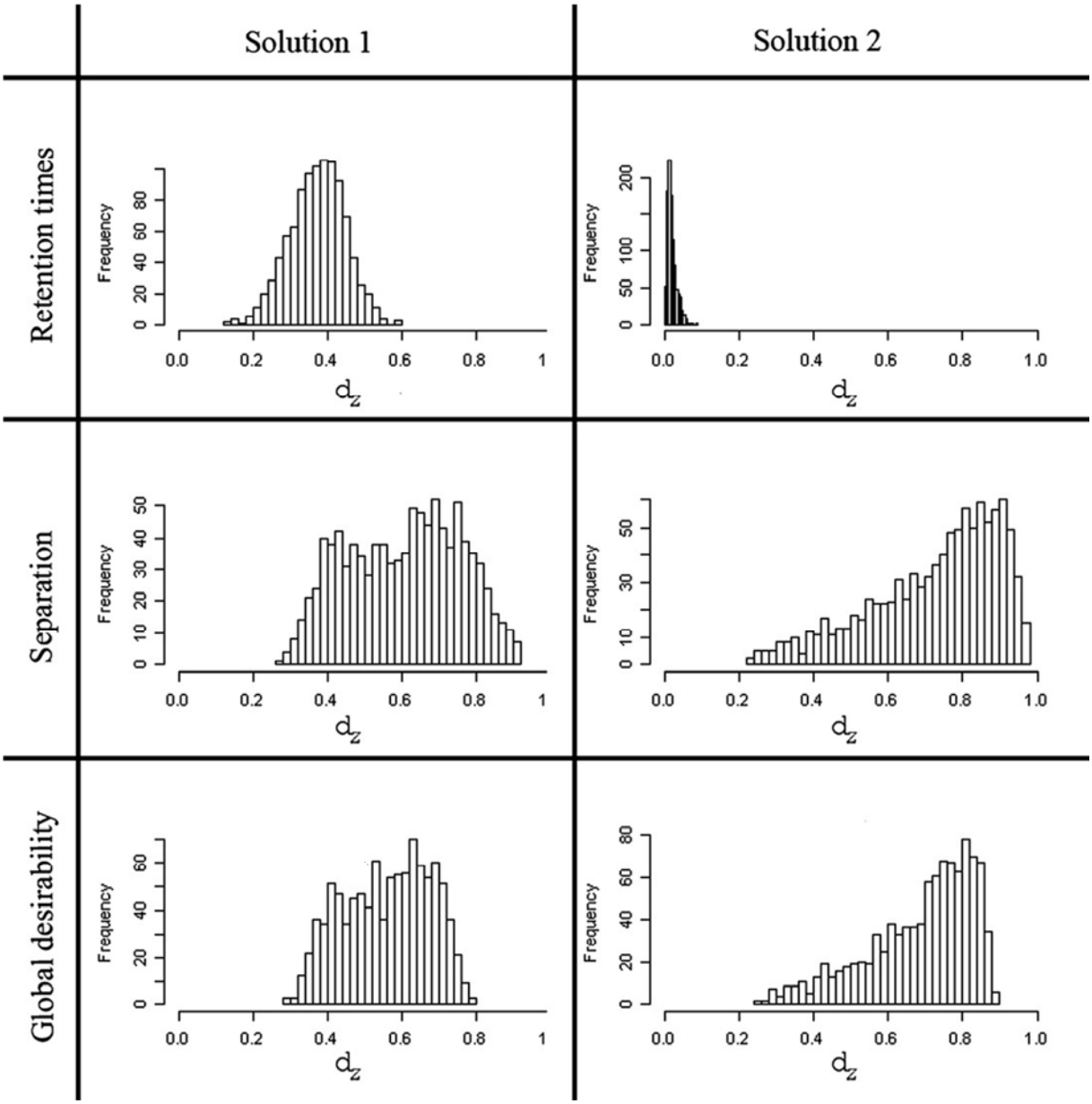


Fig. 9 Distributions of the two estimated criteria (desirability functions of maximum retention times and minimal separation) and global desirability for the two solutions envisaged.

4.8. Equivalence zones

Fig. 10 shows the equivalence zones at 95% on the contour plot of the Figs. 7 and 8. The first graph (left) shows clearly differentiated values. This leads to the conclusion that this optimal point has a limited equivalence zone. The optimal equivalence zone is better (in term of desirability) than elsewhere in the design of experiment. The second graph cannot be used to give a similar conclusion.

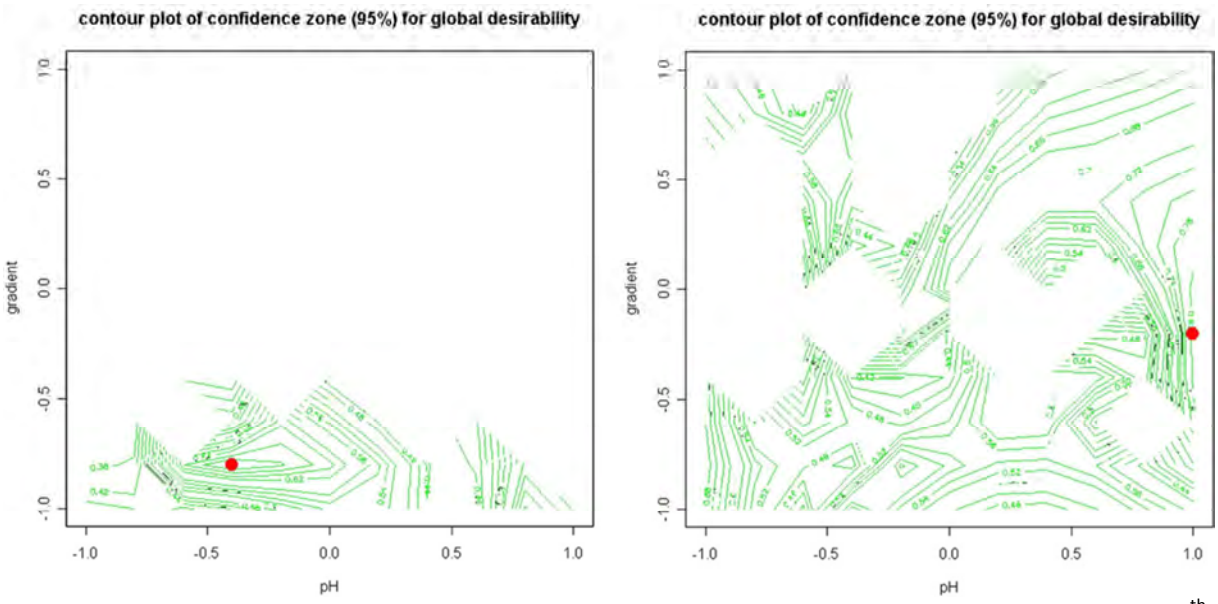


Fig. 10 Contour plot showing the set of solutions giving an estimated desirability greater than the 5th percentile of the simulated global desirability index under error of the models (at optimal solution). Left: solution 1. Very few conditions give desirability likely close to the optimal solution. Right: solution 2. Over the domain, most conditions are likely to give desirability index close to the optimal one.

Table 3 Design space: limit values for the criteria and confidence in design space

Criteria	Value
Separation	0.0 min (minimum)
Total retention time	20 min (maximum)
Probability γ	60%

4.9. Design Space

The final objective was to observe the existence, or nonexistence, of a design space in the experimental domain. A design space is considered as a zone of robustness in the experimental domain because it allows to tolerate variability of materials and slight changes in the process. The Table 3 shows the selected values for criteria to be fulfilled. According to these criteria, a small design space is found (see Section 3.10 for mathematical details), as shown in the Fig. 11. It is certain that the used level of confidence and the chosen limits for the criteria are rather weak. Solution 2 is included in this design space.

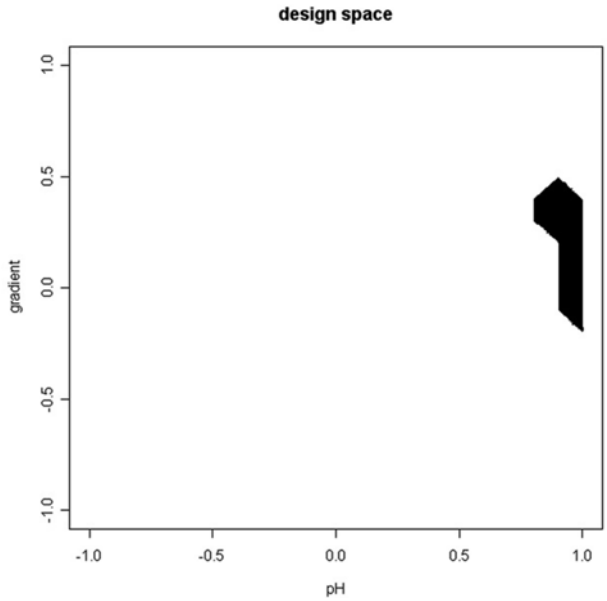


Fig. 11 Contour plot of the design space (dark area) for the values presented in Table 3.

4.10. Validation

The second solution has been validated. Fig. 12 compares the chromatogram obtained with the proposed optimal conditions (right) with the one predicted using solution 2 (left). The positions of the apexes of the peaks do not suffer from excessive imprecision but the width of the peaks is clearly not well predicted. This was foreseeable due to the bad fitting of the model for the width ($Y^{(2)}$).

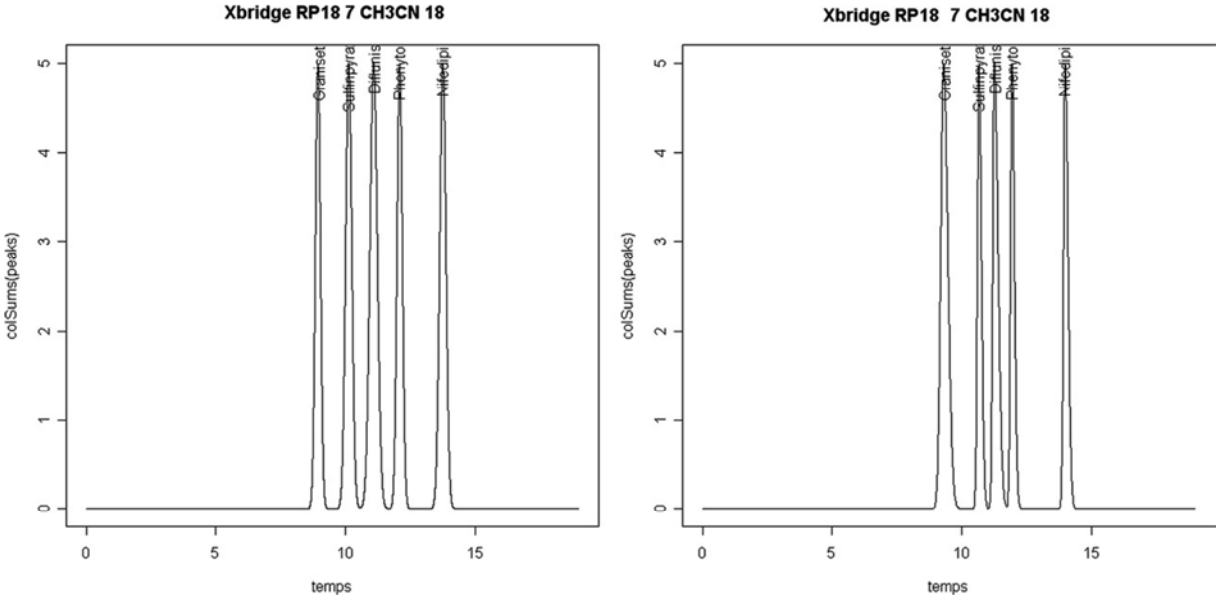


Fig. 12 Validation of optimal solution 2. Left: predicted chromatogram. Right: chromatogram obtained when using optimized HPLC parameters (Column XBridge RP18, pH 7, Solvent CH_3CN, and gradient time 18 min).

5. Conclusions

A very flexible methodology for modelling chromatography has been proposed and shown in order to optimize analytical conditions. This is an important step to permit the automation of methods development. The discretization of retention times is so far the main manual part of the job but evidences show that this process can also be automated. Given the flexibility of the proposed methodology, it is now possible to envisage modelling for automation on real process.

The use of the Normal CDF is a very convenient and automated way to combine various criteria into a global desirability index. This also allows flexibility in regards of the chosen criteria. Furthermore, they can be weighted according to the need of the analysis.

Finally, assessing the way the uncertainty propagates into global desirability index is simple and effective to verify the confidence in optimal condition(s). Design space can also be computed with confidence, using a similar methodology of models error propagation, on each point of the experimental domain. Robustness can finally be assessed or validated analytically with new experiments in the design space around optimal solution(s).

Acknowledgements

This paper has been written for the ADAM project. The acronym ADAM stands for Automated Development of Analytical Methods. The project is born from a cooperation between Eli Lilly company and the Institute of Pharmacy, laboratory of analytical chemistry, University of Liege. The authors thank the Eli Lilly Company and the Walloon Region (FIRST Doctorant Entreprise International project ADAM no. 516130) for their research grants. We would also like to thank Professor Paul Eilers, University of Leiden, The Netherlands, for the precious advice given during this work.

References

[1] P. Schoenmakers, Optimization of chromatographic selectivity: a guide to method development, Anal. Chim. Acta 208 (1988) 357.

[2] L. Snyder, J. Kirkland, J. Glajch, Practical HPLC Method Development, second edition, Wiley-Interscience, 1997.

[3] D. Massart, et al., Chemometrics: A Textbook (Data Handling in Science and Technology), Elsevier Science, 1990.

[4] P. Vanbel, Development of flexible and efficient strategies for optimizing chromatographic conditions, J. Pharm. Biomed. Anal. 21 (1998) 603.

[5] W. Dewé, R. Marini, P. Chiap, P. Hubert, J. Crommen, B. Boulanger, Development of response models for optimizing HPLC methods, Chemom. Intell. Lab. Syst. 74 (2004) 263.

[6] G. Cox, W. Cochran, Experimental Designs, 2nd edition, Wiley, 1957.

[7] B. Winer, Statistical Principles in Experimental Design, 2nd edition, McGraw-Hill, 1962.

[8] ICH, Q8 Draft Guidance on Pharmaceutical Development, 2004 Ver 4.3.

[9] A. Khuri, J. Cornell, Response Surfaces: Designs and Analyses, Marcel Dekker, 1987.

[10] G. Vivó-Truyols, New Strategies for Optimisation and Data Treatment in HPLC. PhD thesis, University of Valencia, 2005.

[11] B. Boulanger, Utilisation de l'Analyse en Composantes Independantes (ICA) pour la separation numerique des pics et la quantification automatique en CLHP-UV, Conference Chimiometrie 2006, Paris, December, 2006.

[12] H. Martens, M. Martens, Experimental Designs, 2nd edition, Wiley, 2001.

[13] E.Harrington,The desirability function, Ind.Qual.Control 21 (1965) 494.

[14] G. Derringer, R. Suich, Simultaneous optimization of several response variables, J. Qual. Technol. 12 (4) (1980) 214.

[15] C. Le Bailly de Tilleghem, B. Govaerts, Uncertainty propagation in multiresponse optimization using a desirability index, Tech. Rep., vol. 0532, Universite catholique de Louvain, Louvain-la-Neuve, 2005.

[16] C. Le Bailly de Tilleghem, B. Govaerts, Distribution of desirability index in multicriteria optimization using desirability functions based on the cumulative distribution function of the standard normal, Tech. Rep., vol. 0531, Universite catholique de Louvain, Louvain-la-Neuve, 2005.

[17] H. Trautmann, C.Weihs, Uncertainty of optimum influence factor levels in multicriteria optimization using the concept of desirability, Tech. Rep., SFB 475, vol. 23/04, Dortmund University, 2004.

Section III.1.2

Application de l'analyse en composantes indépendantes à la chromatographie liquide

B. Debrus, P. Lebrun, A. Ceccato, G. Caliaro, B. Govaerts, B.A. Olsen, E. Rozet, B. Boulanger, Ph. Hubert, A new statistical method for the automated detection of peaks in UV-DAD chromatograms of sample mixture, Talanta 79 (2009) 77.

Résumé

Une des limitations majeures dans le contexte du développement entièrement automatisé de méthodes chromatographiques consiste en la détection et l'identification automatique des pics provenant d'échantillons complexes tels que des formulations pharmaceutiques comprenant plusieurs composés ou les études de stabilité de ces dernières. Le même problème peut également survenir lors de l'analyse d'extraits de plantes ou d'échantillons biologiques. Cette étape est donc cruciale et chronophage, principalement quand une approche par plans d'expériences est utilisée pour générer les chromatogrammes. L'utilisation de plans d'expériences maximise les variations des conditions analytiques afin d'explorer un domaine expérimental. Cependant, cela fournit des chromatogrammes très différents qui peuvent être difficiles à interpréter, et donc qui compliquent les étapes de détection et d'appariement des pics. Dans ce contexte, l'analyse en composantes indépendantes (ACI), une nouvelle méthode d'analyse statistique de signal a été utilisée pour pallier ce problème. Le principe de l'ACI est d'assumer que le signal observé est la résultante de plusieurs phénomènes (également appelés sources), et que ces sources sont statistiquement indépendantes. Selon ces hypothèses, l'ACI est capable d'estimer les sources qui ont une grande probabilité de représenter les composantes constitutives d'un chromatogramme. Dans la présente étude, l'ACI a été appliquée avec succès pour la première fois à des chromatogrammes CLHP-UV. Il a été démontré que l'ACI permet la différenciation des composantes relatives au bruit et aux artéfacts de celles d'intérêts en appliquant des méthodes de classification grâce à l'utilisation des statistiques de moments d'ordre élevé calculées sur ces composantes. De plus, sur la base de la stratégie décrite, il a également été possible de reconstruire des chromatogrammes "épurés" où l'influence du bruit et des artéfacts de la ligne de base sont minimaux. Cette méthodologie peut représenter une avancée majeure en ce qui concerne l'objectif de fournir des outils utiles pour le développement automatisé de méthodes chromatographiques. Finalement, il semble que la recherche en chimie analytique utilisant la chromatographie pourrait être accélérée en utilisant ce type de méthodologie.

Summary

One of the major issues within the context of the fully automated development of chromatographic methods consists of the automated detection and identification of peaks coming from complex samples such as multi-component pharmaceutical formulations or stability studies of these formulations. The same problem can also occur with plant materials or biological matrices. This step is thus critical and time-consuming, especially when a Design of Experiments (DOE) approach is used to generate chromatograms. The use of DOE will often maximize the changes of the analytical conditions in order to explore an experimental domain. Unfortunately, this generally provides very different and "unpredictable" chromatograms which can be difficult to interpret, thus complicating peak detection and peak tracking (i.e. matching peaks among all the chromatograms). In this context, Independent Components Analysis (ICA), a new statistically based signal processing methods was investigated to solve this problem. The ICA principle assumes that the observed signal is the resultant of several phenomena (known as sources) and that all these sources are statistically independent. Under those assumptions, ICA is able to recover the sources which will have a high probability of representing the constitutive components of a chromatogram. In the present study, ICA was successfully applied for the first time to HPLC–UV/DAD chromatograms and it was shown that ICA allows differentiation of noise and artefact components from those of interest by applying clustering methods based on high-order statistics computed on these components. Furthermore, on the basis of the described numerical strategy, it was also possible to reconstruct a cleaned chromatogram with minimum influence of noise and baseline artefacts. This can present a significant advance towards the objective of providing helpful tools for the automated development of liquid chromatography (LC) methods. It seems that analytical investigations could be shortened when using this type of methodologies.

1. Introduction

In many frameworks such as the automated development and optimization of analytical methods, the use of Design of Experiments (DOE) is of great interest to generate data over a broad experimental space. This requires the choice of a well-structured set of experiments to allow the modelling of the studied responses. However, analytical conditions of a liquid chromatography (LC) method must be varied significantly in order to cover an experimental space as large as possible. This can lead to very different chromatograms and peak tracking from one chromatogram to another is often problematic. Peak detection in the case of overlapping peaks can also present a problem. Even if this step can be done manually by an experienced analyst, there is a possibility of error in peak assignment. The automation of this process is therefore one of the first problems to solve to achieve the fully automated development of LC methods.

In the literature, different approaches to peak deconvolution and tracking have been described. A mathematical treatment like deconvolution (one-dimensional or less frequently multidimensional) is a widely used technique to accurately estimate the overlapping of peaks. It only needs the number of peaks and some information about the peak shapes as basic input parameters in the computer assisted peak deconvolution procedure [1–3]. However, the knowledge of overlapped peak shapes is sometimes difficult to determine, and poor results can sometimes be obtained [4–6].Methods such as alternating least-square multivariate curve resolution [7,8], mutual automated peak matching [9] or factor analysis [10] have been recently developed and seem to be very efficient in the multidimensional numerical separation domain applied to HPLC–diode array detection (DAD) data. However, examples with real data demonstrate that some work is still necessary to achieve better automated detection and tracking of peaks [9–10].

Independent Component Analysis (ICA) is an interesting alternative that may be used to achieve peak detection and tracking. ICA is used in many domains due to its ability to separate successfully many types of signals [11]. In this context, ICA is a blind source separation method. The term blind indicates that both the source signals, and the way they are mixed, are unknown. This processing method has been successfully applied to the analysis of mixed sounds, satellite signals, to biomedical signal processing problems such as electroencephalographic (EEG) data [12,13], and used in order to eliminate artefacts from EEG data [14] or functional magnetic resonance imaging data [15]. Recent papers have shown that ICA has been applied in the field of chromatography for metabolite peak detection [16] or to resolve overlapping signals from gas chromatographic–mass spectrometry [17,18] or 3D-fluorescence spectroscopy [19] and for the signal analysis in proteomic and metabolomic investigations [20,21]. Independent components (sources) can be obtained by applying this statistical method which maximizes the statistical independence of the estimated sources. To measure the independence between the sources, the calculation of non-Gaussianity can be used, as explained by Hyvärinen and Oja [22]. Another popular criterion for measuring statistical independence of signals can be mutual information between sources. Typical algorithms for ICA use pre-processing steps in order to

simplify and reduce the complexity of the problem. Data centering, whitening (i.e. signals decorrelation, accomplishable by using a classical principal component analysis method) and dimensionality reduction are generally used. ICA can be carried out by various algorithms which include JADE [23] (joint approximate diagonalization of eigenmatrices), FastICA [24] and OGWE [25] (optimized generalized weighted estimator). It is important to know that being based on different independence criteria, each algorithm may lead to slightly different results.

In fact, every multidimensional recorded signal which can be considered as a combination of primary independent signals (sources) can be treated by ICA, which tries to extract these independent sources and thus estimates the linear combination, i.e. a mixing matrix, which leads to the observed multidimensional signal. Consequently, recorded signals in LC can be regarded as the sum of independent signals that constitute a DAD-chromatogram (i.e. peaks, noise, baseline drift,…). The use of UV-DAD detection and thus the acquisition of multi-wavelength data fulfil the requirements to consider ICA numerical separation of such chromatograms. It is thus very interesting to be able to separate peaks from noise or drift, and co-eluted peaks can be numerically separated for further processing. To carry out the numerical treatment in an automated way, it is advisable to find a methodology which allows the use of ICA in a generic manner, i.e., to find the number of components/sources that ICA will try to make independent. In this paper, our objective is thus to investigate the possibility to perform automatically the pre-treatment of data, independent component analysis, classification of independent components, peak detection, extraction of UV spectra for each component, and reconstruction of cleaned DAD-chromatograms.

2. Methodology

2.1. Independent component analysis

Different definitions can be given for ICA. The most classical is referred to as noise-free or noiseless ICA [22,24,26] whose simplified summary of its application on UV–DAD chromatograms is given below. In UV-DAD chromatography, the recorded data are matrices, $\boldsymbol{X}$, whose dimensions are $T \times J$, where T is the analysis time divided by the acquisition frequency and J is the number of wavelengths. Therefore each matrix can be expressed as a random vector, $\boldsymbol{x} = (x_1, \ldots, x_j, \ldots, x_J)$ where $x_j = x_j(t)$ represents the absorbance values at wavelength j as a function of the time t. Then ICA is carried out on $\boldsymbol{x}$ and consists of estimating the following generative model:

$$\boldsymbol{x} = \boldsymbol{A}\,\boldsymbol{s} \tag{1}$$

where $\boldsymbol{A}$ is a constant ($J \times n$) mixing matrix, with n the number of sources to be computed and $\boldsymbol{s}$ is the sources random vector $\boldsymbol{s} = (s_1, \ldots, s_n)$. Both $\boldsymbol{A}$ and $\boldsymbol{s}$ must be estimated and n must be chosen. To solve this problem, an unmixing matrix, $\boldsymbol{W}$, is introduced and the ICA algorithm

estimates it such that the elements of the vector ***s*** are as statistically independent as possible.

$$\boldsymbol{s} = \boldsymbol{W}\,\boldsymbol{x} \tag{2}$$

An important step consists in normalizing sources to avoid infinity of solutions, when estimating ***W***. The *i*th component, c_i, whose dimensions are *T*×*J*, is then defined as the product between the *i*th row, a_i, of the estimated mixing matrix ***A*** and the *i*th element of the estimated vector of sources ***s*** : $c_i = s_i a_i$. The sum of the *n* components is then equal to ***x***, $\boldsymbol{x} = \sum_{i=1}^{n} s_i a_i$. It is important to bear in mind that the term source designates the vectors on which are calculated the independence and which are used to calculate the high-order moments and statistics. The term component indicates the vectors some of which can be assigned to chemical compounds and which are used to reconstruct "cleaned" chromatograms. In this paper, no modifications were made to the original FastICA algorithm. More details on this algorithm can be found in the book of Hyvärinen et al. [27].

2.2. High-order moments and statistics

Since noise usually follows a Normal distribution, the computed independent sources given by ICA are investigated in order to find which one may be considered as noise or not. It has been described in the literature that the kurtosis of a distribution is a good criterion to reject artefacts and noise [15]. The use of kurtosis and other statistics or moments of the sources was used in the present study.

Thus, the following statistics or moments can be used to check the normality of the distribution of each source. If normality is not observed, it should imply that the source is unlikely to be noise and therefore likely corresponds to an exogenous phenomenon such as a compound being detected. In summary, sources of interest such as peaks are assumed to have a non-normal distribution, and will be identified as such.

2.2.1. Kurtosis

The kurtosis is the fourth standardized moment of a random variable and is a measurement of the peakedness (i.e. the flattening of a distribution) of the probability distribution of this variable, and can be estimated as:

$$K_i = \frac{m_4}{m_2^2} - 3 = \frac{\frac{1}{n}\sum_{t=1}^{T}(s_i(t) - \bar{s}_i)^4}{\left(\frac{1}{n}\sum_{t=1}^{T}(s_i(t) - \bar{s}_i)^2\right)^2} - 3 \tag{3}$$

m_k represents the moment (about the mean) of order *k* and $\bar{s}_i$ is the mean of the *i*th source s_i.

2.2.2. Shapiro–Wilk statistic

The non-parametric Shapiro–Wilk (S–W) test examines the null hypothesis that a sample $s_i(1)$, . . ., $s_i(T)$, comes from a Normally distributed population. This hypothesis is rejected if the test statistic is too small. This test statistic or the associated p-value can be used to characterize the distribution. More details about this test can be found in the literature [28].

Kurtosis and S–W statistics can then be used to describe the degree of normality of a distribution and, hence, to identify noise sources. However, these are far from the only statistics that allow this and other simple or complex statistics (skewness of the distribution of the source, range, Kolmogorov–Smirnov Normality test, etc.) can also be used in this context.

2.3. k-Means clustering

k-Means is a non-supervised method to cluster objects into *k* partitions on the basis of their attributes. The objects are the sources computed by ICA and the attributes are the estimated statistics computed on these sources, namely the kurtosis and Shapiro–Wilk statistic. The objective is to use these computed characteristics to discriminate the relevant sources (peaks) from sources representing noise or irrelevant artefacts, chromatogram by chromatogram. The concept is to compare the Euclidian distance between the objects. A short distance (slight difference in the computed attributes) is a sign of closeness between objects and the closest objects are put together in the same cluster.

Different implementations exist for *k*-means clustering. The Hartigan and Wong algorithm is used here, because it generally provides better results than other implementations [29]. However, this application of clustering is rather simple because it implies a low number of variables (i.e. Kurtosis and S–W) and observations. The clustering results will then not depend on the implementation selected but on the data distribution.

In practice, *k*-means with $k = 2$ is used. The partitioning should result in two clusters, one with sources having a high Kurtosis and low S–W values corresponding to the components of interest and conversely, another with low Kurtosis and high S–W value corresponding to noise and irrelevant artefact components. The outcome of the k-means clustering algorithm may vary from one run to the other because the *k* centres are randomly placed at the start of the algorithm. To avoid this dependence on initial conditions, the algorithm is repeated a hundred times for each data set (coming from one chromatogram) and the selected clusters are those with the highest frequency.

2.4. Peak detection methodology

This section presents the methodology to detect peaks for a given value of *n*, the number of estimated sources. The DAD-chromatograms are referenced with the index *p* (p=1,…, P) corresponding, for example, to *P* experimental conditions. As the treatment for

each DAD-chromatogram is the same, the following explanations are given for an unspecified *p*th chromatogram. First, data for DAD-chromatograms are truncated at 14, 24, and 34 min for overall run times of 16, 26, and 36 min, respectively. This was performed in order to remove the perturbation of the end of the gradient containing only noise or irrelevant artefacts. Second, the ICA algorithm is applied on the resulting DAD-chromatogram and high-order statistics (i.e. Kurtosis and S–W) are computed for each of then sources. Before applying clustering, attributes are divided by their standard deviation computed from all the sources available in order to give each attribute the same weight in the clustering decision process. This is illustrated by the following equation, applied to kurtosis. The same computation was done for each considered statistic.

$$K_{Si} = \frac{K_i}{\sqrt{\sum_{i=1}^{n}(K_i - \overline{K})^2}} \quad (4)$$

where K_{Si} is the standardized Kurtosis value for the *i*th source and $\overline{K}$ is the mean kurtosis computed from all the sources of one DAD-chromatogram. Third, *k*-means clustering (with *k* = 2) is applied on this bi-dimensional distribution. Fourth, the number of sources found in the relevant cluster for the *p*th DAD-chromatogram is defined as $c_r(n,p)$, as it also depends on *n*. These $c_r(n,p)$ relevant sources correspond to the $c_r(n,p)$ detected peaks in the *p*th DAD-chromatogram. Finally, as this methodology is applied on each of the *P* DAD-chromatograms, $c_r^*(n)$ is defined as the mode of the $c_r(n,p)$ distribution and corresponds to the most probable value of $c_r(n,p)$.

2.5. Automated adjustment of *n* methodology

As previously explained, the main parameter of the ICA is the number, *n*, of sources chosen to make up the signal. By default, it can be arbitrarily chosen by the operator if the number of compounds in the sample is known. In the present section, an automated process to adjust the number of sources is presented when the number of compounds to identify is unknown. In a DOE framework, the same mixture is injected several times using different analytical conditions. Thus, each DAD-chromatogram should contain the same information, i.e. the same number of relevant sources. Finding an optimal value of *n* for the application of ICA is similar to finding the most likely number of sources counted in the relevant cluster for all DAD-chromatograms. This method is presented in Fig. 1. An initial value for *n* must first be given. One can use an estimated number of pure components, trough signal matrix singular value ratios computation [30], for instance. Another easier technique is to begin the ICA computation with a very low number of sources, such as *n* = 3. Then, the methodology presented in the previous section is applied on each of the *P* DAD-chromatograms by incrementing the value of *n*, and then each value of $c_r^*(n)$ is recorded. It is possible to plot $c_r^*(n)$ against *n* (see Fig. 5). One can observe that $c_r^*(n)$ stabilizes when *n* increases. The smallest value of *n* when $c_r^*(n)$ reaches a plateau is defined as the optimal number of sources, n_{opt}, to carry out ICA. The value of $c_r^*(n)$ at the ceiling is defined as c_c^* and will be

assumed to be the optimal estimation of the number of compounds in the mixture. The optimal number of sources, n_{opt}, maximizes the probability to have (for any chromatogram) $c_r(n,p) = c_c^*$ and highly depends on the number of compounds in the sample.

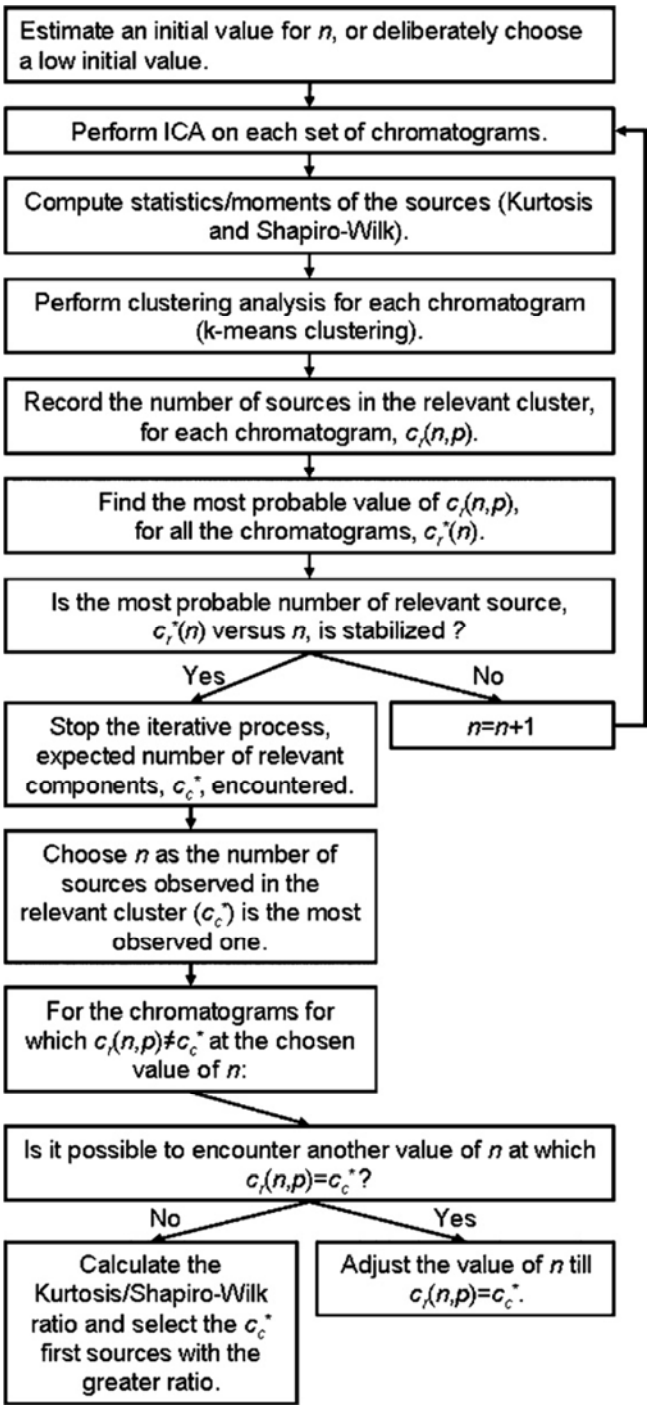


Fig. 1 Methodology for the selection of number of sources for the ICA computation.

3. Experimental

3.1. Chemicals

Methanol and acetonitrile were HPLC grade from Sigma (St. Louis, MO, USA). Ultra–pure water was obtained with a Millipore (Billerica, MA, USA) Milli-Q Academic A10. In the present study, two test mixtures were selected in order to develop and to test the presented algorithms and methodologies. These mixtures mimic an unknown mixture. The first mixture contains atenolol, phenytoin, sulfinpyrazone and warfarin and was used to develop the methodology. The second mixture was used to put the methodology to the test. It contains atenolol, pindolol, warfarin, indoprofen, naproxen, propranolol, retinoic acid, an impurity of the retinoic acid and an unidentified compound. These mixtures and compounds were obtained from the Eli Lilly Company (Indianapolis, IN, USA). Acetic acid (>98%) was purchased from Merck (Darmstadt, Germany), ammonium formate (99%) was purchased from Alfa Aesar (Karlsruhe, Germany) and ammonium hydrogencarbonate (99.7%) was purchased from VWR (Fontenay-sous-Bois, France).

3.2. Sample preparation

The first mixture (sample 1) was prepared as follow. 35 mg of atenolol, 35 mg of phenytoin, 15 mg of sulfinpyrazone and 15 mg of warfarin were dissolved in a 10 ml volumetric flask with a mixture of water-methanol (50:50, v/v). The second mixture (sample 2) was obtained directly from Eli Lilly Company and was prepared as follow. 103 mg of warfarin, 97 mg of atenolol, 100 mg of pindolol, 98 mg retinoic acid, 101 mg of naproxen sodium, 99 mg of indoprofen and 103 mg of propranolol hydrochloride were dissolved in a 100 ml volumetric flask with a mixture of water-acetonitrile (50:50, v/v). Then, from these two samples, two aliquots were filtered with 0.20 µm PTFE syringe filtration disks into some vials for injection in the HPLC system. The injection volume was 2 µl for sample 1 and 0.5 µl for sample 2.

3.3. Experiments

A full factorial design [31] was applied on three HPLC factors: the pH of the aqueous part of the mobile phase (four levels from 2.6 to 10.0), the gradient time (T_G) (three levels from 10 to 30 min) and the nature of organic modifier (two levels, acetonitrile or methanol). For sample 1, this design was repeated on five analytical columns: XBridge C18, XBridge C8, XBridge RP18, XBridge Phenyl columns (100×2.1 mm i.d.; particle size 3.5µm), all from Waters (Milford, MA, USA) and a Cogent Bidentate C18 column (100×2.1mm i.d.; particle size 4.0µm) from Microsolv (Eatontown, NJ, USA). For the sample 2, the design was only carried out on the XBridge C18 with methanol for a total of 12 experimental conditions. The experiments were carried out at a flow rate of 0.25 ml/min and at 30 °C. The buffers consisted of 10 mM pH 2.6 formic acid, pH 5.0 ammonium formate, pH 7.0 ammonium

formate and pH 10.0 ammonium hydrogencarbonate. The pH was adjusted to the selected value with concentrated formic acid or ammonia 35% aqueous solution. The shapes of the linear gradients are described in Table 1.

Table 1 Linear gradient shapes with organic modifier percentage (O.M.%) in mobile phase.

T_G = 10 min		T_G = 20 min		T_G = 30 min	
Time (min)	O.M.%	Time (min)	O.M.%	Time (min)	O.M.%
0	5	0	5	0	5
10	95	20	95	30	95
10.5	95	20.5	95	30.5	95
10.6	5	20.6	5	30.6	5
16	5	26	5	36	5

Chromatographic separations were performed on a Waters 2695 separation module coupled to a Waters selector valve 7678 and a Waters 996 Photodiode array detector. All the DAD-chromatograms were recorded between 210 nm and 400 nm with an estimated step of 1.2 nm (158 points) and with an acquisition time of 500 ms. They were finally exported by Empower 1.0 (Waters) in ASCII files containing the UV-DAD data.

The experimental design leads to 2·4·3 (i.e. 2 organic modifiers, 4 pH, 3 gradient times) = 24 experiments which were repeated for the five columns which leads to a total of 120 chromatograms. It is important to state that this full factorial design was not only set up to challenge ICA abilities to analyse DAD-chromatograms recorded under various conditions but also to develop and to test new mathematical models of retention times [32]. Such a large set of experiments is not required for ICA and for conventional or automated optimization of analytical methods. In most complex separations, classical or D-optimal design with 12–24 experiments will largely suffice to accomplish a DOE and ICA driven separation.

3.4. Software

An in-house computer program was developed to perform the analysis presented in the previous sections. The coding was carried out with R 2.4.1 statistical language for Windows, freely distributed at http://www.r-project.org. These codes can be run on compatible PC (3GHz in this case) or other environments where R is available. For ICA computations, the FastICA algorithm for R, developed at the Helsinki University of Technology, was used.

4. Results and discussion

4.1. ICA algorithm results and components clustering

Each chromatogram matrix ***X*** can be described as a (*T*×158) matrix where *T* is either 1680, 2880 or 4020, depending on the gradient time (10, 20 and 30 min, respectively). In this section, the results obtained (for sample 1) by applying the methodology described in Section 2.4, for one chromatogram, are presented. A summary of the results obtained for sample 2 can be found in Section 4.7. The arbitrarily selected value for *n* is equal to 12 and the chosen DAD-chromatogram is the one recorded with sample 1 on the XBridge C18 column with methanol and buffer at pH 7.0 with a gradient time of 10 min (Fig. 2(a)). In fact, in this case *n* = 12 is a good value for *n* because it succeeds in finding the four compounds and the dead volume perturbation (i.e. the small baseline perturbation observed at t_0) of the considered DAD-chromatogram.

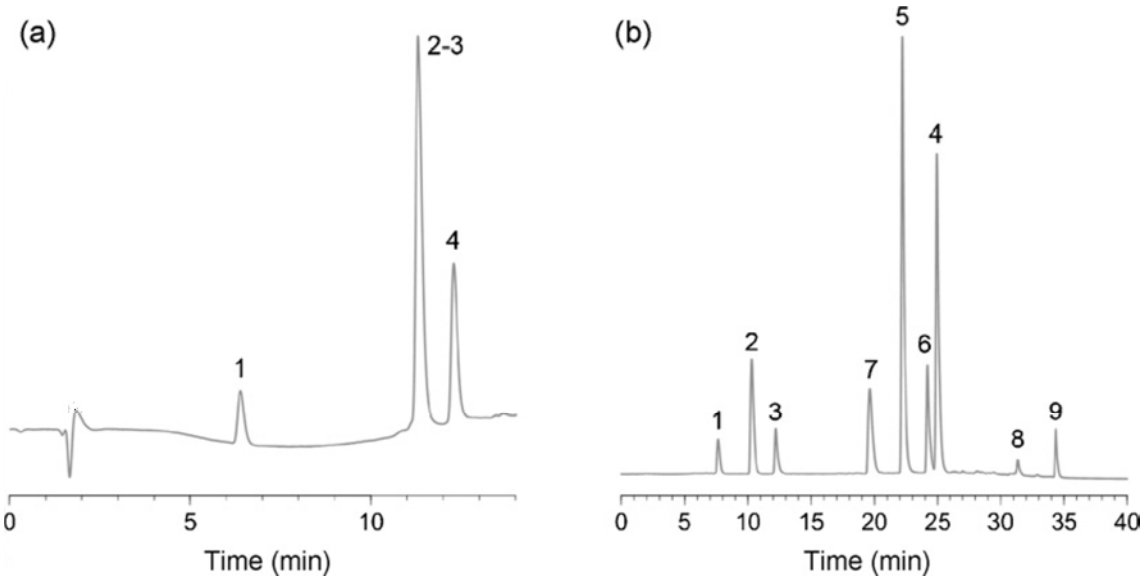


Fig. 2 (a) Chromatogram (sample 1) recorded on an XBridge C18 with methanol and buffer at pH7.0 with a gradient time of 10 min observed at 240 nm. Compounds numbered from 1 to 4 are respectively atenolol, phenytoin, sulfinpyrazone and warfarin. (b) Chromatogram (sample 2) recorded on an XBridge C18 with methanol and buffer pH 2.9 with a gradient time of 30 min observed at 280 nm. Compounds numbered from 1 to 9 are respectively atenolol, pindolol, an unidentified compound, warfarin, indoprofen, naproxen, propranolol, an impurity of the retinoic acid and the retinoic acid.

The twelve independent components are depicted in Fig. 3 at the wavelength of 240 nm. At this wavelength, each peak absorbs and gives approximately the same signal to noise ratio. The components corresponding to peaks are labelled by a surrounding square (sources 3, 5, 7, 11 and 12). The seven other components correspond to noise and are of much lower amplitude.

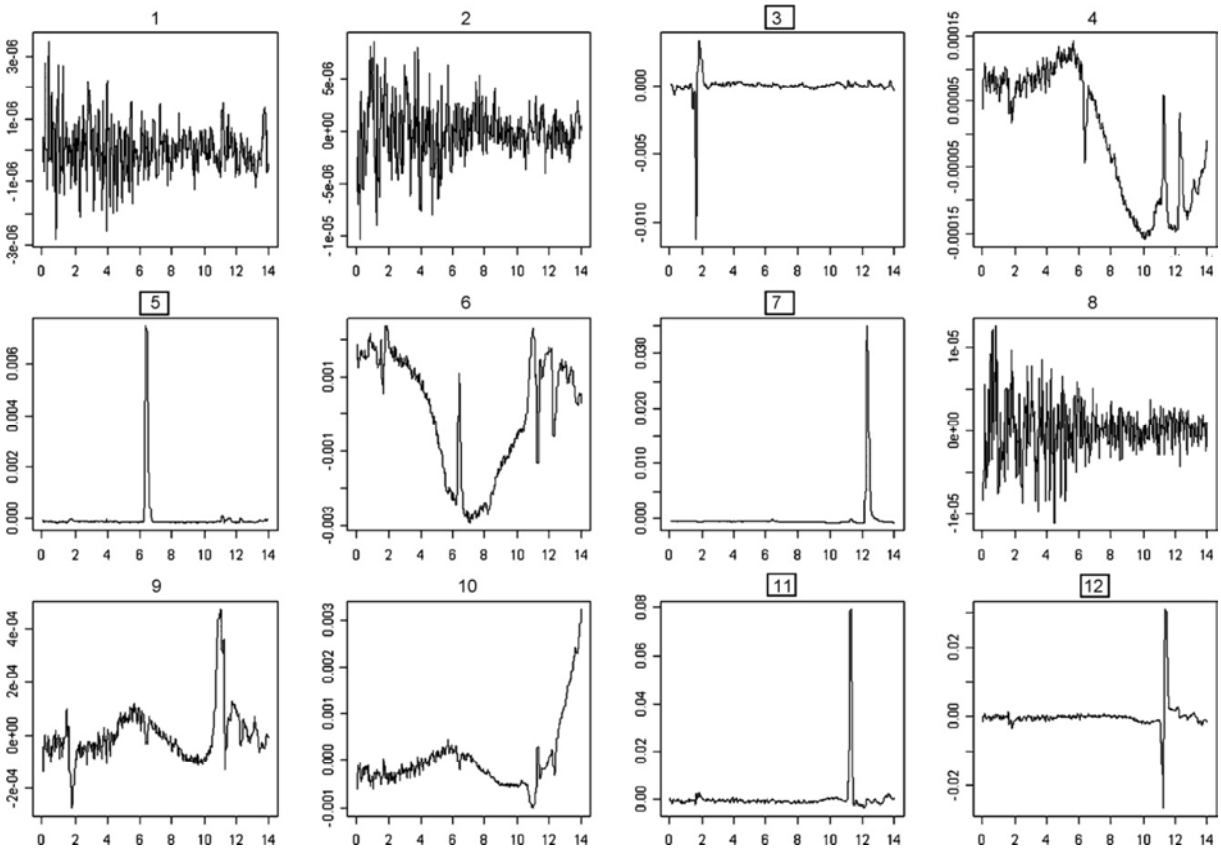


Fig. 3 Components computed by ICA with n = 12 and at 240 nm. The components whose numbers are labelled with a square are the ones that correspond to peaks. Other components correspond to noise or irrelevant artefacts. The X-axis represents the time (min) and the Y-axis is the absorbance.

First, one can observe that the 12thcomponent is distorted and it seems to issue from the strong coelution with the eleventh component. This type of distortion was observed an each set of coeluting peaks for all the chromatograms. A similar distortion is not visible on the 11th component which is probably due to its higher size. However, the distortions which can exist from components which result from coeluted peaks do not seem to affect the detection of these peaks even if their size ratio is about 2.5:1 in this case. Second, some components which were classified as corresponding to noise depict small peaks at the same retention times as the relevant components (e.g., component 6). This results from the fact that ICA algorithm is seldom able to generate components which are totally independent. It apparently does not affect peak detection either. The classification of components determined from k-means clustering is shown in Fig. 4(a). The crosses indicate the centre of the clusters and the relevant one (grey ellipse) is located at the bottom right and contains five components ($c_r(n,p)$ = 5). The apex of each relevant component has been found and its retention time has been automatically placed on the chromatogram (Fig. 4(b)).

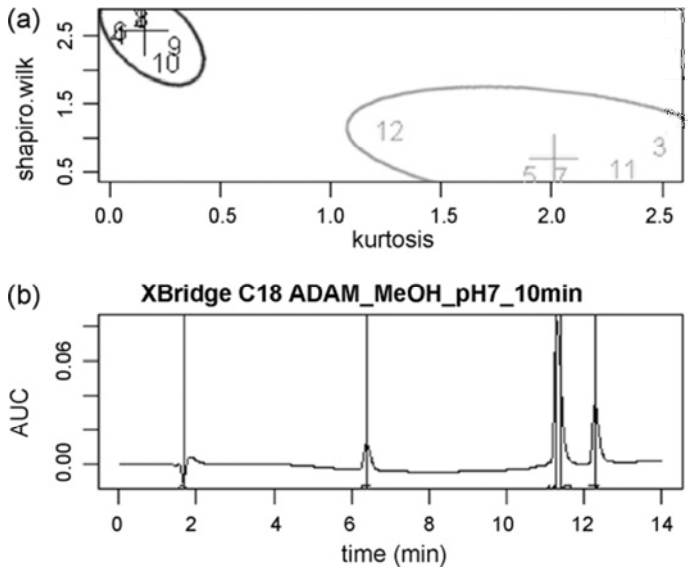


Fig. 4 (a) Clustering k-means realized on standardized kurtosis and Shapiro–Wilk statistics. The right-bottom cluster (cluster 1, in grey) contains the relevant sources. Crosses indicate the clusters centres. (b) Original DAD-chromatogram at 240 nm with automatically picked apexes by a vertical line.

4.2. Optimal number of sources

When the 120 chromatograms are treated as in the previous section, a distribution of 120 values of $c_r(n,p)$ is obtained. With n = 12, the mode of this distribution is equal to five ($c_r^*(n)$ = 5) and its observed frequency is equal to 37.5%. The distribution of $c_r(n,p)$ obtained with n = 20 strengthens the fact that (from a given value) when n increases the mode of the distribution is not modified (still equal to 5 with a frequency of 29.2%).

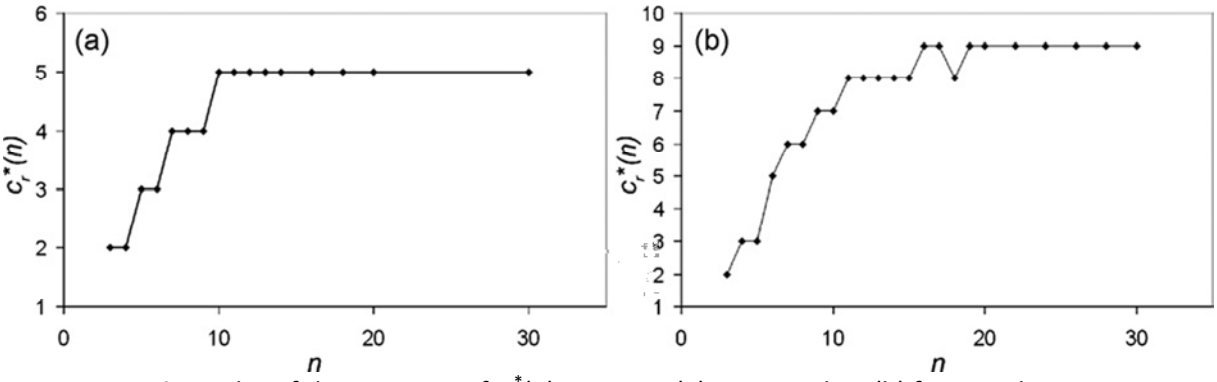


Fig. 5 Plot of the variation of $c_r^*(n)$ versus n. (a) For sample 1 (b) for sample 2.

To find the optimal value of sources which make up the chromatogram, n was incremented from 3 to 30. The variation of $c_r^*(n)$ against n is plotted in Fig. 5(a). This figure clearly shows first, an increasing variation and second, a stabilization of relevant components number, $c_r^*(n)$. The cap of $c_r^*(n)$ is observed at the value of 5 (c_c^* = 5) and starts with a value of n equal to twelve, so n_{opt} = 12. Therefore, this result indicates that the sample seemingly contains five independent sources of interest (five compounds or four compounds and the dead volume baseline perturbation).

4.3. Final adjustment of *n* and Kurtosis/Shapiro–Wilk ratio

However, for $n = n_{opt}$, $c_r(n,p)$ differs from c_c^* in 62.5% of the clustering results. It means that for 62.5% of the results (75 chromatograms), the number of components counted in the relevant cluster is not equal to five. Then, for these inconvenient results, the value of n is adjusted till $c_r(n,p)$ reaches c_c^* whose the computation is explained in the previous section. However, for some chromatograms, when n is incremented, $c_r(n,p)$ never equals five. So, the components are ranked in descending order accordingly to their K_S/S–W ratio. Indeed, if this ratio is high then the corresponding components are close to the bottom right corner of the clustering plot, which means that the probability for this component to belong to the relevant cluster is higher. So, the c_c^* components with the higher ratios are defined as being the relevant ones.

4.4. Results of automated peak detection

For the 120 DAD-chromatograms, a total of 516 components are classified as relevant, 457 of whom are detected peaks and 59 correspond to the detection of the dead volume baseline perturbation. As the dead volume baseline perturbation is detected in about 50% of the chromatograms and as the foretold number of relevant components by chromatogram is equal to five ($c_c^* = 5$), then it means that among these five components stands the dead volume baseline perturbation, so the number of expected peaks in the mixture is equal to 4. As 457 peaks are detected on the 120 chromatograms, 95.2% of the peaks are automatically detected. To confirm these results, peak detection was done manually and four peaks were detected for each chromatogram. However, even done by an experienced analyst, the manual peak detection cannot always be error free (but we assume it is error free due to the small number of compounds in the mixture). The dead volume baseline perturbation is of particular interest (defining the dead time of a HPLC system), but the detection of this relevant component is not consistent. This can be due to its random shape when using different mobile phase as in DOE, to its small size and also to its undifferentiated UV signature with the mobile phase.

4.5. Extracting component UV-spectra

Finally, the identification of the compounds can be done manually. In the case of phenytoin and sulfinpyrazone, which are coeluting (Fig. 2(a)), the UV spectra extracted from the original chromatogram are mixtures of the “pure” spectra (i.e. the spectra that will be obtained if the compounds were not coeluting). As the obtained independent components have the same dimensions than the original chromatogram, UV spectra can be extracted from the respective independent components. These spectra are good estimations of the “pure” spectra and the compounds identification is made easier. Fig. 6(b) depicts, for one chromatogram, the UV spectra estimated by ICA extracted from the corresponding relevant component (which are displayed on Fig. 6(a)) at the apex of the peak. These extracted UV

spectra can be compared to the UV spectra coming from the original chromatogram (Fig. 6(c)). One can observe that some slight differences exist between the extracted and the original UV spectra. It is probably due to the interference of the mobile phase (organic modifier, buffer composition and buffer pH) and to the strong coelution existing between phenytoin and sulfinpyrazone.

4.6. Reconstruction of DAD-chromatogram

As the sum of the *n* components gives (approximately) the original chromatogram, the sum of the $c_r(n,p)$ relevant components should give a chromatogram free from noise and artefacts represented by the other components. Fig. 7(a) illustrates the original DAD-chromatogram compared to the reconstructed one (Fig. 7(c)), both observed at 240 nm. The peaks (identified by the triangles in the top of the chromatograms) have been automatically detected. However, at this wavelength, the noise reduction of the chromatogram is not really observable. The chromatograms of Fig. 7(b) and (d) illustrates that, for other wavelengths, the reconstructed chromatogram (d) has been cleaned from noise, irrelevant artefacts and gradient baseline drift.

4.7. Test of the methodology with a more complex sample

Sample 2was directly shipped from Eli Lilly Company and seven compounds were described in the sample preparation procedure. This mixture was used to test the ability of the methodology to detect peaks in a more complex mixture. The global methodology was exactly applied in the same way as for sample 1. As depicted in Fig. 2(b), nine compounds were manually detected. Thus the mixture contained at least nine compounds which was more than foreseen. The analysis of the variation of $c_r^*(n)$ against *n*, which is displayed on Fig. 5(b), shows that the cap of $c_r^*(n)$ is observed at the value of 9 ($c_c^* = 9$) and starts with a value of *n* equal to nineteen, so $n_{opt} = 19$. As the baseline perturbation at t_0 was not detected, this result confirms that nine compounds were contained in sample 2. Among the two unexpected compounds, one is still unidentified and the second has been identified as an impurity of the retinoic acid. It seems that the baseline perturbation at t_0 was not detected due the really small injected volume (0.5 µl). For the twelve recorded DAD-chromatograms, a total of 103 components were classified as relevant. As the number of compounds in the mixture is equal to nine, it means that 95.4% of the peaks were automatically detected. Fig. 8(a) depicts a chromatogram where compound 7 and compound 8 are coeluting. Despite this coelution and the difference between the peaks size, ICA was able to differentiate and to detect these compounds as displayed on Fig. 8(b). Furthermore, these results obtained with sample 2 tend to confirm that the percentage of detected peak (about 95%) seems not to be affected by the sample complexity.

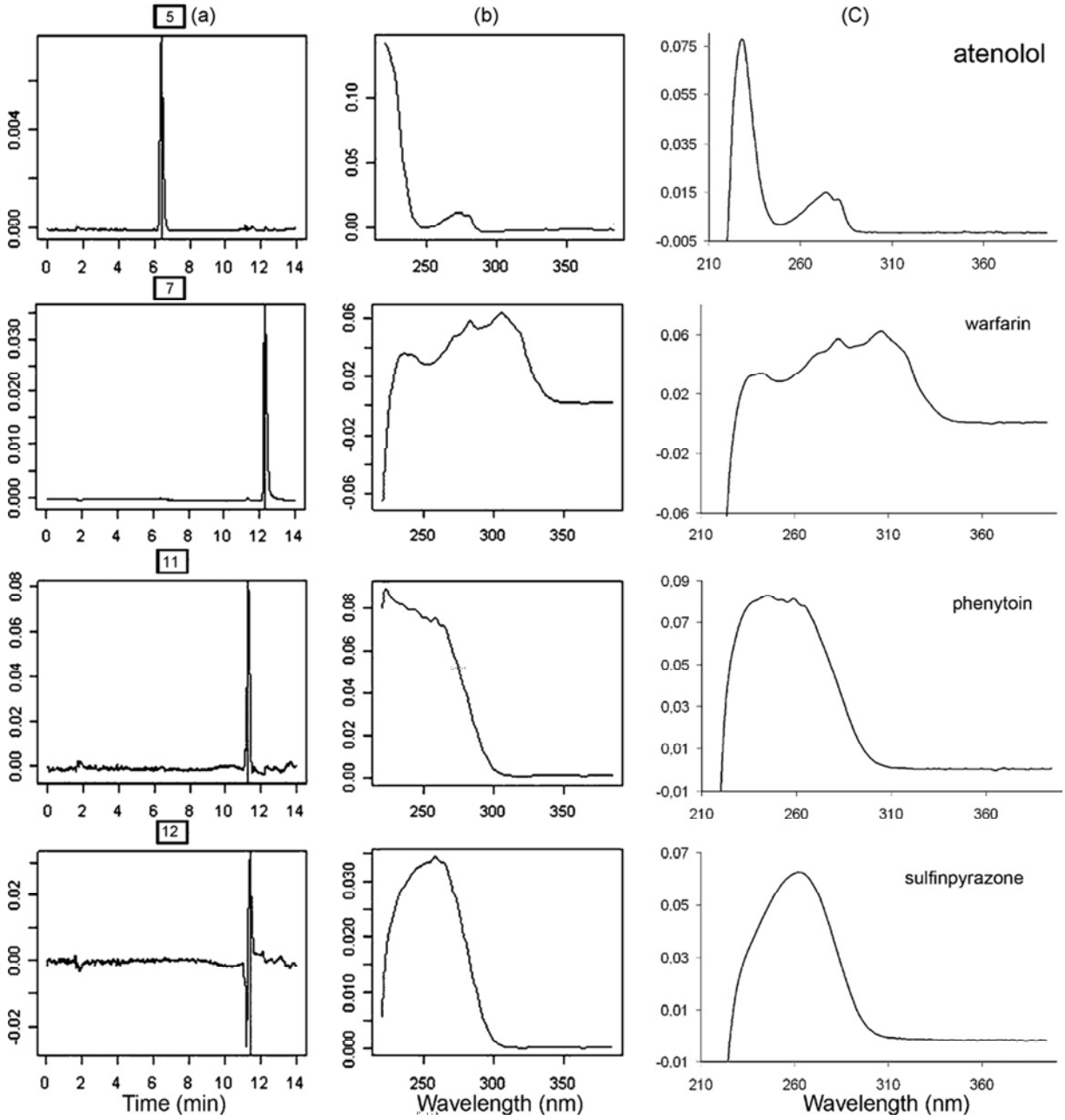


Fig. 6 Comparison between the obtained relevant components UV-spectra and the reference ones. (a) Components contained in the relevant cluster, observed at 240 nm; (b) estimated UV spectra of each independent component at apex (210–400 nm); (c) UV spectra (210–400 nm) as observed in the original chromatogram. Those spectra were recorded from chromatograms obtained from each individual reference solution using the same experimental condition.

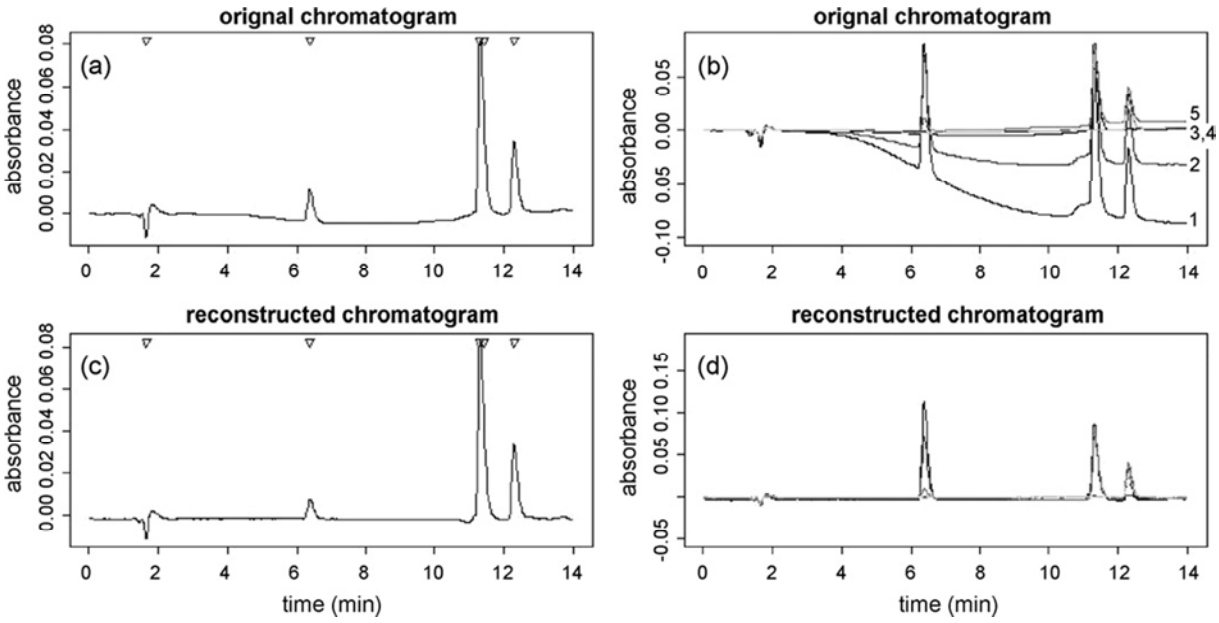


Fig. 7 (a–c) Comparison between original DAD-chromatogram (a) and reconstructed one with *n* = 12 (c). The original is recorded at 240 nm, on an XBridge C18 with methanol and pH 5 buffer with a gradient of 20 min. Apexes have been automatically picked. (b–d) Comparison at different wavelengths (1: 214, 2: 220, 3: 240, 4: 268 and 5: 327 nm) of the same original (b) and reconstructed (d) chromatograms.

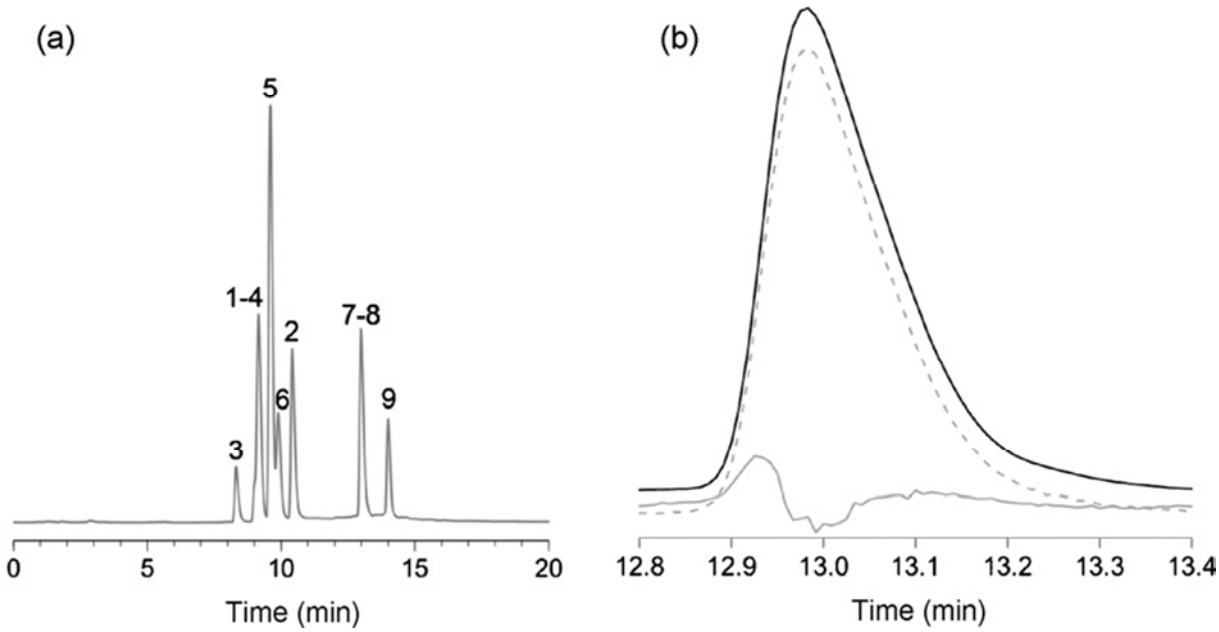


Fig. 8 (a) One chromatogram obtained (with sample 2) in the experimental design depicted at 280 nm. Nine compounds were actually automatically detected. (b) Magnification of the zone between 12.8 and 13.4 min. The solid black line represents the original recorded chromatogram, the dashed grey line depict the component corresponding to compound 7 (propranolol) and the solid grey line display to the component corresponding to compound 8 (an impurity of the retinoic acid).

5. Conclusions

The automated detection of peaks is a very crucial step in the automated development of analytical methods. A new and original approach combining ICA, high-order statistics and clustering was successfully used for the determination of a test mixture of pharmaceutical compounds in the framework of a DOE methodology. The present approach could be envisaged in high throughput screening experiments, such as those resulting from synthetic process development (purity assessment) or for the development of stability indicating methods when impurities are not necessarily known. Moreover, it does not require expensive equipment, such as a mass spectrometer, to detect the compounds, as long as the compounds absorb in the UV. Clustering methods allow separating very efficiently the noise components from the relevant ones, using adequate summary statistics. The technique to find an optimal number of sources is very convenient but is time consuming. Fortunately, the time needed for the numerical data treatments presented in this study is much smaller than the sample analysis time. Indeed, the required times to obtain the ICA results, for one chromatogram, when *n* is incremented from 3 to 30 as shown in this paper, are 35, 55 and 75 seconds for the respective gradient times of 10, 20 and 30 min. Therefore, as the chromatograms are numerically treated one by one, this gives the opportunity to easily implement the numerical data treatments in a semi-concurrent mode, i.e. just after the sample analysis.

On the other hand, this process could also be performed only on sub-parts of a DAD-chromatogram, according to the analyst interest; for instance, searching for coeluted impurities in peaks of interest allowing a significant reduction in the time devoted to the computational process.

Eventually, the combination of the original aforesaid strategies described in this paper could give a powerful tool and seems to begin to open new prospects for the fully automated development of LC–UV/DAD methods. Nevertheless, further work concerning the possibilities for automated peak tracking is required.

Acknowledgements

The authors are very grateful to the reviewers for providing important comments that led to significant improvements of this article. The authors would also like to thank the Walloon Region of Belgium for the FIRST-DEI convention funds no. 516130 and the Eli Lilly Company for partial funding of ADAM project and equipments. A research grant from the Belgium National Fund for Scientific Research (FRS-FNRS) to E. Rozet is also gratefully acknowledged.

References

[1] J.R. Torres-Lapasio, M.C. Garcia-Alvarez-Coque, J. Chromatogr. A 1120 (2006) 308.

[2] J. Krupcik, J. Mydlova, I. Spanik, B. Tienpont, P. Sandra, J. Chromatogr. A 1084 (2004) 80.

[3] G. Vivó-Truyols, J.R. Torres-Lapasió, M.C. García-Alvarez-Coque, J. Chromatogr. A 991 (2003) 47.

[4] V.B. di Marco, G.G. Bombi, J. Chromatogr. A 931 (2001) 1.

[5] P. Nikitas, A. Pappa-Louisi, A. Papageorgiou, J. Chromatogr. A 912 (2001) 13.

[6] J.R. Torres-Lapasio, J.J. Baeza-Baeza, M.C. Garcia-Alvarez-Coque, Anal. Chem. 69 (1997) 3822.

[7] G. Vivó-Truyols, J.R. Torres-Lapasió, M.C. García-Alvarez-Coque, P.J. Schoenmakers, J. Chromatogr. A 1158 (2007) 258.

[8] R. Tauler, Chemom. Intell. Lab. Syst. 30 (1995) 133.

[9] A. Bogomolov, M. McBrien, Anal. Chim. Acta 490 (2003) 41.

[10] A. de Juan, R. Tauler, J. Chromatogr. A 1158 (2007) 184.

[11] G.Wang, Q. Ding, Z. Hou, TrAC, Trends Anal. Chem. 27 (2008) 368.

[12] S. Makeig, T. Jung, A.J. Bell, D. Ghahremani, T.J. Sejnowski, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 10979.

[13] T.-P. Jung, C. Humphries, T.-W. Lee, S. Makeig, M. McKeown, V. Iragui, T.J. Sejnowski, Adv. Neural Inform. Process. Syst. 10 (1998) 894.

[14] A. Delorme, T.J. Sejnowski, S. Makeig, NeuroImage 34 (2007) 1443.

[15] M. McKeown, S. Makeig, G. Brown, T.-P. Jung, S. Kindermann, T.W. Lee, T.J. Sejnowski, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 803.

[16] H. Yamamoto, K. Hada, H. Yamaji, T. Katsuda, H. Ohno, H. Fukuda, Biochem. Eng. J. 32 (2006) 149.

[17] G.Wang, W. Cai, X. Shao, Chemom. Intell. Lab. Syst. 82 (2006) 137.

[18] M. Vosough, Anal. Chim. Acta 598 (2007) 219.

[19] D. Jouan-Rimbaud Bouveresse, H. Benabid, D.N. Rutledge, Anal. Chim. Acta 589 (2007) 216.

[20] D. Mantini, F. Petrucci, P. Del Boccio, D. Pieragostino, M. Di Nicola, A. Lugaresi, G. Federici, P. Sacchetta, C. Di Ilio, A. Urbani, Bioinformatics 24 (2008) 63.

[21] S. Wienkoop, K. Morgenthal, F. Wolschin, M. Scholz, J. Selbig, W. Weckwerth, Mol. Cell. Proteomics 7 (2008) 1725.

[22] A. Hyvärinen, E. Oja, Neural Networks 13 (2000) 411.

[23] J.-F. Cardoso, Neural Comput. 11 (1999) 157.

[24] A. Hyvärinen, E. Oja, Neural Comput. 9 (1999) 1483.

[25] J.J. Murillo-Fuentes, R. Boloix-Tortosa, F.J. Gonzalez-Serrano, Fourth International Symposium on Independent Component Analysis and Blind Signal Separation [ICA2003], 2003, p. 1053.

[26] A. Hyvärinen, IEEE Trans. Neural Networks 10 (1999) 626.

[27] A. Hyvärinen, J. Karhunen, E. Oja, Independent Component Analysis, Wiley, New York, 2001.
[28] S.S. Shapiro, M.B. Wilk, Biometrika 52 (1965) 591.
[29] J.A. Hartigan, M.A.Wong, Appl. Statist. 28 (1979) 100.
[30] H. Yuzhu, S. Weiyang, Y. Weifeng, D.L. Massart, Chemom. Intell. Lab. Syst. 77 (2005) 97.
[31] G. Cox,W. Cochran, Experimental Designs, 2nd ed., McGraw-Hill, 1962.
[32] P. Lebrun, B. Govaerts, B. Debrus, A. Ceccato, G. Caliaro, Ph. Hubert, B. Boulanger, Chemom. Intell. Lab. Syst. 91 (2008) 4.

Section III.1.3

Implémentation de la méthodologie ICA au DoE-DS pour l'optimisation robuste de méthodes chromatographiques

Sous-section A

B. Debrus, P. Lebrun, E. Rozet, I. Nistor, A. Ceccato, G. Caliaro, R. Oprean, B. Boulanger, Ph. Hubert, Nouvelle méthodologie pour le développement automatisé de méthodes analytiques en chromatographie liquide pour l'analyse de mélanges de composés inconnus, Spectra Analyse 268 (2009) 28.

Résumé

De nos jours, de nombreuses stratégies d'optimisation de méthodes chromatographiques sont disponibles. Néanmoins, le développement de méthodes chromatographiques reste l'étape la plus limitante dans les processus de synthèse ou d'identification de nouvelles molécules pouvant conduire à des agents thérapeutiques ou à de nouveaux biomarqueurs malgré la disponibilité de nouvelles technologies tant en chimie (chimie combinatoire, *High throughput screening*…) qu'en biochimie analytique (protéomique, métabolomique, *herbal fingerprinting*…). L'objectif de l'étude présentée dans ces pages est d'éprouver une nouvelle méthodologie de développement automatisé de méthodes chromatographiques combinant la planification expérimentale, l'analyse en composantes indépendantes, l'analyse de la propagation de l'erreur prédictive et la modélisation par régression linéaire multiple. Finalement, cette méthodologie automatisée a permis de séparer avec succès les composés d'un mélange inconnu.

Summary

Nowadays, many strategies to optimize chromatographic methods are available. However, the development of chromatographic methods remains the most limiting step in the process of synthesis or identification of new molecules that could lead to therapeutic agents or new biomarkers despite the availability of new technologies both in chemistry (chemical combinatorial, high throughput screening...) and in analytical biochemistry (proteomics, metabolomics, herbal fingerprinting...). Therefore, the aim of this study is to test a new methodology for developing automated chromatographic methods combining experimental planning, independent component analysis, analysis of predictive error propagation and multiple linear regression modelling. Finally, this automated methodology has enabled us to successfully separate the components of an unknown mixture.

1. Introduction

Dans le domaine pharmaceutique, les méthodes analytiques de quantification occupent une place fondamentale à plus d'un égard. En effet, de nouvelles méthodes et technologies tant en chimie (chimie combinatoire, High throughput screening…) qu'en biochimie analytique (protéomique, métabolomique, herbal fingerprinting…) mènent aujourd'hui à la synthèse ou à l'identification d'un nombre considérable de nouvelles molécules pouvant conduire à des agents thérapeutiques potentiels ou de nouveaux biomarqueurs d'un état pathologique. A cette fin, les techniques séparatives et plus particulièrement la chromatographie liquide haute performance (HPLC) sont depuis plusieurs années des techniques de choix pour identifier et quantifier ces molécules d'intérêts. Néanmoins, les paramètres régissant la qualité de la séparation des composés contenus dans un échantillon sont multiples et la mise au point d'une méthode HPLC peut parfois être laborieuse surtout quand la nature de ces molécules est inconnue. De plus, parmi les méthodes d'optimisation existantes [1-4], certaines peuvent requérir des informations préliminaires à propos des propriétés physico-chimiques des composés étudiés et de ce fait limiter l'automatisation des processus d'optimisation dans le cas de mélanges inconnus. Dans cet article est présentée une méthodologie novatrice [5,6] s'acquittant de prérequis et permettant d'optimiser la séparation d'un mélange de composé inconnu.

2. Stratégie d'optimisation

2.1. Planification expérimentale

Ces dernières années, de nombreuses avancées ont été réalisées dans le domaine du développement de méthodes analytiques. Il est aujourd'hui possible d'envisager l'utilisation de diverses stratégies afin de trouver des séparations optimales. Néanmoins, ces stratégies doivent être susceptibles de fournir des prédictions exactes et des optima robustes avant les étapes ultérieures de validation et/ou de transfert de méthodes. La planification expérimentale est l'une de ces stratégies. La première étape de la planification expérimentale consiste à sélectionner des facteurs caractérisés par un effet marqué sur la réponse qui sera modélisée. Dans l'exemple développé dans ces pages, les facteurs sélectionnés sont le pH de la partie aqueuse de la phase mobile et le temps nécessaire pour passer linéairement de 5 % à 95 % en méthanol (ce facteur est communément appelé temps de gradient, T_G). Afin de maximiser les chances d'identifier une séparation optimale, le domaine expérimental est élargi au maximum : la gamme de pH s'étend de 2,6 à 10,0, celle du temps de gradient de 10 à 30 minutes. Des études préliminaires ont démontré que les plans expérimentaux offrant de bonnes propriétés d'orthogonalité et de rotatabilité sont à privilégier lors de la modélisation de réponses chromatographiques. Les plans factoriels complets (figure 1) fournissent de meilleures prédictions par rapport aux plans D-optimaux.

Dans un second temps, il convient de sélectionner les réponses qui seront modélisées par des fonctions polynomiales. Certaines précautions doivent être prises lorsque la résolution (R_S) est utilisée comme réponse. La résolution est généralement calculée sur les deux pics les plus proches dans un chromatogramme (la paire critique). Cependant, comme les conditions opératoires comprises dans le domaine expérimental engendrent des sélectivités très différentes, l'ordre des pics change et le calcul de la résolution de la paire critique se base sur des composés différents. Ce problème d'inversion des pics induit des discontinuités dans les réponses. Dès lors, la modélisation de ces réponses grâce à des équations linéaires multiples devient non cohérente et des erreurs de prédiction très significatives peuvent en découler. De plus, postérieurement à la modélisation, seule une résolution optimale peut être prédite et aucune information supplémentaire concernant le chromatogramme optimal n'est obtenue. Certains travaux antérieurs ont permis de démontrer l'avantage d'utiliser des réponses mieux adaptées. Ainsi, la modélisation des temps au début, au sommet (temps de rétention, t_R) et à la fin de chaque pic est donc réalisée. Ces réponses présentent les avantages d'être modélisables par des fonctions polynomiales et de rendre possible la prédiction d'un chromatogramme optimal complet sur lequel peut être calculé n'importe quels critères chromatographiques (résolution, séparation entre les pics, largeur des pics…).

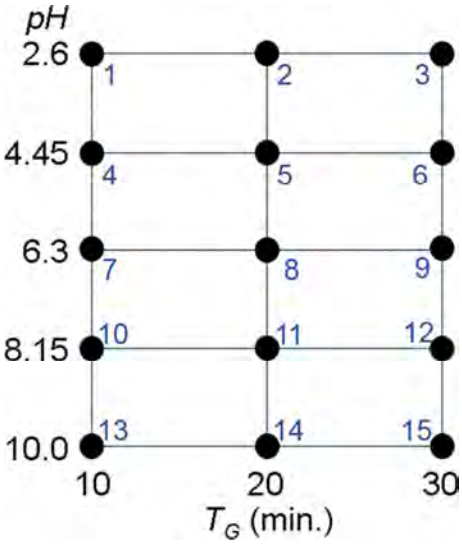


Fig. 1 Plan expérimental factoriel complet à deux facteurs. Les conditions expérimentales ont arbitrairement été numérotées en bleu.

2.2. Lecture automatique des chromatogrammes

L'utilisation de plans expérimentaux et donc de conditions opératoires drastiquement différentes induit des sélectivités très distinctes. Dès lors, la détection et l'identification des pics sont généralement des étapes délicates et qui nécessitent beaucoup de temps. L'analyse en composantes indépendantes (ACI) permet de séparer les composantes d'un chromatogramme enregistré en UV-DAD. Même coélués, les pics dont les spectres UV sont nettement différents et qui présentent des temps de rétention distincts peuvent être numériquement séparés. L'ACI est donc utilisée pour séparer les pics, les détecter et les identifier de manières automatiques. Il ne relève pas du propos de cet article de donner le détail de la méthode ACI, le lecteur intéressé par celle-ci pourra trouver dans la littérature, publications de travaux réalisés par Hyvärinen et al. [7] ou encore par notre équipe [8], de plus amples informations sur celle-ci.

2.3. Modélisation mathématique et prédiction

Les réponses modélisées sont les logarithmes des facteurs de rétention (équation 1).

$$\log(k) = \log\left(\frac{t_R - t_0}{t_0}\right) \quad (1)$$

où k est le facteur de rétention, t_R est le temps de rétention du composé étudié et t_0 est le temps de rétention pour un composé non-retenu. La modélisation mathématique de ces logarithmes est effectuée à l'aide d'équations linéaires multiples (équation 2).

$$\log(k) = \beta_0 + \beta_1 \times pH + \beta_2 \times pH^2 + \beta_3 \times pH^3 + \beta_4 \times pH^4 + \beta_5 \times T_G + \beta_6 \times {T_G}^2 + \beta_7 \times pH \times T_G + \varepsilon \quad (2)$$

où β_0,…, β_7 sont les paramètres du modèle mathématique, *pH* et T_G étant les facteurs du plan expérimental et ε est l'erreur de prédiction. Une régression pas à pas est utilisée afin de maximiser le coefficient de détermination multiple ajusté ($R^2_{\text{ajusté}}$) [9]. Le comportement chromatographique de chaque composé est modélisé par trois équations. Une modélisant le logarithme du facteur de rétention et deux modélisant les logarithmes de la demi-largeur à gauche (du début au sommet du pic) et à droite (du sommet à la fin du pic). L'erreur prédictive, dont les sources sont systématiques et aléatoires, est également analysée. Une analyse de Monte Carlo est conjointement effectuée et permet d'estimer l'ampleur que cette erreur aura sur la prédiction des réponses. Il est alors possible de connaitre la distribution des réponses pour chaque condition opératoire donnée et de propager cette erreur sur les critères (e.g. la résolution). L'utilisation de la résolution comme critère peut se révéler être inadéquat. En effet, la résolution est définie comme étant le rapport entre la différence des temps de rétention de deux pics et la moyenne des largeurs de ces pics. Lors de la propagation de l'erreur au critère, cette division peut engendrer des valeurs infinies et une erreur exagérée entache alors la prédiction. La séparation entre les pics (*S*) se définissant comme la différence entre le temps au début d'un pic et le temps à la fin du pic précédent est donc utilisée. La distribution des réponses étant connue, il est possible de calculer, à chaque condition opératoire, la probabilité que le critère atteigne un seuil prédéfini. Cette information est essentielle car elle permet d'évaluer le risque que l'écart entre valeurs prédites et observées soit trop important. Dans cette optique, plutôt que de représenter des surfaces de réponses, des surfaces de probabilités sont calculées.

On peut alors définir le Design Space (DS) (équation 3) qui est la zone dans laquelle la probabilité pour qu'un critère atteigne le seuil prédéfini soit supérieure à une valeur sélectionnée. En d'autres termes, toutes les conditions opératoires comprises à l'intérieur du DS présenteront, par exemple, une probabilité d'au moins 90 % que la séparation entre les pics les plus proches soit d'au moins 1 minute. Le DS représente une avancée majeure car il permet d'estimer la robustesse des conditions opératoires optimales.

$$\text{DS} = \left\{x_0 \in \chi | E_{\hat{\theta}}\left[\text{P}(S > \lambda)|\hat{\theta}\right] \geq \pi\right\} \quad (3)$$

où x_0 est un point du domaine expérimental, χ. λ est la limite d'acceptation pour le critère sélectionné (*S*) et π est le niveau de qualité. $\hat{\theta}$ est l'ensemble des paramètres estimés du

modèle. P représente l'estimateur de probabilité et E, l'estimateur de l'espérance mathématique.

3. Matériels et méthodes

3.1. Réactifs et échantillons

Le méthanol (HPLC gradient grade), l'acide formique (> 98 %) et l'hydrogénocarbonate d'ammonium (99,7 %) ont été fournis par VWR (Fontenay-sous-Bois, France). Le formiate d'ammonium (99 %) a été acheté chez Alfa Aesar (Karlsruhe, Allemagne). L'eau de qualité ultra-pure (18,2 MΩ·cm) a été obtenue grâce à un appareil Milli-Q Plus 185 de Millipore (Billerica, Etats-Unis). L'échantillon a gracieusement été fourni par Eli Lilly Company (Indianapolis, Etats-Unis). Le volume d'injection était de 0,5 µL. Les solutions tampons consistaient en des solutions de concentration de 10 mM en acide formique à pH 2,6 ; en formiate d'ammonium à pH 4,45 et pH 6,3 ; et en hydrogénocarbonate d'ammonium à pH 8,15 et pH 10,0. Le pH a été ajusté aux valeurs désirées grâce à de l'acide formique concentré et à une solution 35 % en ammoniaque. La concentration des tampons (10 mM) a été ajustée afin de maintenir un pouvoir tampon suffisant et de minimiser la modification des comportements chromatographiques (i.e. des temps de rétention) lors du changement de la nature de tampon. Le plan expérimental factoriel complet comportait 15 conditions expérimentales (figure 1) dont la condition opératoire centrale a été répétée indépendamment (préparations des nouveaux tampons) à deux reprises.

3.2. Appareillages

La colonne chromatographique est une XBridge C18 (100x2.1 mm i.d. ; 3,5 µm) de la société Waters (Milford, USA). Les chromatogrammes ont été enregistrés sur un équipement Waters Alliance 2695 couplé à un détecteur UV-DAD 996, à un débit de 0,25 mL min^{-1} et à une température de 30 °C. Tous les chromatogrammes ont été enregistrés de 210 nm à 400 nm avec une résolution de 1,2 nm et une fréquence d'acquisition de 2 Hz. L'intégration des pics a été réalisée à 280 nm afin d'obtenir un rapport signal sur bruit suffisant pour chacun des pics.

3.3. Logiciel

Un algorithme a été développé pour réaliser l'ACI et les régressions linéaires multiples. Il a été écrit dans le langage statistique R 2.8.1 pour Windows® qui est gratuitement distribué à l'adresse http://www.r-project.org. Ces programmes ont été exécutés sur un ordinateur personnel.

4. Résultats et discussions

4.1. Plan d'expériences

Lors de l'obtention des 17 chromatogrammes aux 15 conditions expérimentales différentes, un maximum de 9 pics a pu être observé. Le mélange inconnu devait donc contenir au minimum 9 composés. Certains des chromatogrammes obtenus ainsi que les conditions opératoires correspondantes sont présentées en Figure 2.

4.2. Lecture automatique par ACI

Chaque chromatogramme a été traité par ACI afin de pouvoir réaliser les intégrations sur les composantes indépendantes plutôt que sur les pics pouvant résulter de composés coéluants. En effet, lorsque deux composés coéluent, la détermination des limites d'intégration est biaisée car même si la coélution est légère ($1{,}0 < R_S < 1{,}5$), les limites d'intégration sont généralement obtenues à l'aide d'un séparateur placé au minimum de la vallée entre les deux pics. Cette approximation restreint la justesse de cette mesure et empêche l'obtention de modèle mathématique reflétant, le plus exactement possible, le comportement chromatographique des composés présents dans le mélange. Dans le cas d'une coélution importante ($0 < R_S < 1$), l'estimation des largeurs des pics est faussée et entachera les prédictions fournies par les modèles mathématiques sur l'ensemble du domaine expérimental. Dès lors, les bornes d'intégration (i.e. le début et la fin d'un pic) sont déterminées automatiquement sur les composantes indépendantes. En partant du sommet d'un pic et en se déplaçant vers la gauche ou vers la droite, le premier point dont la hauteur est inférieure à deux fois le bruit est respectivement sélectionné comme le début ou la fin du pic. La Figure 3 illustre certains des résultats obtenus pour la séparation numérique (par ACI) de pics coéluants. Le traitement de l'ensemble des chromatogrammes par ACI a d'ailleurs permis de confirmer la présence de 9 composés dans le mélange inconnu.

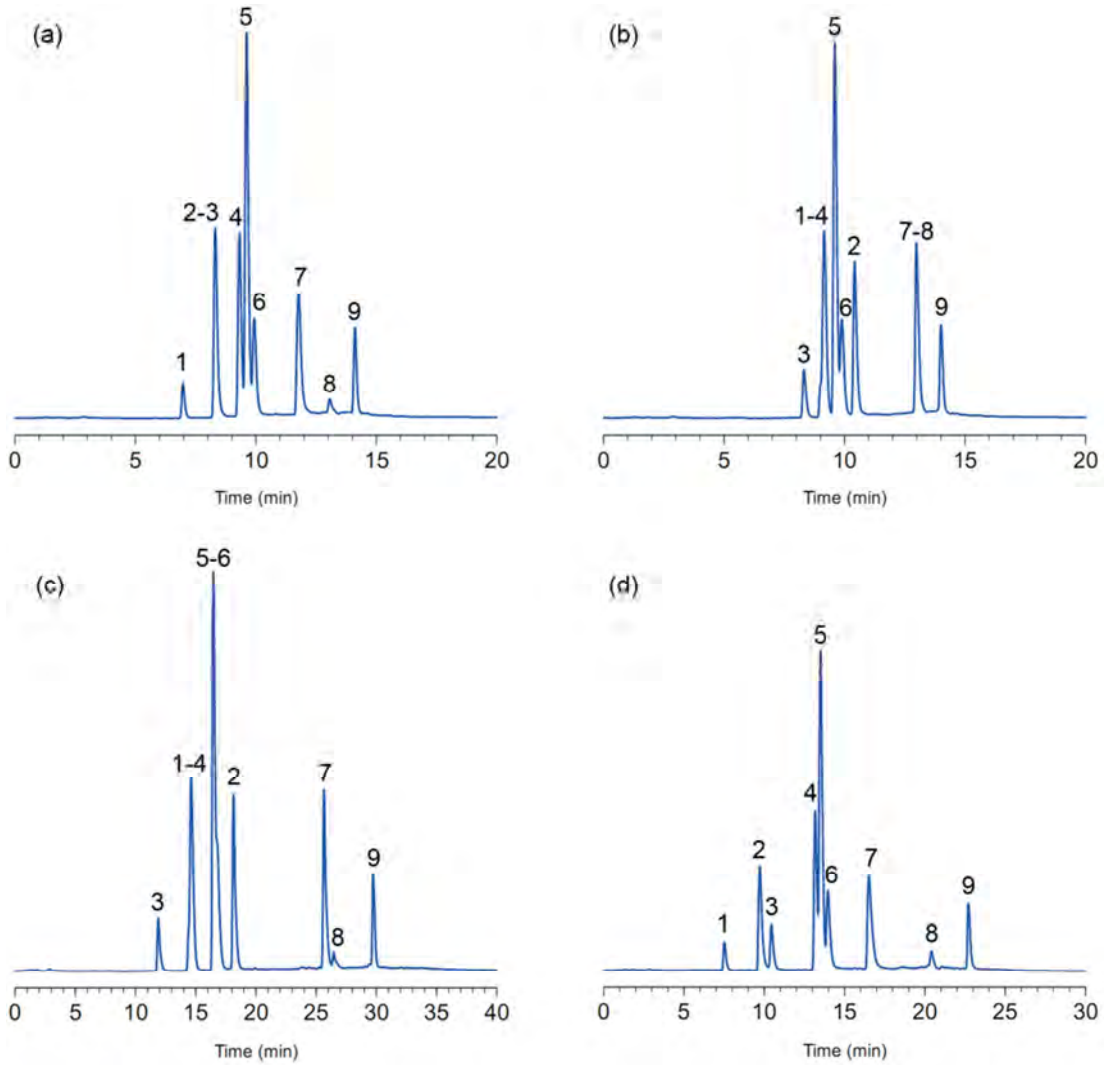


Fig. 2 Exemples de chromatogrammes obtenus lors de la réalisation du plan expérimental. (a) condition expérimentale 10, (b) condition expérimentale 13, (c) condition expérimentale 15, (d) condition expérimentale 8.

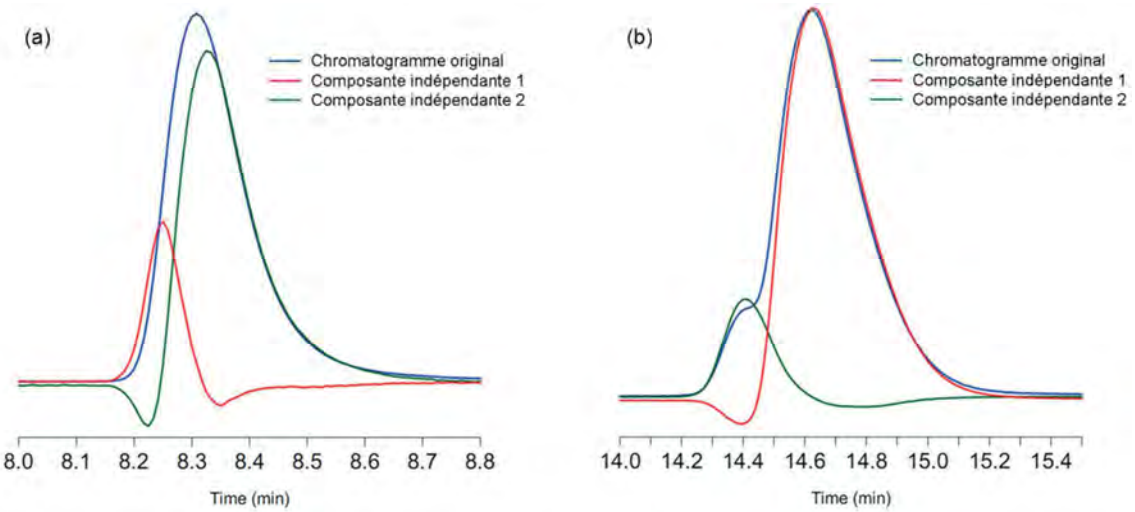


Fig. 3 Exemple de séparations numériques obtenues par ACI pour des composés coéluants. (a) coélution des composés 2 et 3 du chromatogramme représenté à la Figure 2a. (b) coélution des composés 1 et 4 du chromatogramme représenté à la Figure 2c.

4.3. Modélisation mathématique et prédiction

Afin de vérifier la qualité des régressions linéaires multiples, les coefficients de détermination ajusté ($R^2_{\text{ajusté}}$) obtenus pour chaque composé sont repris à la Table 1. L'adéquation existant entre les temps de rétention prédits par les modèles et ceux observés lors de l'enregistrement des chromatogrammes permet également de confirmer la qualité de l'ajustement. La Figure 4 illustre les relations entre réponses observées et réponses prédites, ainsi que les résidus correspondants. Il est à noter que la distribution Normale des résidus ainsi que la réalisation d'un test de manque d'ajustement (*lack of fit test*) ont également permis de confirmer la qualité de l'ajustement. Dès lors que la qualité des régressions linéaires multiples a été vérifiée, c'est-à-dire que les réponses ont été correctement modélisées, il est possible de réaliser une optimisation sur les critères sélectionnés. Dans notre cas, et pour les raisons énoncées précédemment, le critère retenu est la séparation entre les pics de la paire critique, *S*. La Figure 5 représente la surface de probabilité que *S* soit plus grand que 0. Sur cette figure est également indiqué le DS pour un niveau de qualité π égal à 85 %. Afin de pouvoir confirmer la capacité du DS à prédire des conditions opératoires offrant une séparation au moins égale à 0, un ensemble de chromatogrammes ont été enregistrés à des conditions opératoires se trouvant à l'intérieur ou à proximité de ce dernier. La Table 2 résume les résultats obtenus.

Table 1 Coefficient de détermination ajusté ($R^2_{\text{ajusté}}$) pour le logarithme des facteurs de rétention ($\log(k)$), pour le logarithme de la demi-largeur à gauche ($\log(w_a)$) et à droite ($\log(w_b)$).

Composé N°		1	2	3	4	5	6	7	8	9	Moyenne
$R^2_{\text{ajsuté}}$	$\log(k)$	0,991	0,993	0,999	0,997	0,999	0,998	0,999	0,998	0,999	0,997
	$\log(w_a)$	0,901	0,801	0,453	0,959	0,643	0,797	0,935	0,512	0,303	0,700
	$\log(w_b)$	0,955	0,969	0,444	0,780	0,942	0,833	0,768	0,491	0,372	0,698

L'écart existant entre les valeurs de *S* prédites et celles observées est clairement non négligeable. Plusieurs raisons sont à l'origine de cette différence. Premièrement, parmi les réponses modélisées, les logarithmes de demi-largeur de pic sont celles qui offrent la moins bonne justesse en prédiction (cf. table 1). Quoiqu'il en soit, le choix de ces réponses est motivé par leurs indépendances, en effet, des réponses indépendantes sont préférables pour une bonne analyse de l'erreur prédictive globale. Les demi-largeurs modélisées permettent de recalculer les temps en début et en fin de pic grâce à la valeur du temps de rétention qui lui-même est affecté de son erreur de prédiction. C'est ainsi que la source de l'erreur observée pour les temps en début et en fin de pic est double, car dépendant de l'erreur de prédiction sur le facteur de rétention et de l'erreur de prédiction propre aux modélisations de ces demi-largeurs.

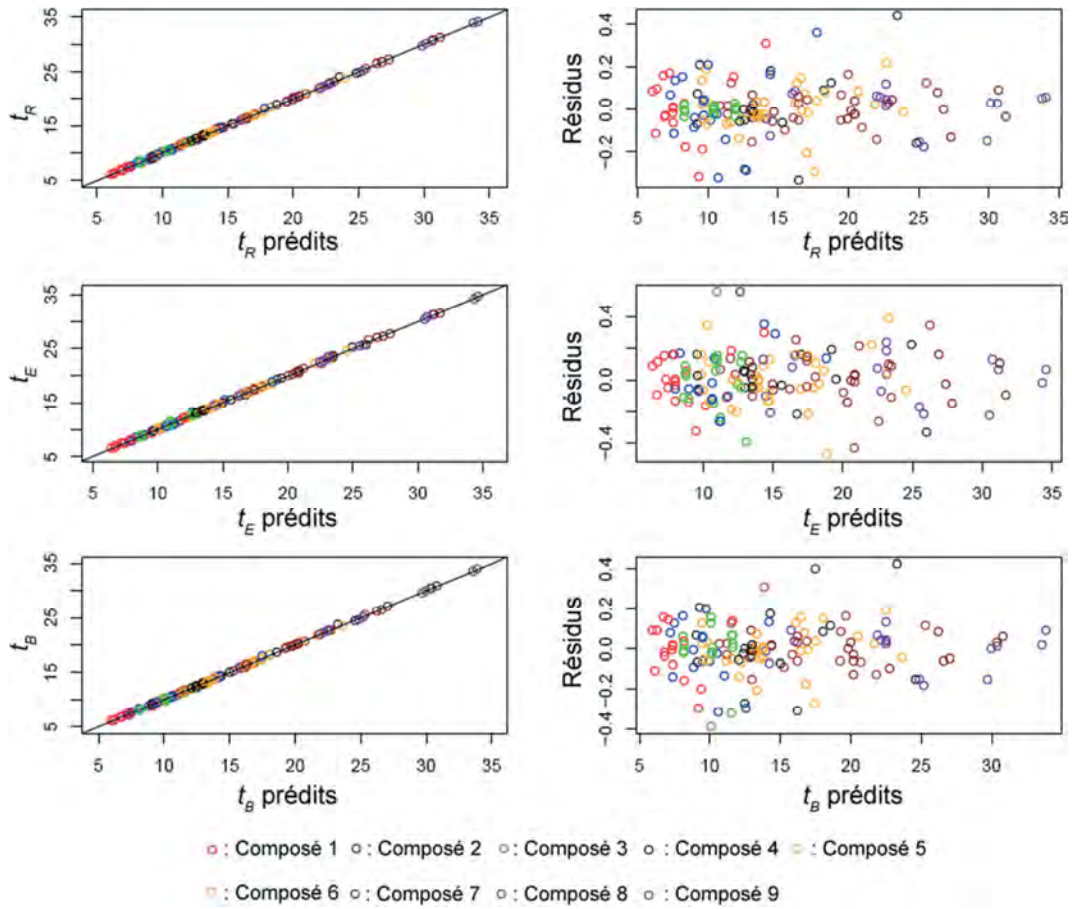


Fig. 4 Corrélations obtenues pour les temps de rétention (t_R) prédits et observés, ainsi que pour les temps au début des pics (t_B) et à la fin des pics (t_E).

Table 2 Adéquation entre séparations (S) prédites et observées pour les 6 conditions expérimentales définies à la Figure 5.

N°	pH	T_G (min)	P($S > 0$)	$S_{prédit}$ (min)	$S_{observé}$ (min)
Optima 1	3,14	30,0	88%	0,37	-0,25
Optima 2	2,9	30,0	86%	0,34	-0,09
Optima 3	3,65	30,0	85%	0,27	0,11
Optima 4	2,9	27,5	80%	0,26	-0,28
Optima 5	2,9	25,0	77%	0,19	-0,19
Optima 6	2,9	20,0	70%	0,10	-0,18

Par ailleurs, S étant directement calculé à partir des temps en début et fin de deux pics, sa valeur se trouve également entachée d'une erreur (cependant moins que R_S) ayant une double origine. En ce qui concerne la résolution, les valeurs prédites sont nettement inférieures à 1,5 et les écarts entre R_S observés et R_S prédits sont beaucoup plus significatifs. Deuxièmement, comme observé à la Figure 3, les composantes indépendantes correspondant à des composés coéluants peuvent être affectées d'une déformation vers des valeurs négatives à l'endroit de la coélution. Ceci peut être clairement remarqué sur les composantes indépendantes 1 et 2 de la Figure 3a. Cette déformation génère une sous-estimation des largeurs de pics, ce qui engendre une surestimation de la séparation prédite.

Des travaux sont en cours afin de pallier ce problème. Néanmoins, les prédictions sont en moyenne très exactes. Le coefficient de corrélation linéaire entre les temps de rétention prédits et observés rapportés sur la Figure 6a est de 0,998. Le chromatogramme enregistré à la condition opératoire optimale 2 (cf. figure 6b) permet également de confirmer que la coélution existant entre les composés 4 et 6 n'empêche pas de quantifier ceux-ci. Finalement, la majorité des composés ont été identifiés grâce à leur spectre UV. Initialement numéroté de 1 à 9, les composés sont respectivement, l'aténolol, le pindolol, un composé inconnu, la warfarine, l'indoprofène, le naproxène, le propranolol, une impureté de l'acide rétinoïque et l'acide rétinoïque.

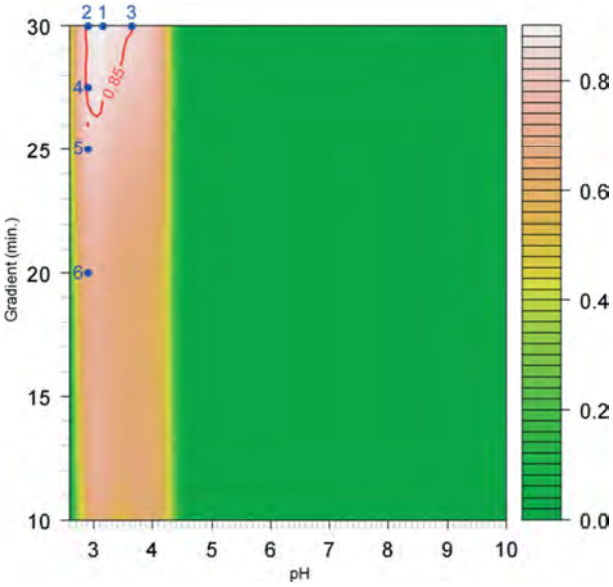


Fig. 5 Surface de probabilité d'atteindre $S > 0$ min. Le DS avec un niveau de qualité de 85 % est entouré par la ligne rouge. Les conditions expérimentales en bleu sont numérotées arbitrairement et définissent les conditions optimales testées.

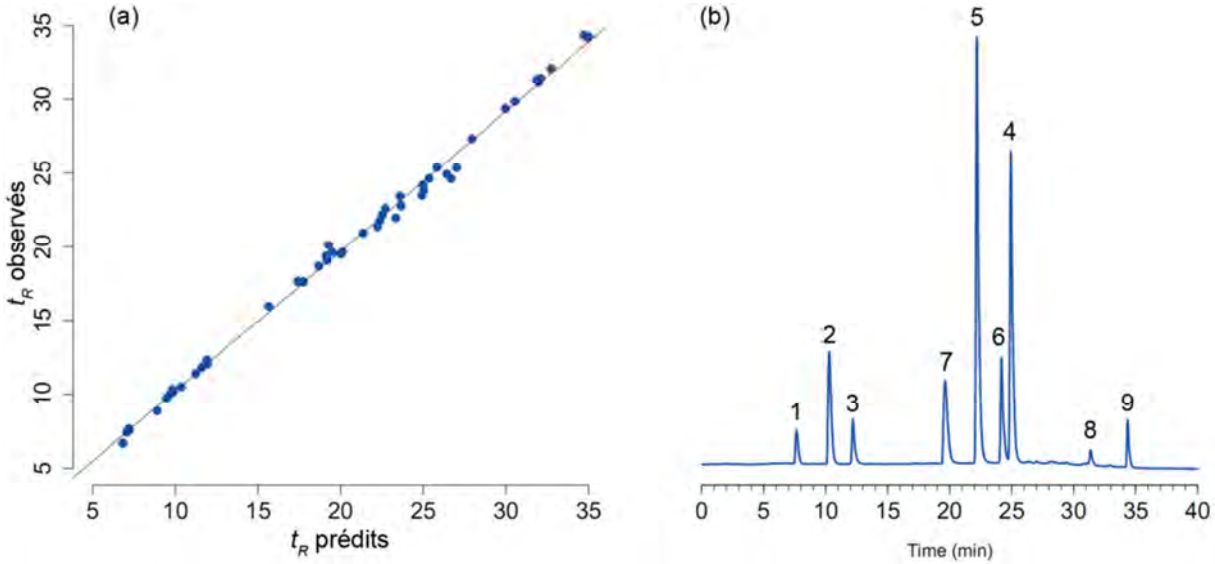


Fig. 6 (a) Corrélation entre les temps de rétention prédits et observés pour les chromatogrammes enregistrés aux conditions opératoires optimales (1, 2, 3, 4, 5, 6). (b) Chromatogramme enregistré à la condition optimale 2.

5. Conclusions

L'optimisation automatisée de séparations réalisées en chromatographie liquide constitue la première étape critique dans le domaine du développement automatisé de méthodes chromatographiques. Cette méthodologie qui repose sur les notions de planification expérimentale, d'analyse en composantes indépendantes, de régressions linéaires multiples, de propagation de l'erreur prédictive et de Design Space a été appliquée avec succès pour séparer les composés d'un mélange inconnu. D'autres travaux pourront permettre d'améliorer la qualité des prédictions, grâce à l'amélioration, d'une part, de la modélisation des réponses corrélées, et d'autre part, de la qualité des séparations numériques obtenues par ACI. Quoi qu'il en soit, cette méthodologie globale représente déjà un outil puissant susceptible d'améliorer la qualité et la quantité des résultats fournis par un laboratoire de recherche et/ou d'analyse dans des domaines aussi variés que les sciences biomédicales et pharmaceutiques, l'agroalimentaire, l'industrie chimique, la pharmacognosie et le chimiotaxonomie.

Références

[1] R. Berges, V. Sanz-Nebot, J. Barbosa, J. Chromatogr. A 869 (2000) 27.

[2] M.C. Garcia-Alvarez-Coque, J.R. Torres-Lapasio, J.J. Baeza-Baeza, Anal. Chim. Acta 579 (2006) 125.

[3] P. Nikitas, A. Pappa-Louisi, J. Chromatogr. A 1216 (2009) 1737.

[4] L.R. Snyder, J.J. Kirkland, J.L. Glajch, Practical HPLC Method Development, Wiley & Sons, 2nd edn., New York, 1997.

[5] P. Lebrun, B. Govaerts, B. Debrus, A. Ceccato, G. Caliaro, Ph. Hubert, B. Boulanger, Chemom. Intell. Lab. Syst. 91 (2008) 4.

[6] W. Dewé, R. Marini, P. Chiap, Ph. Hubert, J. Crommen, B. Boulanger, Chemom. Intell. Lab. Syst. 74 (2004) 263.

[7] A. Hyvärinen, J. Karhunen., E. Oja, Independent Component Analysis, Wiley, New York, 2001.

[8] B. Debrus, P. Lebrun, A. Ceccato, G. Caliaro, B. Govaerts, B.A. Olsen, E. Rozet, B. Boulanger, Ph. Hubert, Talanta 79 (2009) 77.

[9] H. Martens, M. Martens, Experimental Designs, 2nd edn., Wiley, 2001.

Section III.1.3

Implémentation de la méthodologie ICA au DoE-DS pour l'optimisation robuste de méthodes chromatographiques

Sous-section B

B. Debrus, P. Lebrun, A. Ceccato, G. Caliaro, E. Rozet, I. Nistor, R. Oprean, F.J. Rupérez, C. Barbas, B. Boulanger, Ph. Hubert, Application of new methodologies based on design of experiments, independent component analysis and design space for robust optimization in liquid chromatography, Anal. Chim. Acta *In press* (2011).

Résumé

Les séparations en CLHP d'un mélange inconnu et d'une formulation pharmaceutique ont été optimisées en utilisant une méthodologie chimiométrique initialement proposée par W. Dewé *et al*. en 2004 et récemment améliorée par P. Lebrun *et al*. en 2008. Cette méthodologie est basée sur les plans d'expériences qui sont utilisés pour modéliser les temps de rétention des composés étudiés. Ensuite, la précision de la prédiction et la robustesse de la séparation optimale, incluant l'étude de l'erreur, ont été évaluées. Finalement, "l'espace de conception" (*the design space*, directive ICH Q8(R1)) a été calculé comme la probabilité qu'un critère se trouve dans des limites d'acceptation. De plus, les chromatogrammes ont été lus automatiquement. La détection et l'appariement des pics ont été réalisés en utilisant une méthodologie précédemment développée et basée sur l'analyse en composantes indépendantes, publiée en 2009 par B. Debrus *et al*. Les présentes applications couronnées de succès, renforcent le haut potentiel de ces méthodologies en vue du développement automatisé de méthodes chromatographiques.

Summary

HPLC separations of an unknown sample mixture and a pharmaceutical formulation have been optimized using a recently developed chemometric methodology proposed by W. Dewé et al. in 2004 and improved by P. Lebrun et al. in 2008. This methodology is based on experimental designs which are used to model retention times of compounds of interest. Then, the prediction accuracy and the optimal separation robustness, including the uncertainty study, were evaluated. Finally, the design space (ICH Q8(R1) guideline) was computed as the probability for a criterion to lie in a selected range of acceptance. Furthermore, the chromatograms were automatically read. Peak detection and peak matching were carried out with a previously developed methodology using independent component analysis published by B. Debrus et al. in 2009. The present successful applications strengthen the high potential of these methodologies for the automated development of chromatographic methods.

1. Introduction

The optimization of liquid chromatography (LC) operating conditions for the separation of several compounds in complex samples is often intricate. Indeed, it can be tricky to separate compounds due to their similar chromatographic behaviours or fussy to elute all of them well separated when some have widely distinct physico-chemical properties (e.g. polarities, pKa, log P). Since past decades and mainly those last years, a lot of improvements in the fields of method development have been done [1-3] and it is thus possible to foresee different strategies that could be applied to find optimal separations in a rather automated way. The main difficulty is that these strategies need to give accurate and robust predictions prior to the validation and transfer of these analytical methods. ICH Q8(R2) guideline [4] provides a harmonized guidance to improve the robustness and reliability of pharmaceutical development. In this guideline the design space (DS) is defined as "*the multidimensional combination and interaction of input variables (e.g. material attributes) and process parameters that have been demonstrated to provide assurance of quality*". To decrypt this definition, it can be assumed that, the "*multidimensional combination and interaction of input variables and process parameters*" is a multidimensional space (or subspace) whose dimensions are the factors used during the method development. Afterwards, the guideline mentions "*that have been demonstrated to provide assurance of quality*" which means that the size of this space is defined by the set of combinations of factors ranges wherein a process provides quality results. In the LC field, and more precisely when the separation optimization is the main aim, the DS predicts a space wherein the separation is achieved taking into account the measurements, process and models uncertainties. The separation quality can be evaluated by a chromatographic criterion such as the resolution ($R_S = 2 \cdot (t_{R,2} - t_{R,1})/(w_{b,1} + w_{b,2})$; with $t_{R,2} > t_{R,1}$ and $w_{b,1}$, $w_{b,2}$ are the peaks widths at baseline; $t_{R,1}$ and $t_{R,2}$ being the retention times of the critical pair peaks) and the method uncertainty is estimated by the probability to reach a given criterion threshold (e.g. the probability for R_S to be higher than 1.5) in future uses of the method. It also states that "*working within the design space is not considered as a change*". Therefore, in the framework of separation methods development, DS can be clearly considered as a zone of theoretical robustness since modifications of the method parameters will not significantly affect the separation quality. Thus, no decrease of quality in the separation should be observed while working in DS. Consequently, to optimize HPLC separations and advisedly compute DS, design of experiments (DoE) is one of the most appropriate strategies. Examples of DS for HPLC assays can be found in references [5,6].

Thus, based on DoE strategy, a recent methodology [7] was used to model and predict the retention times according to selected chromatographic factors (e.g. mobile phase composition, gradient time, mobile phase pH) and subsequently optimize the separation. Afterwards, the prediction error was also estimated in order to allow the DS computation. However, DoE involves the recording of chromatograms at very different operating conditions. By this way, very distinct selectivities are obtained. Peak detection and identification in each chromatogram can thus become tedious and time-consuming. Hence,

an additional methodology based on independent component analysis (ICA) has been developed to detect and match peaks among the chromatograms resulting from DoE [8].

Two test samples were selected to evaluate these complementary methodologies. The first sample (sample 1) was sent from Eli Lilly and Company and was considered as an unknown sample. A similar test mixture was already used by Biswas et al. to conduct a method screening study [9]. However, in this work, the amount of compounds and their nature were deliberately occulted to test out both methodologies. The second sample (sample 2) is a common-cold pharmaceutical formulation. Studies involving some of these compounds were previously published: the separation of the three active drugs using a poly(ethyleneglycol) column [10], the separation of the three active drugs using cyano columns [11], the comparison between five HPLC columns [12], the comparison between electrophoresis and liquid chromatography separation [13] and the validation of a HPLC method for the quantification of the active drugs [14]. Furthermore, a review of analytical methods published for the separation of some of these substances can be found in the works of Marín et al. [14]. These abundant publications highlight the interest that still remains in separating and quantifying these compounds. Nevertheless, as the sweeteners of a new sugar-free formulation were not taken into account in the aforementioned methods, a new method development was necessary.

2. Materials and methods

2.1. Chemicals

Methanol (HPLC grade) was purchased from Sigma (St-Louis, MO, USA). Ultra–pure water was obtained with a Millipore (Billerica, MA, USA) Milli-Q Academic A10. Formic acid (>98%) was purchased from Merck (Darmstadt, Germany), ammonium formate (99%) was purchased from Alfa Aesar (Karlsruhe, Germany) and ammonium hydrogen-carbonate (99.7%) was purchased from VWR (Fontenay-sous-Bois, France). The unknown sample (sample 1) was provided by Eli Lilly & Co (Indianapolis, IN, USA) and was ready for injection. The pharmaceutical formulation (sample 2) came from the Cinfa Laboratory (Huarte, Spain). It consisted in a common-cold formulation which contained phenylephrine hydrochloride, acetaminophen, chlorpheniramine maleate, sodium saccharin, sodium cyclamate, mannitol, orange flavor, sunset yellow FCF and polyvinylpyrrolidone (PVP-K30).

2.2. Sample preparation

2.2.1. Sample 1 – unknown sample

This sample was provided in solution which was a mixture of water-acetonitrile (50:50, v/v). The solution was then filtered with a 0.20 µm syringe filtration disk to a vial for injection in the HPLC system. The injection volume was 0.5 µL. The rather low injection

volume was selected to prevent any peak distortion such as peak fronting or retention times shift due to high elution strength of this solubilization mixture (i.e. 50% of acetonitrile) compared to the elution strength at the beginning of the gradient tested in the DoE (i.e. 5% of methanol). In addition, for 0.5 µL injections, injected volume repeatability was tested and the resulting estimated precision was very good (CV% = 0.5%).

2.2.2. Sample 2 – pharmaceutical formulation

800 mg of the formulation powder were dissolved in a 10 mL volumetric flask with a mixture of water-acetonitrile (80:20, v/v). After 10 min of magnetic stirring an aliquot was filtered with a 0.20 µm syringe filtration disk to the vial for injection in the HPLC system. In accordance with the theoretical formulation content, the concentrations of the active drugs were: acetaminophen (13 mg mL^{-1}), phenylephrine (0.16 mg mL^{-1}) and chlorpheniramine (0.06 mg mL^{-1}). The injection volume was 1.0 µL.

2.3. HPLC experiments

Both separations were carried out on an Alliance 2695 separation module coupled with a UV-DAD 2996 detector from Waters (Milford, MA, USA). The analytical column was an XBridge C18 (100x2.1 mm i.d.; particle size 3.5 µm) from Waters. A C18 column was selected because this kind of columns usually offers relatively good retention properties for most organic compounds generally having a hydrophobic character. XBridge columns also sustain highly acidic and alkaline conditions (pH 1–12) thanks to the trifunctionally bonded C18 and the ethylene bridges within the silica matrix. In order to keep suitable column lifetime, the pH range was slightly narrowed from 2.6 to 10 (see Table 1). The experiments were carried out at a flow rate of 0.25 mL min^{-1} at 30 °C. The buffers consisted in 10 mM pH 2.6 formic acid, pH 4.45 ammonium formate, pH 6.3 ammonium formate, pH 8.15 ammonium hydrogencarbonate and pH 10.0 ammonium hydrogencarbonate. The pH was adjusted to the selected value with concentrated formic acid or ammonia 35% aqueous solution. The buffers concentration (10 mM) was adjusted to avoid pH modification and to minimize modification of the chromatographic behaviour due to the change of buffer composition. Linear gradients were carried out from 5% to 95% of methanol and an isocratic step at 95% was hold during 10 min before column equilibration for the next injection. This final isocratic step was kept this long to avoid compound elution after the gradient and to ensure compound detection in a given time window for most unknown mixtures. If a mixture composition is known, the final isocratic step could obviously be reduced by a factor 5.

All the chromatograms were recorded between 210 nm and 400 nm with an estimated step of 1.2 nm and an acquisition frequency of 2 Hz. The peaks integration was carried out at 280 nm and 215 nm for sample 1 and 2, respectively. For coeluting peaks, the integrations were made on their respective independent components resulting from ICA numerical separation.

2.4. HPLC robust separation optimization

2.4.1. Responses and criteria

The first critical step is the selection of the response. Indeed, the choice of the factors directly depends on this response which also needs to be modelled by a multiple linear function. The use of resolution (R_S) of the critical pair (i.e. the two most proximate peaks in a chromatogram) as the response of the model can be hazardous due to its non-linear and discontinuous behaviour [7]. Thus, the retention times at the beginning, the apex and the end of each peak (respectively t_B, t_R, t_E) were measured and were transformed into retention factors (k_{tR} = (t_R-t_0)/t_0), with t_0 equal to the column dead-time. For each compound, the modelled responses were the logarithm of the retention factor (k_{tR}) and the logarithm of the semi-widths (i.e. $w_l = k_{tR} – k_{tB}$ and $w_r = k_{tE} – k_{tR}$). Based on these responses, the computation of the predicted value of some chromatographic criteria (resolution, total analysis time, etc.) only involved simple parameters (i.e. t_B, t_R and t_E). As the further computation of error propagation can be delicate when R_S is selected as a criterion (see section 2.4.5.), a simpler one was introduced: the separation criterion, S, defined as the difference between t_B of the second peak and t_E of the first peak of the critical pair. Even if S and R_S are highly correlated, the computation of S and its associated uncertainty is easier. Regardless of the critical pair peaks size ratio, a value of S equal or higher than 0 means that the peaks are baseline resolved which is similar to a value of R_S approximately equal or higher than 1.5.

2.4.2. Experimental factors, ranges and levels

The choice of the factors and their respective range is primordial. These factors have to affect the selected response in order to modify the selectivity enough to achieve a total separation of all the compounds. Furthermore, the investigated experimental domain (i.e. the factors ranges multidimensional space or workspace) needs to be large enough to create a response variation which allows encountering acceptable criteria values (i.e. S >0 or R_S > 1.5). As the response variation is not known a priori, the selection of a large experimental domain minimizes the risk of not finding any good separations. Furthermore, the number of levels determines the polynomial order (or polynomial degree) of the multi-linear equation used to model the response. For example, if two levels are selected, the modelling equation can only be linear. Generally, if no prior information about the response variation is known, preliminary experiments should be carried out to estimate the range and the order of variation of each factor.

In the present study, the selected factors were the pH of the aqueous part of the mobile phase (*pH*) and the gradient time (T_G) to change the organic modifier proportion from 5% to 95%. In order to orthogonally modify the selectivity, it is common use to change the column or the organic modifier. However, these qualitative factors (i.e. which can only take some given values) considerably increase the number of experiments in DoE. Hence, selectivity modification was mainly based on *pH* effect. Besides, even if T_G did not affect the

selectivity as considerably as *pH*, this factor was selected to mainly adjust the retention times in order to get acceptable analysis times. This consideration is of first importance because the ICA-DoE-DS methodology is intended to be as generic as possible.

Generally, one can observe a sigmoidal shape of the theoretical relationship between t_R and *pH* for monoprotic molecules. So five *pH* levels were selected to use a polynomial order of four (i.e. until pH^4). Furthermore, three levels of gradient time were chosen to estimate its quadratic effect (i.e. T_G^2).

2.4.3. Choice of a design

The selected DoE needs to have good statistical properties as orthogonality and/or rotatability and to maintain the number of experiments as low as possible. DoE can be split up in two categories. Screening designs allow estimating the factors effect on the selected response. When too many factor seems to affect the response (i.e. four factor or more), screening designs are mainly used to select factors that have the higher effects (i.e. the factors that create the higher response variation). In this category, a well-known design is the Plackett and Burman design. In liquid chromatography, Plackett and Burmann are also used to estimate the robustness of an optimal separation [15-19]. The second category corresponds to optimization designs (i.e. used to model the response and to optimize response and/or criteria). These DoE are mainly full factorial (i.e. complete combination of all factors levels), fractional factorial designs (i.e. a statistically selected subpart of full factorial design), D-optimal (i.e. tend to minimize the parameter estimates covariance then maximizing the D factor) and central composite designs (i.e. tend to place the experiments on a circular/spherical shape around the experimental domain centre to obtain a good rotatability property) [20]. Nowadays, several softwares can be used to determine the experimental designs according to the analyst's choices.

In the present study, one full factorial design was repeated for each sample (see Table 1). The central operating condition (i.e. pH 6.3 – T_G=20 min) was further independently repeated twice (i.e. preparation of new buffers and mobile phases). The DoE thus counted a total of 17 experiments including 2 independent repetitions at the centre (i.e. the central point was performed in triplicates as usually recommended).

Table 1 Full factorial design used in the present study.

Factors	Levels				
pH	2.6	4.4	6.3	8.1	10.0
T_G (min)	10	20	30		

2.4.4. Response modelling

The response modelling is achieved by fitting a multiple linear equation (i.e. polynomial model) whose mathematical terms depend on the selected factor level number (e.g. linear for two terms, quadratic for three terms). Nevertheless, at least one degree of

freedom must theoretically be kept to estimate the error. Practically, two or more degrees of freedom are kept to accurately estimate this error. For example, for a simple linear model, two experiments are mandatory to estimate the equation parameters but a third one is needed to approximate the model error which depends on parameters error. On the other hand, the addition of too much high order terms or crossed terms can lead to overfitting (i.e. noise/random error modelling). For this purpose, stepwise regression was applied to maximize the adjusted coefficient of determination (R^2_{adj}) [21]. In the present case, two factors were selected (*pH* and T_G) and the model used for multiple linear regressions for both samples is shown with Eq. 1.

$$\log(k_{tR}) = \beta_0 + \beta_1 \times pH + \beta_2 \times pH^2 + \beta_3 \times pH^3 + \beta_4 \times pH^4 + \beta_5 \times T_G + \beta_6 \times {T_G}^2 + \beta_7 \times pH \times T_G + \varepsilon \quad (1)$$

with β_0...β_7 the model parameters and ε the estimated error.

In order to evaluate the models adequacy, lack of fit tests were also performed. Lack of fit tests consist in comparing experimental and predicted variances. A normal zero-centred residuals distribution was already a good clue for adequate models (see section 3.1.2. and 3.2.2.). χ^2 and anova tests were performed and confirmed the models adequacy (i.e. p-values > 0.05).

At this step, it is also important to mention that the models could be used to highlight or to confirm some information about the physico-chemical properties (e.g. pKa, logP). However, this kind of interpretations was not the aim of the present study. The main objective was to demonstrate the applicability of the ICA-DoE-DS methodology for robust LC method optimization.

2.4.5. Quality criteria, error propagation, design space computation and robustness estimation

Following the ICH Q2(R1) guideline [22], "*the robustness of an analytical procedure is a measure of its capacity to remain unaffected by small, but deliberate variations in method parameters*". In a separative framework, it means that the criteria – that quantify the separation quality – will remain unchanged around the optimal separation area. Robustness can be evaluated after the optimization process [15,16]. Using DoE-DS methodologies, robustness assessment is concurrently performed during the optimization phase. The error was assumed normal and was defined by a Gaussian law as shown in Eq. 2.

$$\log(k_{tR,i}) = \log(\hat{k}_{tR,i}) + \mathrm{N}(0, \sigma_i) \quad (2)$$

with $\log(k_{tR,i})$, the estimate distribution of the *i*th response (corresponding to the *i*th compound); $\log(\hat{k}_{tR,i})$, the mean predicted value for the *i*th modelled response and $N(0,\sigma_i)$, a normal distribution centred on 0 and with a standard deviation equal to σ_i obtained as computed by the *i*th model. In practice, the Monte-Carlo method [23] was then used to propagate the Gaussian estimation of $\log(k_{tR,i})$ error to the separation criterion to assess robustness. For this purpose, 2500 simulations were carried out. Prediction error

computation was carried out in a similar way for w_l and w_r. The distribution of $t_{R,i}$ (i.e. the distribution of t_R for the *i*th compound) can be obtained following Eq. 3. The distribution of $t_{B,i}$ and $t_{E,i}$ can also be obtained by Eq. 4 and 5, respectively.

$$t_{R,i} = \left(\exp\left(\log\left(k_{tR,i}\right)\right) + 1\right) \times t_0 \quad (3)$$

$$t_{B,i} = t_{R,i} - \exp(\log(w_l)) \quad (4)$$

$$t_{E,i} = t_{R,i} + \exp(\log(w_r)) \quad (5)$$

Finally, the error was propagated to the separation criteria, S, with Eq. 6.

$$S = t_{B,2} - t_{E,1} \quad (6)$$

Where, S is the estimated distribution of the separation criterion, $t_{E,1}$ is the end of the first peak of the critical pair and $t_{B,2}$ is the begin of the second peak of the critical pair, considering that this computation was carried out at a given operating condition to simplify notations. Similar computations can also be used to find the distribution of R_S or other selected quality criteria.

If and only if the criteria error is estimated, the computation of the DS can be done. If R_S is selected as criterion, the division taking place in its calculation makes its error more complex than for S which is computed by a simple subtraction [7]. Following the ICH Q8(R2) guideline, the DS can be defined as the set of experimental conditions where the criterion meet acceptance limits (e.g. S > 0.2 min) with a given probability, known as the quality level (e.g. π = 60%), as shown with Eq. 7.

$$\mathrm{DS} = \left\{x_0 \in \chi | E_{\hat{\theta}}\left[\mathrm{P}(S > \lambda)|\hat{\theta}\right] \geq \pi\right\} \quad (7)$$

Where x_0 is a point in the experimental domain, χ. λ is the acceptance limit for criterion S, π is the quality level and $\hat{\theta}$ is the set of estimated parameters of the model. P and E respectively correspond to the estimators of probability and mathematical expectation. The DS is thus assumed to be a zone of theoretical robustness as the predicted criterion values are higher than an advisedly selected acceptance limit (threshold) together with a high probability to at least reach this acceptance limit. In the present study, π was selected as 90% of the optimal probability to get S > 0 min. For instance, for sample 1, P(S > 0) at the optima was equal to 94.4%. The quality level was settled at 85% (i.e. 94.4%×90%) (see section 3.1.3. and Fig. 4). On one hand, a low π do not mean that unacceptable separations could be obtained (i.e. S < 0). The peak could be well separated concurrently with a relatively high error tarnishing the t_R predictions leading to predictions of low probability while peaks could be well separated. On the other hand, a high π could induce obtaining a good separation.

Up to now, it is important to stress that there is still no regulatory documents stating how to conveniently compute or estimate (nor define) the required quality level. Nevertheless, ICH Q8(R2) remains the most appropriated guideline when discussing about DS [4].

2.5. Automatic reading of chromatograms

Before modelling responses, the retention times (t_B, t_R and t_E) have to be obtained for each peak in each tested experimental conditions. Three successive steps can be identified: peak detection (i.e. effective presence of peaks in the chromatogram, corresponding to the investigated compounds) which is not obvious when coelutions are present, peak matching (i.e. concordance between the detected peaks in a given chromatogram and the peaks of another chromatogram obtained in another analytical condition of DoE) and peak identification (i.e. attribution of a given peak to a clearly identified compound). The chromatograms coming from DoE depict various selectivities and therefore complicate peak detection and peak matching. This step is highly time-consuming and its automation can be very useful.

ICA is a mathematical process which maximizes the statistical independence between estimated independent components. ICA has already been used on UV-DAD chromatograms [8] and further details can be found in the works of Hyvärinen et al. [24,25]. In the present study, ICA numerical separation was carried out on each chromatogram to accomplish the integration on the independent components rather than on the original peaks that can result from coeluting compounds. Indeed, when two compounds coelute, the determination of the integration limits is biased. If the coelution is slight ($1.0 < R_S < 1.5$; $S < 0$), the integration limits are usually obtained using a vertical separator at the minimum of the valley between the two peaks (also called the valley drop-line). This approximation restricts the accuracy of this measurement and prevents obtaining mathematical model that reflects, as accurately as possible, the chromatographic behaviour of compounds in the sample. In the case of a major coelution ($0 < R_S < 1$, $S << 0$), estimation of peak widths is highly biased. These inaccurate integration limits generate inaccurate predictions on the whole experimental domain. Therefore, the limits of integration (i.e. the beginning and the end of a peak, t_B and t_E) were automatically determined on the independent components. Starting from the top of a peak and moving left or right, the first point whose height is less than twice the noise was respectively selected as the beginning or end of the peak. By this way, ICA methodology also provides an accurate estimation of the peak number in the sample and ease peak matching [8].

2.6. Software

In-house computer software were developed to run the ICA algorithm (i.e. fastICA R package) and to perform the multiple linear regressions. The coding was carried out with the academic version of Revolution R 1.3.0 (including R 2.7.2), freely distributed at http://www.revolutionanalytics.com/.

3. Results and Discussion

To avoid column equilibration at each mobile phase pH for each injection, experiments at a same pH were carried out in a row. Experiments within a pH level were however conducted in a random order. In fact, it can be assumed that the mean relative error performed during buffer pH measurements is about 0.3% (i.e. corresponding to 3 times the specified pH-meter precision). Therefore, the retention time variability component coming from pH measurement can be considered as negligible – in the framework of the present feasibility study – in comparison to other variability sources coming from the LC equipment (e.g. t_R variability coming from gradient slope or mobile phase composition, etc.). Examples of chromatograms for both samples are depicted in Fig. 1.

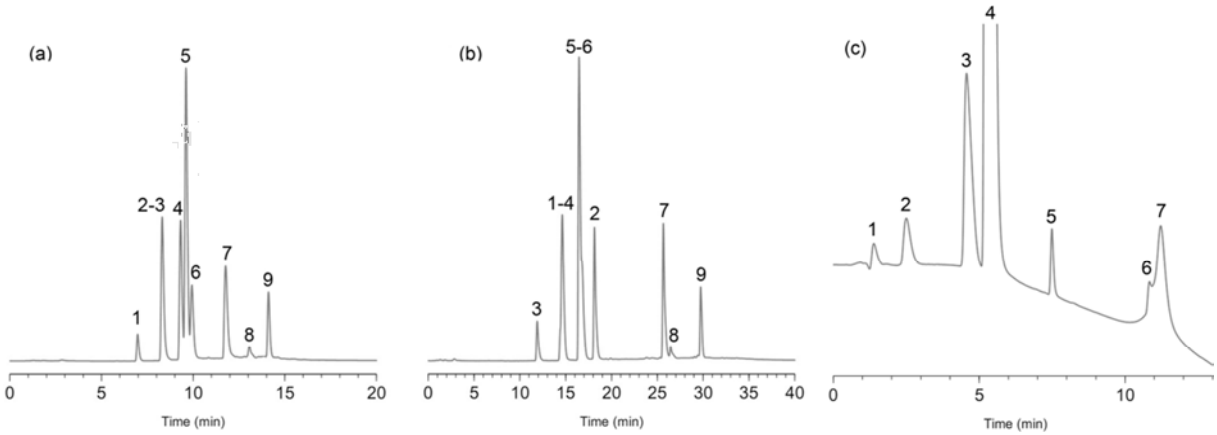


Fig. 1 Examples of chromatogram obtained during the experimental design. (a) Chromatogram of sample 1 at pH 8.1 and T_G = 10 min displayed at 280 nm, (b) chromatogram of sample 1 at pH 10 and T_G = 30 min displayed at 280 nm and (c) chromatogram of sample 2 at pH 6.3 and T_G = 10 min displayed at 215 nm. Sample 2 peak numbering: 1 =maleate, 2 = phenylephrine, 3 = saccharin, 4 = acetaminophen, 5 = sunset yellow FCF, 6 = chlorpheniramine, 7 = PVP-K30.

3.1. Sample 1 – unknown sample

As sample 1 was unknown, peaks were numbered in accordance with their elution order for the first tested experimental condition (i.e. pH 6.3 – T_G=10 min).

3.1.1. Automatic reading of chromatograms

When the experimental design was carried out, a maximum of nine peaks was observed. At this stage, sample 1 should then contain at least nine compounds. Fig. 2 presents two examples of numerical separation of coeluting peaks obtained when carrying out the experimental design. ICA numerical separation also confirmed the presence of nine compounds in sample 1. Finally, the retention times (i.e. t_B, t_R and t_E) were automatically obtained for each independent component.

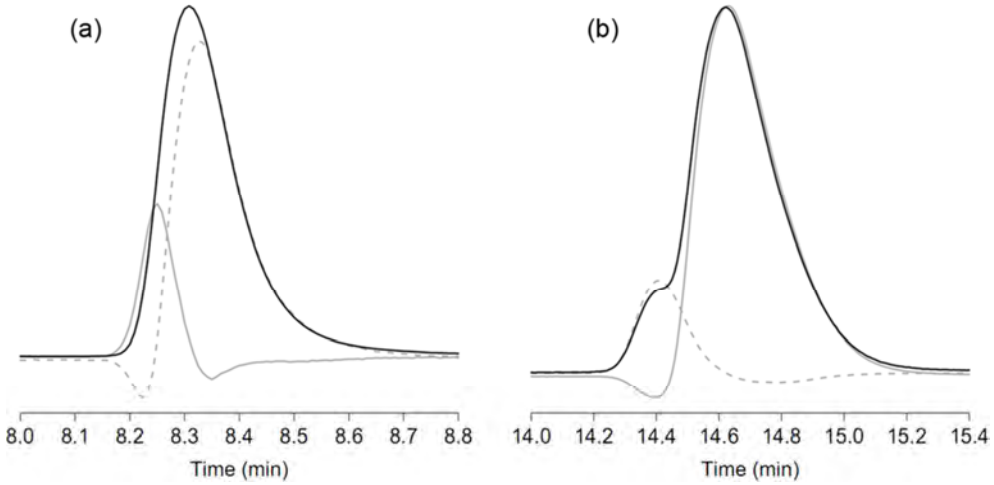


Fig. 2 Examples of a numerical separation of coeluted peaks carried out by ICA. The black line represents the recorded signal. (a) Numerical separation of compound 2 (dashed line) and compound 3 (grey line) for the chromatogram depicted in Fig. 1a. (b) Numerical separation of compound 1 (dashed line) and compound 4 (grey line) for the chromatogram displayed in Fig. 1b.

3.1.2. Responses modelling and criterion computation

In order to assess the polynomial fit quality, the adjusted coefficients of determination (R^2_{adj}) are presented in Table 2. The computed R^2_{adj} resulted from the models (see Eq. 1) rather than the correlation between experimental and predicted retention times on Fig. 3a.

Table 2 Adjusted coefficient of determination (R^2_{adj}) and residuals standard deviation (σ) for the logarithm of the retention factors ($\log(k_{tR})$), for the left half-width ($\log(w_l)$) and for the right half-width ($\log(w_r)$) for sample 1.

Compound N°		1	2	3	4	5	6	7	8	9	Mean
R^2_{adj}	$\log(k_{tR})$	0.991	0.993	0.999	0.997	0.999	0.998	0.999	0.998	0.999	0.997
	$\log(w_l)$	0.901	0.801	0.453	0.959	0.643	0.797	0.935	0.512	0.303	0.700
	$\log(w_r)$	0.955	0.696	0.444	0.780	0.942	0.833	0.768	0.491	0.372	0.698
σ	$\log(k_{tR})$	0.023	0.020	0.003	0.016	0.008	0.011	0.007	0.004	0.006	0.011
	$\log(w_l)$	0.117	0.220	0.450	0.196	0.213	0.172	0.093	0.279	0.173	0.212
	$\log(w_r)$	0.296	0.318	0.414	0.553	0.330	0.363	0.388	0.189	0.205	0.339

Additionally, the adequacy between the retentions times predicted by the models and those observed also allowed confirming fits quality (Fig. 3). This latter figure clearly illustrates the good relationship between the experimental responses and the predicted ones as well as the corresponding residuals. In addition, residuals Gaussian distributions also corroborated the quality of the multiple linear regressions and obtained models. For each response, a Shapiro-Wilk test was carried out on the residuals distribution to verify its normality. The p-values were equal to 0.111, 0.129 and 0.080 for $\log(k_{tR})$, $\log(w_l)$ and $\log(w_r)$, respectively. The residuals were thus normally distributed (p-values > 0.05). Once responses were correctly modelled, the separation criterion computation was carried out. Using a grid search method, the separation criterion and its corresponding error were computed over

the whole experimental domain. Furthermore, using Monte-Carlo simulations, the probabilities for *S* to be greater than 0 min (i.e. baseline resolved peaks) were also calculated as it can be seen in Fig. 4.

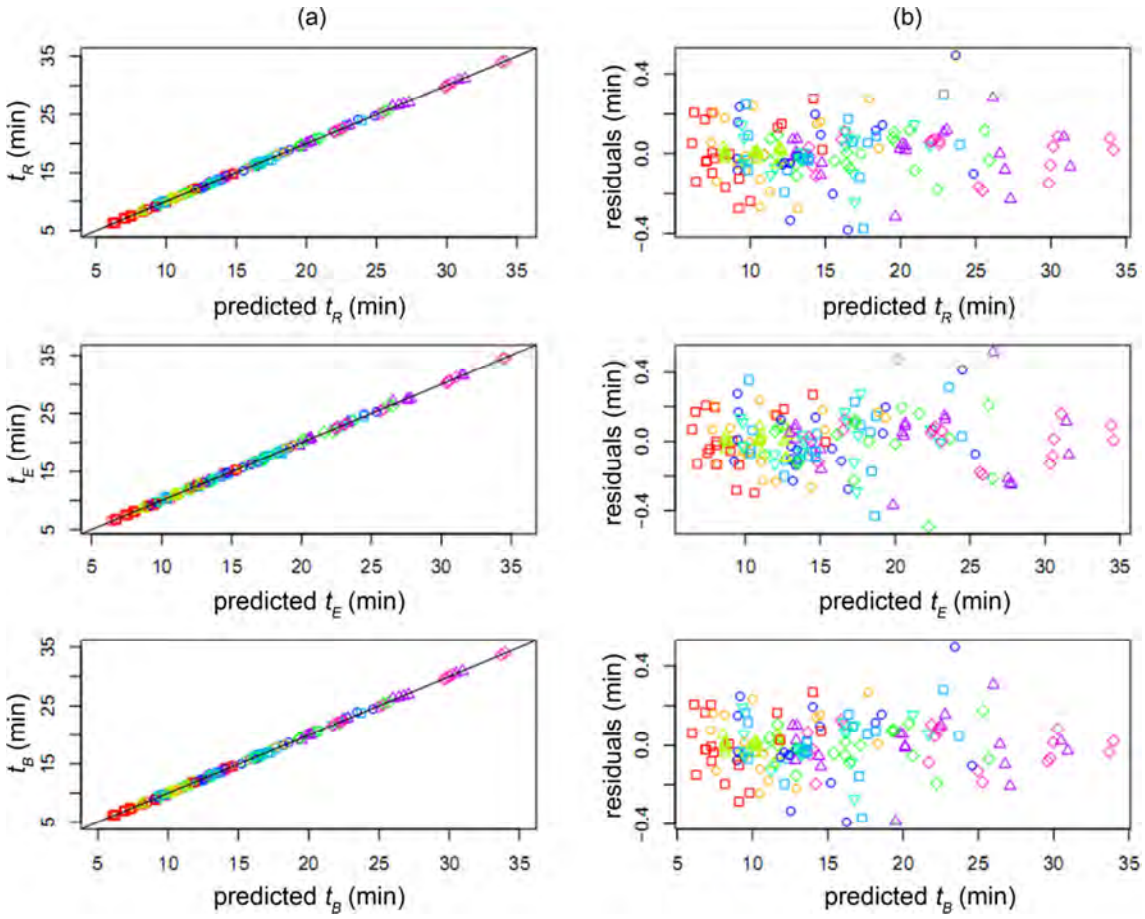


Fig. 3 Modelling results for sample 1. (a) Predicted versus experimental values for t_R, t_E and t_B. (b) Corresponding residuals plots. Compound assignation: (red circle) compound 1, (orange circle) compound 2, (green triangle) compound 3, (blue circle) compound 4, (blue triangle) compound 5, (blue square) compound 6, (green diamond) compound 7, (purple triangle) compound 8 and (pink diamond) compound 9.

3.1.3. Optimal separation

One optimal experimental condition was identified at pH 3.1 and with T_G=30 min (Fig. 4, point 1). Fig. 5b shows the chromatogram recorded at this optimal condition. The predicted *S* was 0.34 min (i.e. predicted R_S = 2.06) and the observed one was -0.09 min (i.e. observed R_S = 1.23). Some reasons are behind the difference between predicted and experimental values. First, among the modelled responses, the logarithms of peak half-width were those that offer the lowest accuracy in prediction (see Table 2). The modelled half-widths allowed calculating the beginning and end of the peak with the value of retention time, which itself was affected by its error. Thus, the observed errors for t_B and t_E have a double source. It depends on the error of the retention factor and the specific half-widths

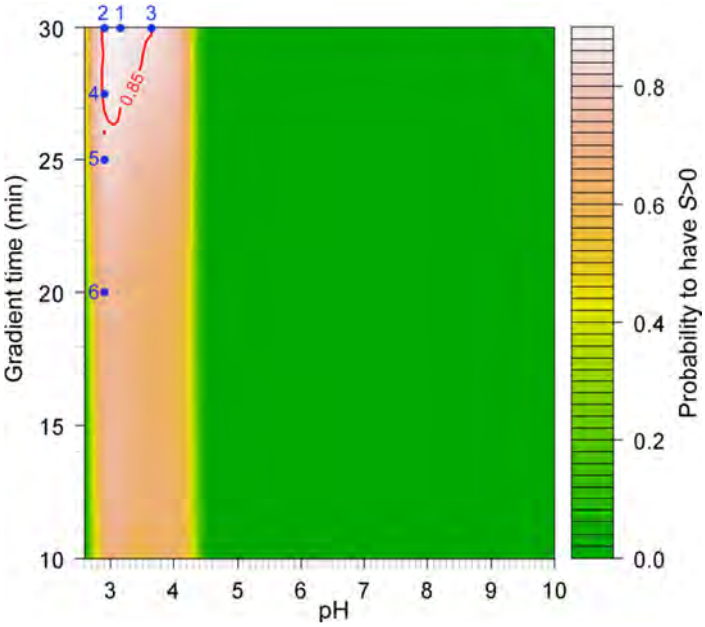


Fig. 4 Probability surface for the separation criterion (acceptance limit λ = 0min) with the design space hedged-in the black line (π = 85%). Blue dots depict additional tested experimental conditions.

errors. Moreover, S was directly calculated from the beginning and end times of two peaks, its value is also tarnished (though less than R_S) with a double error coming from the modelling of t_R and w_l or w_r. Secondly, as observed in Fig. 2, independent component coming from coeluted peak can show a local negative baseline on the side where coelution was initially observed. This can be clearly observed on the independent components in Fig. 2a. This deformation generates a relatively small underestimation of peak widths, resulting in a relatively small overestimation of predicted separation.

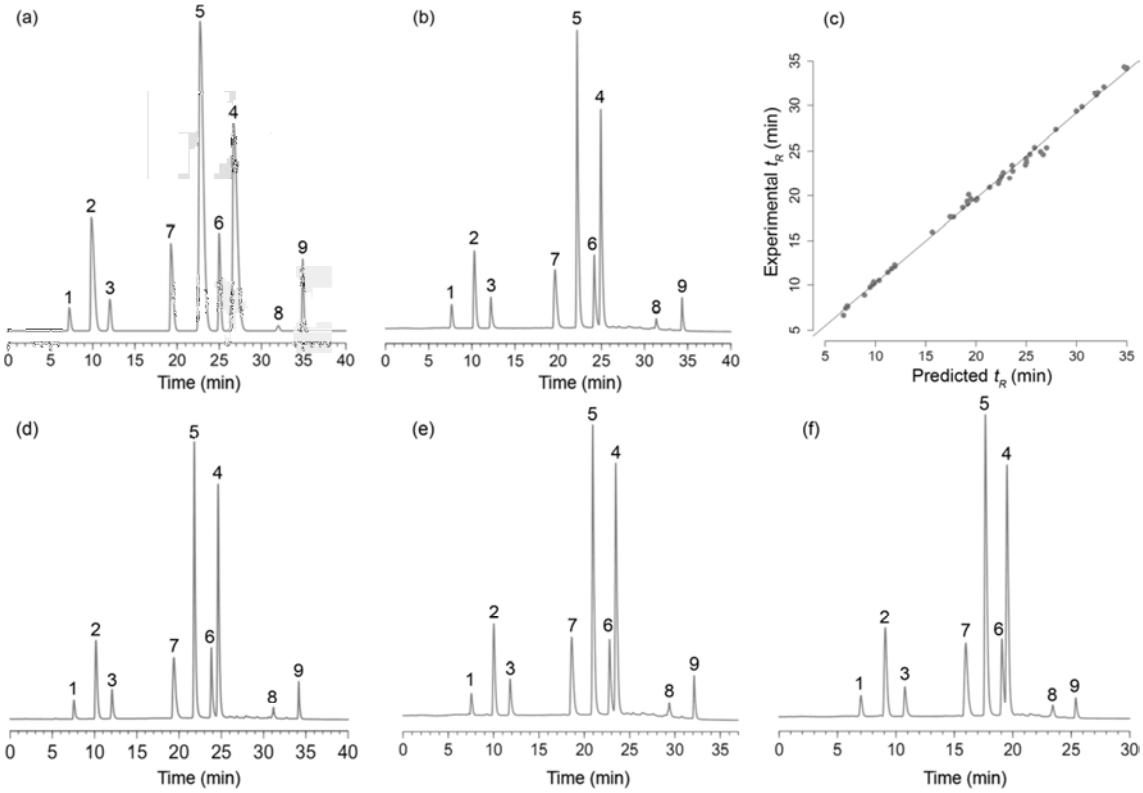


Fig. 5 (a) Predicted chromatogram at the optimal condition 1 (pH 3.1 and T_G = 30.0 min), (b) chromatogram recorded at the optima condition 1 (pH 3.1 and T_G = 30 min). (c) Relationship between experimental retention times versus predicted ones for the nine peaks and for the six tested experimental conditions. The black line is a simple line with 0 intercept and slope of 1. (d) Chromatogram recorded at condition 3 (pH 3.7 and T_G = 30.0 min), (e) chromatogram recorded at condition 4 (pH 2.9 and T_G = 27.5 min) and (f) chromatogram recorded at condition 6 (pH 2.9 and T_G = 20.0 min).

Work is underway to overcome this problem. Nevertheless, the predictions are very accurate on average. The correlation coefficient between the predicted retention times (i.e. t_R) and observed ones is 0.998 as reported in Fig. 5c. The chromatogram recorded at the optimal operating condition (see Fig. 5b) also confirms that the slight coelution existing between compound 4 and 6 do not prevent the quantitative determination of these two compounds.

3.1.4. Design space robustness assessment

As it can be seen in Fig. 4 (points 2-6), five experimental conditions were selected to verify models ability to predict and to assess the robustness. The experimental separation criteria, *S*, measured in experimental conditions 2, 3 and 4 were not significantly different from the separation measured at the optimal condition (point 1) chromatograms. Even if the corresponding chromatogram present slight coelution between compound 4 and 6 ($S \sim 0$ min), the observed separations did not change significantly within DS. On the other hand, chromatograms were recorded at experimental condition numbered 5 and 6 (i.e. outside the DS, see Fig. 4) in order to confirm robustness in DS. In both of them, the coelution of compound 4 and 6 was so important that their accurate quantification would be impossible (see Fig. 5f).

Compounds of sample 1 were first matched using their UV spectrum (i.e. each peak was labelled with a number). Secondly, Eli Lilly & Co provided peak identification with structures and names of labelled peaks. Originally numbered from 1 to 9, the compounds were respectively, atenolol, pindolol, a licensed compound, warfarin, indoprofen, naproxen, propranolol, an impurity of retinoic acid and retinoic acid.

3.2. Sample 2 – pharmaceutical formulation

As sodium cyclamate, mannitol and orange flavor do not absorb in the usable UV range (above 200 nm), they were not taken into account in the present optimization process. Also, sunset yellow FCF does not appear in the following sections as it did no absorb UV at acidic pH (at pH 2.6). Peak integration was carried out by the same operator using ICA in order to avoid additional error sources. For each chromatogram, the peaks were integrated and identified using their UV spectra. Then, the measured values of t_B, t_R and t_E were stored in a file for the further computational process.

3.2.1. Automatic reading of chromatograms

When coeluted peaks were observed, ICA was used to numerically separate the peaks to avoid a poor estimation of integration limits. Then, the integrations were carried out on the respective components resulting from the ICA numerical separation. Compounds were manually identified thanks to the comparison between their UV spectra and reference UV

spectra [26]. These identifications were also confirmed by the injections of individual solution of each compound.

3.2.2. Responses modelling and criterion computation

On one hand, the adjusted coefficients of determination (R^2_{adj}) were computed in order to assess the polynomial fit quality. As shown in Table 3, the logarithm of the retention factors were well modelled since the adjusted R^2_{adj} were higher than 0.98. R^2_{adj} were computed from the models (see Eq. 1).

Table 3 Adjusted coefficient of determination (R^2_{adj}) and residuals standard deviation (σ) for the logarithm of the retention factors ($\log(k_{tR})$), for the left half-width ($\log(w_l)$) and for the right half width ($\log(w_r)$) for sample 2. Peak numbering: 1=maleate, 2=phenylephrine, 3=saccharin, 4=acetaminophen, 6= chlorpheniramine, 7=PVP-K30.

Compound N°		1	2	3	4	6	7	Mean
R^2_{adj}	$\log(k_{tR})$	0.993	0.998	0.982	0.999	0.999	0.999	0.995
	$\log(w_l)$	0.524	0.914	0.444	0.841	0.906	0.995	0.771
	$\log(w_r)$	0.912	0.947	0.982	0.845	0.486	0.961	0.856
σ	$\log(k_{tR})$	0.027	0.016	0.013	0.013	0.010	0.002	0.014
	$\log(w_l)$	0.047	0.068	0.110	0.148	0.063	0.035	0.079
	$\log(w_r)$	0.079	0.089	0.044	0.094	0.268	0.082	0.109

On the other hand, Fig. 6 illustrates the relationship between the experimental and the predicted responses as well as the corresponding residuals. The Gaussian distributions of the residuals also corroborated the quality of the multiple linear regressions and allowing accurate predictions of retention times. As previously, a Shapiro-Wilk test was performed on the residuals distribution to verify the response normality. The p-values were equal to 0.056, 0.062 and 0.101 for $\log(k_{tR})$, $\log(w_l)$ and $\log(w_r)$, respectively. The residuals distributions were thus considered normal (p-values > 0.05).

3.2.3. Optimal separation

Two optimal separation conditions were encountered. The first one (optimum n°1) at pH 4.3 and with T_G = 29.5 min and the second one (optimum n°2) at pH 10.0 and with T_G = 10.0 min. When an acceptance limit λ = 0.1 min for S was selected, the quality level had to be adjusted to 40% to find a DS. At this value, two DS were found around each optimum as seen on Fig. 7a and 7b.

The chromatograms recorded at optimal conditions 1 and 2 are presented in Fig. 8a and 8b, respectively. To test the precision of these operating conditions, three chromatograms were independently recorded with new buffer solutions at each optimal experimental condition. The repeatabilities of retention times were very good with CV(%)

values of 0.6% and 0.8% for optimum 1 and 2, respectively. Fig. 8a shows that saccharin (compound 3) and acetaminophen (compound 4) were slightly coeluted.

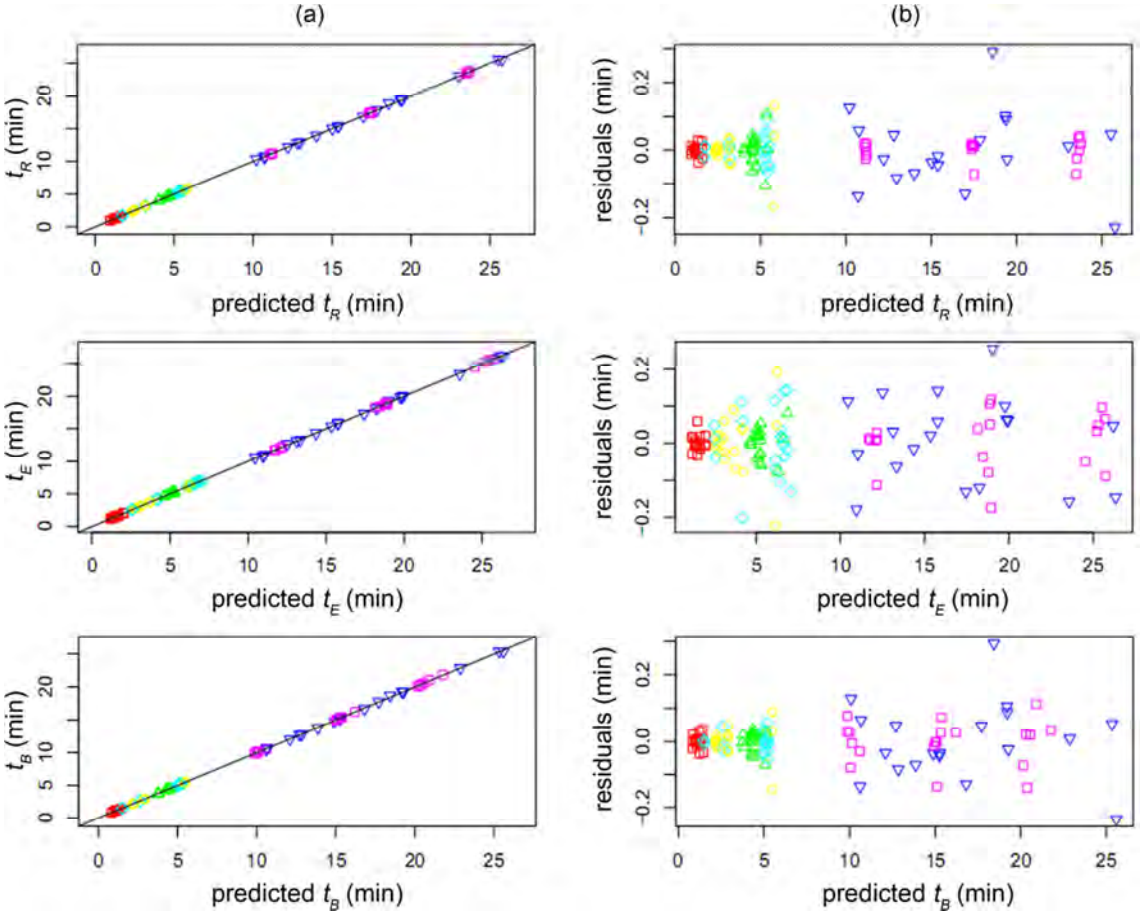


Fig. 6 Modelling results for sample 2. (a) Predicted versus experimental values for tR, tE and tB. (b) Corresponding residuals plots. Compound assignation: (red square) maleate, (yellow circle) phenylephrine, (green triangle) saccharine, (blue diamond) acetaminophen, (blue triangle) chlorpheniramine and (pink square) PVP-K30.

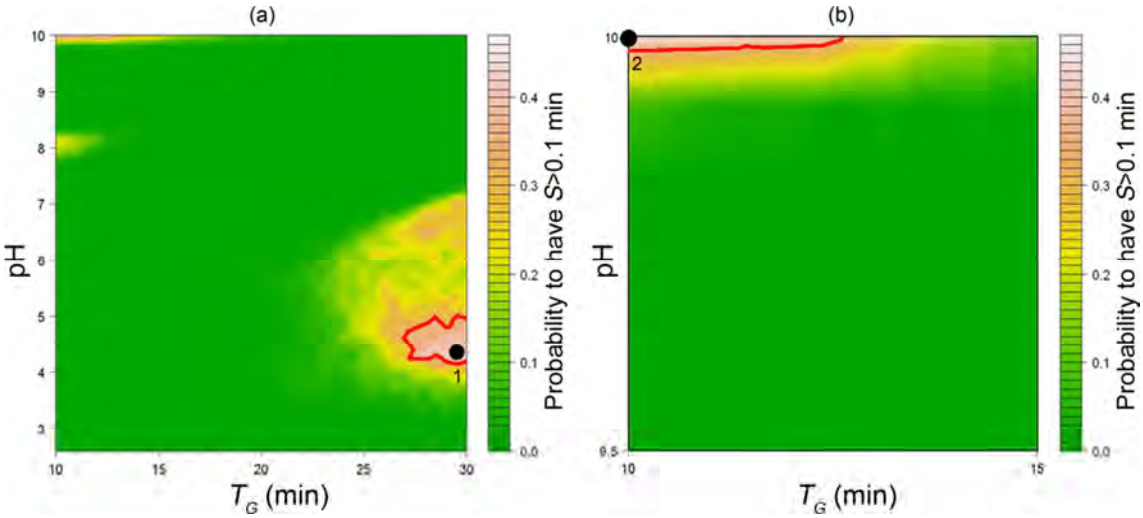


Fig. 7 (a) Probability surface for the separation criterion (acceptance limit λ = 0.1 min) with the design space 1 (π = 40%) hedged-in the red line; optimum 1 is depicted by the black circle (pH 4.3 – T_G = 28.0 min). (b) Magnification of the upper left zone of the probability surface with the design space 2 (π = 40%) hedged-in the red line, optimum 2 is depicted by the black circle (pH 10 – T_G = 10 min).

As the probability of the prediction to get S > 0.1 minute was only of 40%, this small coelution (S < 0) was not surprising. On the contrary, the chromatograms recorded at optimum 2 (Fig. 8b), the separation S of the critical pair was higher than 0.1 minute and seemed to be robust as a significant variation of S was not observed during the independent repetitions.

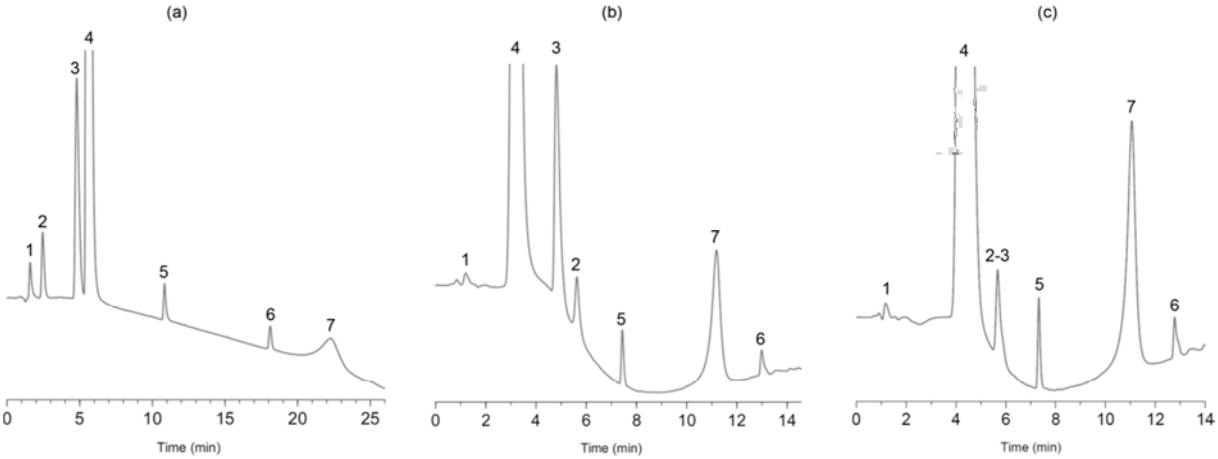


Fig. 8 (a) Chromatogram recorded at optimum 1 (pH 4.3 – T_G = 28.0 min). (b) Chromatogram recorded at optimum 2 (pH 10.0 – T_G = 10 min) (c) Chromatogram recorded outside the DS as a counterexample (pH 9.5 – T_G = 10 min). Peak numbering: 1 =maleate, 2 = phenylephrine, 3 = saccharin, 4 = acetaminophen, 5 = sunset yellow FCF, 6 = chlorpheniramine, 7 = PVP-K30.

3.2.4. Design space robustness assessment

The computed DS were large enough to evaluate the separation robustness. In the DS 1 (corresponding to optimum 1), three experimental conditions were tested, at pH 4.3 with T_G = 28.0 min and at pH 4.8 with T_G = 28.0 and 29.5 min. The measured separations, S, on the corresponding chromatograms were not significantly different from the initial measured separation (optimum 1). As for optimum 1, the separation criterion threshold (i.e. λ = 0.1 min) was not reached. However, these separations were barely acceptable (S < 0 min but > – 0.1 min) as the coelution between peak 3 and 4 was very slight. Moreover, the observed separations did not change significantly within DS involving an expected robustness. The shape of DS 2 (corresponding to optimum 2) was thinner, so only one experimental condition was added, at pH 10.0 and with T_G = 12.0 min, to evaluate the DS. In this chromatogram, a minor decrease of S was noticed but it was still higher than the selected acceptance limit (S > 0.1 min). Therefore, this DS defined a zone of acceptable robust separation. As previously carried out with sample 1, a counterexample was performed outside of the DS, at pH 9.5 and with T_G = 10.0 min (see Fig. 8c). At this experimental condition, a large coelution was observed between saccharin (compound 3) and phenylephrine (compound 2). In addition, a decrease in saccharin peak size between Fig. 8b and 8c was observed. It can be explained by the modification of the molar absorption coefficient with pH change.

4. Conclusions

The automated optimization of chromatographic separation is the first critical step in the framework of the automated development of chromatographic method. In this paper, DoE, ICA, multiple linear regression, error propagation and DS methodologies were successfully applied to separate nine compounds of an unknown sample mixture in less than 40 min and the seven compounds of a pharmaceutical formulation. Furthermore, for this latter sample, the analysis time was shortened to less than 14 min. This global methodology is also very flexible as the choice of each criterion and their respective acceptance limit are made by the analyst. In addition, an evaluation of DS robustness was carried out during the present study. The separation criterion *S* clearly demonstrated its robustness capability within the identified DS. It strengthens the fact that DS defines a space wherein the separations are complete and the method is robust and even more robust if the DS is large. Nevertheless, as the size of DS depends on the value of π, further works are still required to define possible adjustment strategy for this parameter. Eventually, if one compound is discarded from one of the tested mixtures, such as an active ingredient of the pharmaceutical formulation, the present DS would obviously have defined optimal and robust space for the separation of the subsequent mixture without any additional experiments.

Acknowledgements

The author would like to thank the Walloon Region of Belgium for the FIRST-DEI convention funds N°516130, the Eli Lilly and Company for partial funding of ADAM project and the fund Léon Fredericq for the grant allocated to his working period in Madrid. A research grant from the Belgium National Fund for Scientific Research (F.R.S-FNRS) to E. Rozet is also gratefully acknowledged. Bernard A. Olsen and Brian Fahie from Eli Lilly & Co are also gratefully acknowledged for their contribution to this work.

References

[1] L. R. Snyder, J. J. Kirkland, J. L. Glajch : Practical HPLC Method Development, Wiley & Sons, 2nd edn., New York, 1997.

[2] P. Schoenmakers, R.W. Abbott, Anal. Chim. Acta 208 (1988) 357.

[3] W. Dewé, R. Marini, P. Chiap, P. Hubert, J. Crommen, B. Boulanger, Chemom. Intell. Lab. Syst. 74 (2004) 263. http://hdl.handle.net/2268/6036

[4] International Conference on Harmonization (ICH) of Technical Requirements for registration of Pharmaceuticals for Human Use, Topic Q8(R2): Pharmaceutical Development, Geneva, 2005.

[5] J.J. Peterson, J. Qual. Tech. 36 (2004) 139.

[6] B. De Backer, B. Debrus, P. Lebrun, L. Theunis, N. Dubois, L. Decock, A. Verstraete, Ph. Hubert, C. Charlier, J. Chromatogr., B 877 (2009) 4115. http://hdl.handle.net/2268/4442

[7] P. Lebrun, B. Govaerts, B. Debrus, A. Ceccato, G. Caliaro, Ph. Hubert, B. Boulanger, Chemom. Intell. Lab. Syst. 91 (2008) 4. http://hdl.handle.net/2268/1640

[8] B. Debrus, P. Lebrun, A. Ceccato, G. Caliaro, B. Govaerts, B.A. Olsen, E. Rozet, B. Boulanger, Ph. Hubert, Talanta 79 (2009) 77. http://hdl.handle.net/2268/13236

[9] K.M. Biswas, B.C. Castle, B.A. Olsen, D.S. Risley, M.J. Skibic, P.B. Wright, J. Pharm. Biomed. Anal. 49 (2009) 692.

[10] A. García, F.J. Rupérez, A. Marín, A. de la Maza, C. Barbas, J. Chromatogr. B 785 (2003) 237.

[11] B. Olmo, A. García, A. Marín, C. Barbas, J. Chromatogr. B 817 (2005) 159.

[12] A. Marín, C. Barbas, J. Pharm. Biomed. Anal. 40 (2006) 262.

[13] A. Marín, C. Barbas, J. Pharm. Biomed. Anal. 35 (2004) 769.

[14] A. Marín, E. García, A. García, C. Barbas, J. Pharm. Biomed. Anal. 29 (2002) 701.

[15] Y. Vander Heyden, A. Nijhuis, J. Smeyers-Verbeke, B.G.M. Vandeginste, D.L. Massart, J. Pharm. Biomed. Anal., 24 (2001) 723.

[16] B. Dejaegher, Y. Vander Heyden, J. Chromatogr. A 1158 (2007) 138.

[17] B. Dejaegher, M. Dumarey, X. Capron, M.S. Bloomfield, Y. Vander Heyden, Anal. Chim. Acta 595 (2007) 59.

[18] B. Dejaegher, X. Capron, J. Smeyers-Verbeke, Y. Vander Heyden, Anal. Chim. Acta 564 (2006) 184.

[19] B. Dejaegher, J. Smeyers-Verbeke, Y. Vander Heyden, Anal. Chim. Acta 544 (2005) 268.

[20] A.C. Atkinson and A.N. Donev, Optimum Experimental Designs, Clarendon Press, Oxford, 1992.

[21] H. Martens and M. Martens, Multivariate Analysis of Quality. An Introduction, Wiley, Chichester, 2001.

[22] International Conference on Harmonization (ICH) of Technical Requirements for registration of Pharmaceuticals for Human Use, Topic Q2(R1): Validation of Analytical Procedures: Text and Methodology, Geneva, 1996.

[23] M.A. Herrador, A.G. Asuero, A.G. Gonzalez, Chemom. Intell. Lab. Syst. 79 (2005) 115.

[24] A. Hyvärinen, J. Karhunen, E. Oja, Independent Component Analysis, Wiley, New York, 2001.

[25] A. Hyvärinen, E. Oja, Neural Networks 13 (2000) 411.

[26] A.C. Moffat, M.D. Osselton, B. Widdop, Clarke's Analysis of Drugs and Poisons, Pharmaceutical Press, Electronic version, Edition 2004.

Section III.1.4

Relation entre modélisation et théorie de la chromatographie appliquée à un cas réel

B. Debrus, P. Lebrun, J. Mbinze Kindenge, F. Lecomte, A. Ceccato, G. Caliaro, J. Mavar, R. Marini , B. Boulanger, E. Rozet, Ph. Hubert, Innovative HPLC method development for the screening of 19 antimalarial drugs based on generic approach, using design of experiments and design space, J. Chromatogr. A *Submitted for publication* (2011).

Résumé

Une méthodologie novatrice basée sur les plans d'expériences (*design of experiments*: DoE), l'analyse en composantes indépendantes (*independent component analysis*: ICA) et sur "l'espace de conception" (*design space:* DS) a été développée dans des travaux précédents et a été testée à l'aide d'un mélange de 19 antipaludéens. Cette méthodologie globale de développement (DoE-ICA-DS) en chromatographie liquide a été utilisée pour optimiser la séparation de 19 antipaludéens afin d'obtenir une méthode criblage. Cette méthodologie est totalement en accord avec le concept de "qualité par la conception" (*quality by design*). Un plan d'expériences a permis de définir un ensemble de conditions opératoires afin de modéliser les temps de rétention au début, au sommet et à la fin de chaque pic. De plus, la séparation numérique des pics dans le but d'estimer leurs limites d'intégration non biaisées a été réalisée par l'analyse en composantes indépendantes. Le temps de gradient, la température et le pH ont été sélectionnés comme facteurs d'un plan factoriel complet. Les temps de rétention ont été modélisés par des régressions multiples linéaires en mode pas à pas. Un attribut critique de qualité récemment introduit, à savoir le critère de séparation (*S*), a été utilisé pour évaluer la qualité des séparations plutôt que de se baser sur la résolution. En outre, les modèles mathématiques résultants ont également étudiés d'un point de vue chromatographique pour comprendre et investiguer le comportement chromatographique de chaque composé. De bonnes adéquations ont été observées entre ces modèles et les comportements prédits par les théories chromatographiques. Finalement, dans une optique de gestion du risque, l'espace de conception a été calculé comme le sous-espace multidimensionnel où la probabilité que le critère de séparation se trouve dans des limites d'acceptation, soit plus grande qu'un niveau de qualité défini. La propagation de l'erreur de prédiction des réponses modélisées jusqu'au critère de séparation, en utilisant des simulations de Monte Carlo, a permis d'identifier cet espace de conception. Une méthode de criblage robuste pour les 19 antipaludéens sélectionnés, permettant de combattre la contrefaçon, a donc été obtenue à l'aide de la méthodologie DoE-ICA-DS. De plus et sur seule base des données obtenues par le plan d'expériences, une méthode spécifique pour la détermination de trois antipaludéens dans une formulation pharmaceutique a été proposée pour démontrer l'efficacité et la flexibilité de la méthodologie présentée.

Summary

An innovative methodology based on design of experiments (DoE), independent component analysis (ICA) and design space (DS) was developed in previous works and was tested out with a mixture of 19 antimalarial drugs. This global LC method development methodology (i.e. DoE-ICA-DS) was used to optimize the separation of 19 antimalarial drugs to obtain a screening method. DoE-ICA-DS methodology is fully compliant with the current trend of quality by design. DoE was used to define the set of experiments to model the retention times at the beginning, the apex and the end of each peak. Furthermore, ICA was used to numerically separate coeluting peaks and estimate their unbiased retention times. Gradient time, temperature and pH were selected as the factors of a full factorial design. These retention times were modelled by stepwise multiple linear regressions. A recently introduced critical quality attribute, namely the separation criterion (S), was also used to assess the quality of separations rather than using the resolution. Furthermore, the resulting mathematical models were also studied from a chromatographic point of view to understand and investigate the chromatographic behaviour of each compound. Good adequacies were found between the mathematical models and the expected chromatographic behaviours predicted by chromatographic theory. Finally, focusing at quality risk management, the DS was computed as the multidimensional subspace where the probability for the separation criterion to lie in acceptance limits was higher than a defined quality level. The DS was computed propagating the prediction error from the modelled responses to the quality criterion using Monte Carlo simulations. DoE-ICA-DS allowed encountering optimal operating conditions to obtain a robust screening method for the 19 considered antimalarial drugs in the framework of the fight against counterfeit medicines. Moreover and only on the basis of the same data set, a dedicated method for the determination of three antimalarial compounds in a pharmaceutical formulation was proposed to demonstrate both the efficiency and flexibility of the presented methodology.

1. Introduction

Malaria remains one of the most extended illnesses worldwide. According to world health organisation, almost 3.3 billion peoples scattered across hundreds of countries (mostly in the intertropical belt) are at risk of various species of plasmodium [1]. Despite numerous active antimalarial molecules, several reasons are behind malaria resurgence. Over past decades, improper use of antimalarial drugs contributed to widen resistance against malaria parasite to several drugs (e.g. chloroquine) [2,3]. Nevertheless, artemisinin-based combination therapies (ACT) bring new hopes in the fight against malaria [4-7]. Furthermore, in some African countries, up to 80% of medical products are counterfeit [8]. In this context, analytical chemistry and especially chromatographic methods can help to fight against counterfeit medicines.

Screening analytical methods are usually used for several purposes. They can be used to confirm if a targeted active antimalarial ingredient (AAI) is present or absent in a pharmaceutical formulation and to settle if an unappointed AAI is present or absent in this formulation. In this perspective, a screening HPLC method could be very helpful justifying the development of a generic method for the follow-up of the various antimalarial drugs available on the market. The proposed method was thus developed for the screening of 19 AAIs: amodiaquine (AQ), arteether (AE), artemether (AM), artemisinin (ART), artesunate (AS), atovaquone (AT), chloroquine (CQ), cinchonine (CC), dihydroartemisinin (DHA), halofantrine (HF), lumefantrine (LF), mefloquine (MQ), piperaquine (PPQ), primaquine (PQ), proguanil (PG), pyrimethamine (PM), quinine (QN), sulfadoxine (SD) and sulfalene (SL). Calculated pKa are given in Table 1.

Nowadays, HPLC method development can be achieved using different methodologies. Some have already led to some commercial softwares (e.g. Drylab, ACD/LC simulator, Chromsword, Osiris). These softwares make use of chromatography-based theory such as solvophobic theory, linear solvent strength relationship, etc., to optimize the separation of sample mixtures while maintaining the number of test experiments to a minimum [9-15]. These strategies are generally fast and efficient. Nevertheless, in the current trend of Quality by Design (QbD), these softwares do not provide the ability to compute or estimate the robustness, also sometimes called ruggedness in some regulatory documents. Consequently, classical robustness tests have to be carried out at the end of the method validation phase to estimate the method ability to remain unaffected by small, but deliberate variations in method parameters [16-18]. In the present work, a distinct and innovative methodology combining design of experiments (DoE), independent component analysis (ICA) and design space (DS) – with regard to ICH Q8(R2) [19] – was used to simultaneously optimize the separation and estimate the method robustness over the whole experimental domain (i.e. the knowledge space) instead of around the optima only.

In the ICH pharmaceutical development guidelines, Q8(R2) [19], the DS is defined as "*the multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality*". Therefore, the multidimensional combination and interaction of input variable

correspond to a subspace, so-called the DS, where assurance of quality has to be proved. Thus, the DS is necessarily encompassed within the experimental domain which is the multidimensional space formed by the factor ranges used during method development. The main concept lying behind the ICH Q8(R2) definition of DS is assurance of quality (also known as quality risk management). Hence, an HPLC method development process which do not take into account the prediction error (i.e. the prediction accuracy [20] estimated from the experimental data) in order to manage the risk cannot be considered as QbD compliant and will not allow to identify or compute the DS. Even if ICH Q8(R2) is the most appropriated guideline when discussing about QbD and DS applied to pharmaceutical sciences, the given DS definition and the examples shown at the end of the document are divergent. It seems that the identification of the multidimensional zone where a quality criterion lies in given acceptance limits or is higher or lower than an advisedly selected threshold do not define the DS taking into account assurance of quality. In other terms, in LC, to predict the multidimensional region where $R_S \geq 1.5$, does not define a DS as only quality is predicted but not assurance of quality. Indeed, the risk assessment of not being within the acceptance limits is not carried out. On the other hand, using the probability for a given quality criterion to be in the acceptance limits is a way to define DS (e.g. $P(R_S \geq 1.5)$). When computing such a probability, the quality risk management is carried out. Interesting discussions about QbD and DS applied in LC [21] and in pharmaceutical development [22-24] were already published in the scientific literature. Some DoE-DS LC applications were also previously published [25-28].

Therefore, the determination of the DS for a LC method development implies to consider also the error on the studied responses and criteria. The variability of the retention times have to be taken into account during the development phase. These considerations hold completely with the QbD definition [19]. Furthermore, ICH Q8(R2) guideline states that "*working within the design space is not considered as a change. Movement out of the design space is considered to be a change and would normally initiate a regulatory post approval change process*". Consequently, the DS defines a zone of robustness as no significant changes in terms of separation quality should be observed on the resulting chromatograms.

DoE strategy can be considered, by some pure chromatographists, as a "black box". Indeed, fitted mathematical models are only an approximation of the "true" chromatographic behaviour of investigated compounds. Nevertheless, these models are very useful to identify and study the chromatographic behaviour of compounds under investigation which can be unknown molecules. The most interesting advantage is that the DoE strategy is an overall data-driven methodology which does not necessarily imply preliminary knowledge of chromatographic behaviour of compounds under consideration and/or the understanding of their chromatographic parameters before starting the optimization process. In most cases, the application of chromatographic theories will lead to very good results. But when molecules are unknown or when their chromatographic behaviours are trickier than expected, they can lead to unpredicted and inoperable results.

Therefore, in this work, an innovative methodology using design of experiments (DoE) and independent component analysis (ICA) was used to optimise the separation and identify the DS for the above mentioned AAIs. The present study is a useful and relevant application of complementary strategies previously published [25,29]. The overall objective of the present work was both to demonstrate the ability of the DoE-ICA-DS methodology to provide optimal and robust HPLC method and to apply it as a case study to the development of a useful method for the screening of 19 antimalarial drugs in the framework of the fight against counterfeit medicines.

Table 1 Pka (at 25 °C) found on SciFinder®, calculated using Advanced Chemistry Development (ACD/Labs) Software V11.02.

Compounds	pKa		Compounds	pKa	
	Most acidic	Most Basic		Most acidic	Most basic
amodiaquine	9.43±0.50	5.62±0.50	lumefantrine	13.44±0.20	8.71±0.50
arteether	NA	NA	mefloquine	12.81±0.20	9.24±0.10
artemether	NA	NA	piperaquine	NA	8.92±0.50
artemisinin	NA	NA	primaquine	NA	10.38±0.10
artesunate	4.28±0.17	NA	proguanil	NA	11.15±0.10
atovaquone	5.01±0.10	NA	pyrimethamine	NA	7.18±0.10
chloroquine	NA	10.47±0.25	quinine	12.80±0.20	9.28±0.70
cinchonine	12.98±0.20	9.33±0.70	sulfadoxine	6.16±0.50	2.18±0.10
dihydroartemisinin	12.61±0.70	NA	sulfalene	6.61±0.40	1.46±0.10
halofantrine	13.57±0.20	9.44±0.50			

2. Experimental

2.1. Chemicals and reagents

Methanol (HPLC gradient grade), hydrochloric acid, ammonium hydroxide and ammonium hydrogen carbonate (99%) were purchased from Merck (Darmstadt, Germany). Ammonium formate (99%) was provided by Alfa Aesar (Karlsruhe, Germany). Ultrapure water was obtained from a Milli-Q Plus 185 water purification system from Millipore (Billerica, MA, USA).

Artesunate (99.8%), arteether (99.5%), artemisinin (99.6%), artemether (99.8%) and dihydroartemisinin (100.0%) were purchased from Apoteket AB (Stockholm, Sweden). Lumefantrine was gratefully provided by Novartis international pharmaceutical Ltd. (Villevorde, Belgium). Amodiaquine hydrochloride (99%), chloroquine diphosphate (98%), cinchonine (99.9%), mefloquine hydrochloride (99%), quinine hydrochloride (96.1%), sulfadoxine (99.9%) and butylhydroxyanisole (Ph. Eur.) were purchased from Sigma Aldrich (St. Louis, MO, USA). Atovaquone (98.9%), proguanil hydrochloride (99.9%) and halofantrine were gratefully provided by GlaxoSmithKline (Genval, Belgium). Piperaquine was extracted

from the P-Gvaxin formulation from Bliss GVS Pharma (Mumbai, India). Sulfalene, methylparaben (101.1%), propylparaben (101.9%) and butylhydroxytoluene (Eur. Ph.) were purchased from Fagron NV/SA (Waregem, Belgium). Finally, Arte-plus® pharmaceutical formulation was obtained from KIM Pharma Ets. (Kinshasa, D.R. Congo).

2.2. Sample preparations

A mixture of the 19 AAIs was used and was prepared as follow. In a first step, 10 mg of AQ, AT, CC, CQ, HF, LF, QN, MQ, PG, PM, PPQ, PQ, SD and SL, were dissolved in a 100.0 ml volumetric flask with methanol. This solution was annotated S1. In a second step, 10 mg of AE, AM, ART, AS and DHA were dissolved in a 2.0 ml volumetric flask with a mixture S1-methanol (50:50, v/v). An aliquot of this solution was filtered with 0.20 µm PTFE syringe filtration disks into a vial for injection in the HPLC system.

Other groups of compounds were prepared in the same way. Group 1 (see section 3.3.) contained AE, AM, AQ, ART, AS, AT, CC, CQ, HF, LF, QN, MQ, PG, PM, PPQ, PQ, SD and SL. Group 2 (see section 3.3.) contained AE, AM, AQ, AS, AT, CC, CQ, DHA, HF, LF, QN, MQ, PG, PM, PPQ, PQ, SD and SL. One sub-mixture was also used (see section 3.5.) and contained AS, PM and SL only.

2.3. Experiments

Chromatographic experiments were performed on a Waters 2695 separation module coupled to a Waters selector valve 7678 and a Waters 996 Photodiode array detector (Waters, Milford, MA, USA). pH measurements were performed with a SevenEasy S20 pH meter (Mettler Toledo, Columbus, OH, USA). Buffers consisted in 10 mM pH 2.5 formic acid, pH 4.0 ammonium formate, pH 6.0 ammonium formate, pH 8.0 ammonium hydrogencarbonate and pH 10.0 ammonium hydrogencarbonate. The pH was adjusted to the desired value with concentrated formic acid or ammonia 35% aqueous solution.

Chromatograms were recorded at wavelengths ranging from 210 nm to 400 nm with a step of 1.2 nm and with an acquisition time of 500 ms. Separations were carried out on an XBridge C18 250 x 4.6 mm, 5 µm, combined with an XBridge guard column C18 20 x 4.6 mm, 5 µm, both provided by Waters. Experiments were carried out at a flow rate of 1.0 ml/min and injection volume was 10 µl. Peaks were integrated at 230 nm.

The choice of the chromatographic column is justified by two main reasons. First, the chemistry of the column (i.e. trifunctionally bonded C18) allows a relatively good retention for most of the organic compounds. Second, the silica matrix-ethylene bridges and the octadecyl triple bonding strengthen the silica stability in very acidic and alkaline media. Therefore, this column was selected to obtain a generic methodology (i.e. usable for most mixtures in pharmaceutical analytical sciences).

2.4. Design of experiments

In order to model the chromatographic behaviour of each peak, a full factorial design was selected. It comprised three factors: pH of the aqueous part of the mobile phase (*pH*), gradient time (t_G) to linearly modify to proportion of methanol from 5% to 95%, and column temperature (*T*). Factors and their respective levels are summarized in Table 2. A total of 45 experimental conditions (5×3^2) were defined by this full factorial design. As shown in Table 2, temperature range was quite narrow. This factor was investigated to test out how the method robustness depended on this factor rather than optimizing the separation while varying *T*. As the resulting screening HPLC method should be used in different African countries with various or more frequently no column oven, it is of first importance to evaluate the effect of this factor on the studied responses. Moreover, the temperature could not be higher than 35 °C because some molecules were found to be relatively unstable at higher temperature. Gradient time and pH ranges were expanded as much as possible in order to widen the experimental domain and to minimize the risk of not finding any good separation within it. XBridge C18 columns can sustain pH from 1 to 12. pH range was slightly narrowed from 2.5 to 10 in order to maintain suitable column lifetime. Gradient time range was also wide (from 20 to 60 min). These factors were selected to test out their effects on the selectivity (a priori, mainly by modifying pH) and to shorten the time of analysis (a priori, mainly by decreasing t_G while preserving acceptable separation).

The experiments at a same pH were carried out in row for evident practical reasons. Then, the within pH blocks experiments were conducted in a random order. It is preferable to carry out the experiments in a totally random order to avoid experimental biases. Nevertheless, the column equilibration and conditioning times when constantly changing mobile phase pH drastically increase the time devoted to achieve the DoE results. Furthermore, the pH measurement error can be assumed to be equal to 0.1%. Other error sources (i.e. mobile phase composition during gradient, temperature, etc.) generated higher response errors and the pH blocking did not lead to poor predictive errors. Its impact was considered as negligible in the present study.

The central point (i.e. pH 6.0, t_G = 40 min, *T* = 30 °C) was independently repeated twice (i.e. carried out thrice) with the preparation of new buffer and fresh mobile phase. The central points for lower and higher temperatures (i.e. pH 6.0, t_G = 40 min, *T* = 35 °C and pH 6.0, t_G = 40 min, *T* = 25 °C, respectively) were also repeated once (i.e. carried out twice).

Table 2 Factors and levels of the full factorial design

Factors	Levels				
pH	2.5	4.0	6.0	8.0	10.0
Gradient time (t_G, min)	20	40	60		
Temperature (*T*, °C)	25	30	35		

2.5. Independent component analysis

ICA is a statistical method allowing the numerical separation of sources maximizing the independence between them based on non-Gaussianity [30]. In chromatography, ICA was already used to numerically separate coeluting peaks in order to estimate their integration limits (i.e. the times at the beginning and end of a peak) [28,29]. Indeed, for coeluting peaks, when using a drop-line valley separator, the estimation of integration limits is highly biased. Then, the modelling of responses based on these biased times could lead to poor prediction accuracy. In order to avoid this situation, ICA was used to numerically separate coeluting peaks of antimalarial drugs. The integration limits of non–coeluting peaks were obtained manually on the chromatograms.

2.6. Modelling and optimisation methodology

2.6.1. Retention times modelling

In the current trend to provide useful tools to build in quality within the LC method development field, as advocated by QbD guidelines, Dewé et al. [31] and Lebrun et al. [25] provided a new approach for retention times modelling and DS computation. In these works, they showed that the modelling of the resolution ($R_{S,crit} = 2\cdot(t_{R,2} - t_{R,1})/(w_{b,1} + w_{b,2})$; with $t_{R,2} > t_{R,1}$ and $w_{b,1}$, $w_{b,2}$ are the baseline peaks widths; $t_{R,1}$ and $t_{R,2}$ being the retention times of the critical pair peaks (i.e. the two most proximate peaks)) can lead to poor prediction caused by its non-linear and non-continuous behaviour when selectivity drastically changes. Thus, the modelling of retention times and the computation of a separation quality representative criterion (see section 2.6.2.) is a first step towards QbD in LC method development.

Hence, the retention times at the beginning, the apex and the end of each peak (respectively t_B, t_R and t_E) were measured. The studied responses were the logarithm of the retention factor (i.e. $\log(k_{tR})$ with $k_{tR} = (t_R - t_0)/t_0$, t_0 being the column dead time) and the logarithm of both half-widths (i.e. $\log(w_l)$ and $\log(w_r)$; with $w_l = k_{tR} - k_{tB}$ and $w_r = k_{tE} - k_{tR}$). These responses were modelled by an identical polynomial equation (see Eq. 1) using a multiple linear stepwise regression maximizing the adjusted coefficient of determination (R^2_{adj}).

$$\begin{aligned}\log(k_{tR}) = {} & \beta_0 + \beta_1 \times pH + \beta_2 \times pH^2 + \beta_3 \times pH^3 + \beta_4 \times t_G \\ & + \beta_5 \times {t_G}^2 + \beta_6 \times T + \beta_7 \times T^2 + \beta_8 \times pH \times t_G \\ & + \beta_9 \times pH \times T + \beta_{10} \times T \times t_G + \beta_{11} \times pH \times t_G \times T + \varepsilon\end{aligned} \quad (1)$$

2.6.2. Quality criterion, error propagation and design space computation

After the responses modelling, the retention times were predicted using Eq. 1. The prediction error was considered as a Gaussian distribution centred on 0 and with a standard deviation equal to the standard deviation of the residuals distribution (i.e. the difference between observation and prediction). Thus, for each operating condition, the predicted

responses were normally distributed and the predicted retention time for a compound followed a log-Normal distribution.

A quality criterion, also known as critical quality attribute, was selected to assess the quality of a separation. Lebrun et al. [25] proposed to use the separation criterion (*S*) defined as the difference between the beginning of the second peak and the end of the first peak of the critical pair ($S_{crit} = t_{B,2} - t_{E,1}$; with $t_{B,2} > t_{E,1}$; $t_{E,1}$ and $t_{B,2}$ being the end time and the beginning time of the critical pair peaks, respectively). The computation of this latter criterion only requires a subtraction while the one of R_S involves a division, sums and subtractions. The separation criterion is clearly easier to compute and to interpret. If $S_{crit} \geq 0$, critical pair peaks are baseline-resolved.

Then, the prediction errors obtained for the retention times are propagated to the quality criterion. Practically, Monte Carlo simulations were used to obtain, for a given operating condition, the distribution of *S* from the distributions of t_R, t_B and t_E of the critical pair peaks. Finally, using the *S* distribution, the probability for *S* to be higher than a selected threshold was used to determine the DS. In mathematical terms, the DS can be defined by Eq. 2.

$$\mathrm{DS} = \left\{ x_0 \in \chi | E_{\hat{\theta}} \left[\mathrm{P}(S > \lambda) | \hat{\theta} \right] \geq \pi \right\} \quad (2)$$

Where x_0 is a point of the experimental domain (χ). λ is the acceptance limit for criterion *S*, π is the quality level and $\hat{\theta}$ is the set of estimated parameters of the model. P and E respectively correspond to the estimators of probability and mathematical expectation. In other words, the DS defines a subspace where the probability to obtain baseline resolved peaks (i.e. $S > 0$ min) is higher than a predefined quality threshold (e.g. $\pi = 65\%$). In practice, π was selected to be 85% of the optimal probability to have $S > 0$. For instance, if at the optimal operating condition, $P(S > 0) = 70\%$, thus $\pi = 59.5\%$ was selected (i.e. 70%×85% = 59.5%). Obviously, in this example, finding a DS with a quality level of 95% would be more desirable. No regulatory document yet provides guidelines on how to compute or estimate the design space quality level [28]. Nevertheless, a high π can only be obtained with a relatively small prediction error compared to the difference between the end and the beginning of the critical pair peaks. Therefore, a high $P(S > 0)$ induces the achievement of robust chromatographic methods.

2.6.3. Optimal separation prediction

Finally, finding the optimal separation resumes in finding the operating condition which maximises the probability for *S* to be higher than the selected threshold. In practice, the use of nonlinear optimization techniques (e.g. Nelder-Mead method [32]) can lead to local optima. In order to avoid this non-optimal situation, the experimental domain was investigated with a grid search method. A multidimensional grid was defined over the experimental domain. Then, the quality criterion value and its associated prediction error (the distribution of *S* obtained propagating the error) were computed for each of the experimental condition defined by the grid. The optimum was selected as the point of the

grid giving the highest probability value. In order for this approach to accurately find the global optima, the grid has to contain numerous points. Nonetheless, the number of points also determines the computing time and this latter increased with the nth power of the number of points per factors (n being the number of factor). In practice, the number of points was set as high as possible while keeping the total computing time lower than 12 hours (i.e. one night computation).

2.7. Software

An in-house computer program was developed to perform the retention times modelling with stepwise multiple linear regressions, the error propagation and the grid search method. The coding was carried out with R 2.11.1 [33]. ICA-based numerical separations were performed using FastICA algorithm [34].

3. Results and discussion

3.1. Peak detection and peak matching

The first step after recording the chromatograms was the detection and the matching of each peak. This step is usually tedious and time-consuming and this complexity is proportional to the number of compounds and to the similarity between the UV spectra when UV-DAD detection is used. As the compounds from the artemisinin group (i.e. AE, AM, AS and DHA) present very similar and non-specific UV spectra, these five molecules were injected individually to identify and match them. Then, for the rest of the compounds, in case of coelution, ICA was used to determine the times at the beginning (t_B), the apex (t_R, the retention time) and the end (t_E) of each peak [29]. Finally, for the non-coeluting peaks, these times were manually read on the chromatograms. Examples of coeluted peaks numerical separation using ICA can be found in previous works [28,29].

Furthermore, at alkaline pH and high temperature (pH 10, T = 35 °C), DHA peak was split in two coeluting peaks and several unidentified peaks were also observed. These results suggested that DHA degraded in alkaline conditions at 35 °C. Therefore, pH 10 experiments were not used for responses modelling. Full factorial designs offer the advantage to give others full factorial designs when levels are removed as it was the case in the present study.

3.2. Retention times modelling

Retention times modelling were achieved by stepwise multiple linear regressions which selected the terms of Eq. 1 to maximize R^2_{adj}. As three times (t_B, t_R and t_E) of 19 peaks were modelled, 54 models were obtained. The estimated model parameters ($\beta_0 \ldots \beta_{11}$) are presented in Table 3 and the R^2_{adj} are summarized in Table 4.

Generally, the terms of a model are considered statistically significant if their respective student's t-test p-value is lower than 0.01 (i.e. α = 1%). In the present study, the terms were considered highly significant when their p-value was smaller than 0.001 (see Table 3). In Table 3, some p-values can seem high but the corresponding terms were selected as they increased R^2_{adj} which was selected as the stepwise regression criterion [35].

The model parameters presented in Table 3 represent the chromatographic behaviour of the compounds under investigation. Therefore, the model parameters can be used to corroborate the chromatographic behaviours thanks to some physico-chemical properties such as pKa values. These interpretations are listed in following sections (3.2.1. to 3.2.3.). The objective of these considerations is not to highlight some quantitative structure retention relationships but to understand the resulting models thanks to some chromatographic behaviour.

3.2.1. pH effect

The retention time of acidic or basic compounds should respectively decrease or increase with respect to pH following a sigmoidal curve. To ease the interpretation of the model parameters given in Table 3, three cases should be envisaged.

First, if pKa is inside to the modelling pH range (MpHR; from 2.5 to 8), the sigmoidal curve inflexion point should be within MpHR. In this case, the *pH* term (β_1 is positive for bases and negative for acids; it can be close to 0) and mainly the pH^3 term (β_3 is positive for acids and negative for bases) should be highly expressed to mimic the theoretical sigmoidal variation of t_R with respect to pH. AQ, CC, LF, PM, PPQ and QN (basic compounds) and AS (acidic compound) presented t_R variations corresponding to this first case (see Fig. 1a).

Second, if pKa is outside and lower than MpHR, the truncated sigmoidal curve can be thus considered as a decreasing convex quadratic variation for acids and an increasing concave quadratic function for bases. A function is convex if and only if the region above its graph is a convex set. The *pH* term (β_1 is positive for bases and negative for acids) and principally the pH^2 term (β_2 is positive for acids and negative for bases) should be mainly expressed to fit the quadratic variation of t_R with respect to pH. PQ, a basic compound, presented a t_R variation which corresponded to both first and second case. The pKa corresponding to the quinoline function (pKa ~ 4.9) was inside MpHR (corresponding to case 1) but the second pKa (~ 10.4) was higher than MpHR (corresponding to case 2).

Third, if pKa is outside and higher than MpHR, the truncated sigmoidal curve can be thus considered as a decreasing concave quadratic variation for acids and an increasing convex quadratic function for bases. The *pH* term (β_1 is positive for bases and negative for acids) and principally the pH^2 term (β_2 is negative for acids and positive for bases) should be mainly expressed to fit the quadratic variation of t_R with respect to pH. AT, SD and SL (acidic compounds) and CQ, HF and CQ (basic compounds) displayed t_R variations corresponding to this third case.

The rest of the AAIs (AE, AM, ART and DHA) are neutral compounds, which did not present t_R variation with respect to pH. The above mentioned interpretations were in accordance with the predicted t_R variations with respect to pH presented in Fig. 1a.

It should be noted that the given pKa are estimated or measured for aqueous media. Therefore, they merely represent an estimation of the pKa that will determine the equilibrium between protonated and deprotonated forms in hydro–organic phase.

3.2.2. Gradient time and temperature effects

The general shape of t_R variation with respect to t_G is an increasing concave function which can easily be fitted by a second order polynomial (see Fig. 1b). All β_4 (corresponding to t_G) were positive and all β_5 (corresponding to t_G^2) were negative due to the concave curvature.

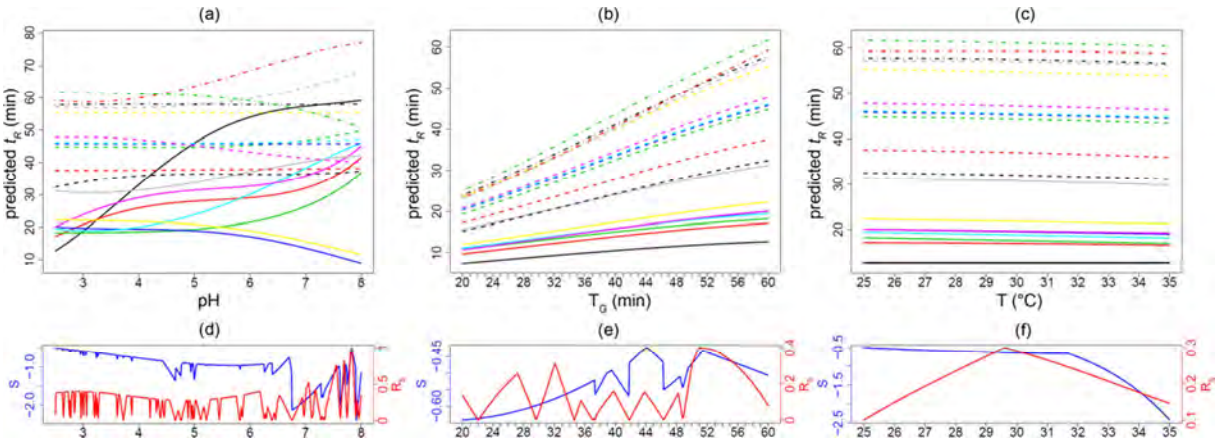


Fig. 1 Predicted retention times (t_R) with respect to DoE's factors. (a) predicted t_R (min) *vs pH* – with t_G=60 min and *T*=25 °C – , (b) predicted t_R (min) vs t_G (min) – with *pH* 2.5 and *T*=25 °C – and (c) predicted t_R (min) vs *T* (°C) – with *pH* 2.5 and t_G=60 min. Compound assignation: (black line) PPQ, (red line) CC, (green line) CQ, (blue line) SL, (cyan line) AQ, (magenta line) QN, (yellow line) SD, (grey line) PM, (dashed black line) PQ, (dashed red line) PG, (dashed green line) MQ, (dashed blue line) ART, (dashed cyan line) DHA, (dashed magenta line) AS, (dashed yellow line) AM, (dashed grey line) HF, (dotted-dashed black line) AE, (dotted-dashed red line) LF and (dotted-dashed green line) AT. (d,e,f) Corresponding predicted *S* (blue line, left axis) and R_S (red line, right axis).

Van't Hoff law defines the variation of an equilibrium constant with respect to the temperature. It is used to define the t_R variation with respect to *T*. In HPLC, a generally linear relationship links log(k_{tR}) with 1/*T*. A decreasing curved function should be observed for log(k_{tR}) *vs*. *T*. Besides, the selected temperature range was narrow and statistically significant t_R variations were not observed premising method robustness against temperature variation as expected (see Fig. 1c).

3.2.3. Models adequacy

In order to visualize the models adequacy, the appropriateness between predicted and experimental data as well as the corresponding residuals were displayed on Fig. 2. As observed on Fig. 2b, the residuals are distributed between –2 and 2 min. The residuals standard deviations were equals to 0.36, 0.41 and 0.36 min for t_R, t_E and t_B, respectively. Given this results, it is thus reasonable to assume that the retention time predictive error will be around 0.8 min (i.e. 2×standard deviation) over the whole experimental domain.

Furthermore, Shapiro-Wilk Normality tests were carried out on the modelled responses residuals. The p-values were all higher than 0.05 which meant that the residuals were actually normally distributed.

From Table 4, one can observe that the R^2_{adj} were higher for t_R than for t_E and t_B. This can be explained by two main reasons. First, the times at the peaks beginning and end could be biased by a poor estimation of integration limits. Even when ICA was applied, the independent component baseline can be distorted at the side of the coelution [28,29] leading to an overestimation of t_B or an underestimation of t_E. Second, w_l and w_r are computed from 2 measurements. For instance, the computation of w_l involved t_B and t_R whereas the computation of k_{tR} only involved t_R thus explaining the lower observed R^2_{adj} values.

The smallest R^2_{adj} value is equal to 0.053 for the right half-width of PM peak. This small value did not reflect a poor data adjustment. It only reflected the fact that PM peak right half-width was poorly influenced by the experimental factors and that the response variability was higher than its average variation. As observed on Fig. 2, no aberrant results were observed for this compound or for others (i.e. no significant outliers observed on residuals plots).

Once the models qualities were checked, the predicted chromatographic behaviours were assessed. Each model was then used to predict t_R according to DoE factors. As expected and presented in Fig. 1a, *pH* is the factor that had the most significant effect on selectivity. The identification of neutral, acidic or basic compounds is also easy. Neutral compounds show no t_R variation with respect to *pH*. Acidic and basic compounds have a respective decreasing or increasing variation with respect to *pH*. The huge t_R increasing variation of PPQ can be explained by the four basic nitrogen on the piperazine cycles lying in the PPQ structures.

Table 3 Estimated model parameters (β_0… β_{11}, see Eq. 1) and p-values (Pr(>|t|)) of their significance tests for log(k_{tR}) modelling of the studied compounds: amodiaquine (AQ), arteether (AE), artemether (AM), artemisinin (ART), artesunate (AS), atovaquone (AT), chloroquine (CQ), cinchonine (CC), dihydroartemisinin (DHA), halofantrine (HF), lumefantrine (LF), mefloquine (MQ), piperaquine (PPQ), primaquine (PQ), proguanil (PG), pyrimethamine (PM), quinine (QN), sulfadoxine (SD) and sulfalene (SL).

	Compounds													
	AQ		AE		AM		ART		AS		AT		CQ	
Term	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)
Intercept	4.557	**< 0.001**	5.323	**< 0.001**	5.282	**< 0.001**	5.107	**< 0.001**	5.081	**< 0.001**	5.358	**< 0.001**	4.323	**< 0.001**
pH	0.594	**< 0.001**	-0.004	0.468	0.000	0.338	0.001	0.558	-0.120	**< 0.001**	-0.050	**< 0.001**	0.212	**< 0.001**
pH^2	0.173	**< 0.001**	NS	NS	0.001	0.112	0.001	0.300	-0.023	**< 0.001**	-0.077	**< 0.001**	0.276	**< 0.001**
pH^3	-0.195	**< 0.001**	0.006	0.327	NS	NS	-0.001	0.712	0.039	**< 0.001**	-0.041	**< 0.001**	0.114	**< 0.001**
t_G	0.336	**< 0.001**	0.439	**< 0.001**	0.428	**< 0.001**	0.406	**< 0.001**	0.406	**< 0.001**	0.445	**< 0.001**	0.309	**< 0.001**
t_G^2	-0.081	**< 0.001**	-0.103	**< 0.001**	-0.103	**< 0.001**	-0.096	**< 0.001**	-0.096	**< 0.001**	-0.106	**< 0.001**	-0.073	**< 0.001**
T	-0.016	0.004	-0.006	**< 0.001**	-0.011	**< 0.001**	-0.013	**< 0.001**	-0.016	**< 0.001**	-0.010	**< 0.001**	-0.021	**< 0.001**
T^2	NS	NS	-0.005	0.052	-0.002	0.006	-0.003	0.006	-0.003	0.207	-0.001	0.457	NS	NS
$pH \times t_G$	0.053	**< 0.001**	NS	NS	0.000	0.610	0.001	0.401	-0.010	**< 0.001**	-0.007	**< 0.001**	0.043	**< 0.001**
pH×T	0.019	0.010	NS	NS	0.001	0.296	NS	NS	-0.003	0.103	NS	NS	0.015	0.036
$t_G \times T$	NS	NS	-0.004	0.042	-0.002	**< 0.001**	-0.002	0.004	-0.003	0.115	-0.001	0.497	NS	NS
$pH \times t_G \times T$	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

	Compounds											
	CC		DHA		HF		LF		MQ		PPQ	
Term	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)
Intercept	4.673	**< 0.001**	5.112	**< 0.001**	5.329	**< 0.001**	5.450	**< 0.001**	5.077	**< 0.001**	5.197	**< 0.001**
pH	0.039	0.010	0.006	0.018	0.101	**< 0.001**	0.253	**< 0.001**	0.023	**< 0.001**	0.428	**< 0.001**
pH^2	-0.058	**< 0.001**	0.013	**< 0.001**	0.074	**< 0.001**	0.043	**< 0.001**	0.052	**< 0.001**	-0.576	**< 0.001**
pH^3	0.352	**< 0.001**	0.004	0.079	0.011	0.178	-0.071	**< 0.001**	0.032	**< 0.001**	0.244	**< 0.001**
t_G	0.337	**< 0.001**	0.405	**< 0.001**	0.435	**< 0.001**	0.408	**< 0.001**	0.422	**< 0.001**	0.375	**< 0.001**
t_G^2	-0.077	**< 0.001**	-0.097	**< 0.001**	-0.100	**< 0.001**	-0.088	**< 0.001**	-0.100	**< 0.001**	-0.085	**< 0.001**
T	-0.012	0.003	-0.013	**< 0.001**	-0.008	**< 0.001**	-0.013	**< 0.001**	-0.011	**< 0.001**	NS	NS
T^2	-0.007	0.310	-0.003	0.012	-0.004	0.262	-0.005	0.314	-0.004	**< 0.001**	NS	NS
$pH \times t_G$	0.056	**< 0.001**	0.001	0.191	-0.022	**< 0.001**	-0.057	**< 0.001**	0.002	0.008	0.102	**< 0.001**
pH×T	NS	NS	0.001	0.267	NS	NS	-0.012	0.003	0.004	**< 0.001**	NS	NS
tG×T	-0.005	0.299	-0.002	0.030	NS	NS	0.003	0.404	-0.001	0.109	NS	NS
$pH \times t_G \times T$	NS	NS	NS	NS	NS	NS	0.006	0.208	NS	NS	NS	NS

	Compounds											
	PQ		PG		PM		QN		SD		SL	
Term	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)	Estim.	Pr(>|t|)
Intercept	4.897	**< 0.001**	4.926	**< 0.001**	4.853	**< 0.001**	4.763	**< 0.001**	4.394	**< 0.001**	4.325	**< 0.001**
pH	-0.005	0.372	0.011	**< 0.001**	0.232	**< 0.001**	0.056	**< 0.001**	-0.177	**< 0.001**	-0.133	**< 0.001**
pH^2	-0.047	**< 0.001**	0.002	0.116	0.048	**< 0.001**	-0.053	**< 0.001**	-0.241	**< 0.001**	-0.331	**< 0.001**
pH^3	0.070	**< 0.001**	-0.001	0.778	-0.098	**< 0.001**	0.306	**< 0.001**	-0.102	**< 0.001**	-0.206	**< 0.001**
t_G	0.379	**< 0.001**	0.385	**< 0.001**	0.370	**< 0.001**	0.356	**< 0.001**	0.260	**< 0.001**	0.226	**< 0.001**
t_G^2	-0.090	**< 0.001**	-0.092	**< 0.001**	-0.086	**< 0.001**	-0.080	**< 0.001**	-0.063	**< 0.001**	-0.058	**< 0.001**
T	-0.015	**< 0.001**	-0.019	**< 0.001**	-0.016	**< 0.001**	-0.015	**< 0.001**	-0.026	**< 0.001**	-0.027	**< 0.001**
T^2	-0.006	0.027	-0.007	**< 0.001**	-0.007	0.029	NS	NS	NS	NS	NS	NS
$pH \times t_G$	0.007	0.003	-0.001	0.204	0.024	**< 0.001**	0.046	**< 0.001**	-0.053	**< 0.001**	-0.071	**< 0.001**
pH×T	-0.002	0.343	NS	NS	0.007	0.006	NS	NS	NS	NS	NS	NS
$t_G \times T$	-0.005	0.025	-0.003	0.005	NS	NS	-0.006	0.168	NS	NS	NS	NS
$pH \times t_G \times T$	0.004	0.167	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

NS: non-selected during stepwise regression.

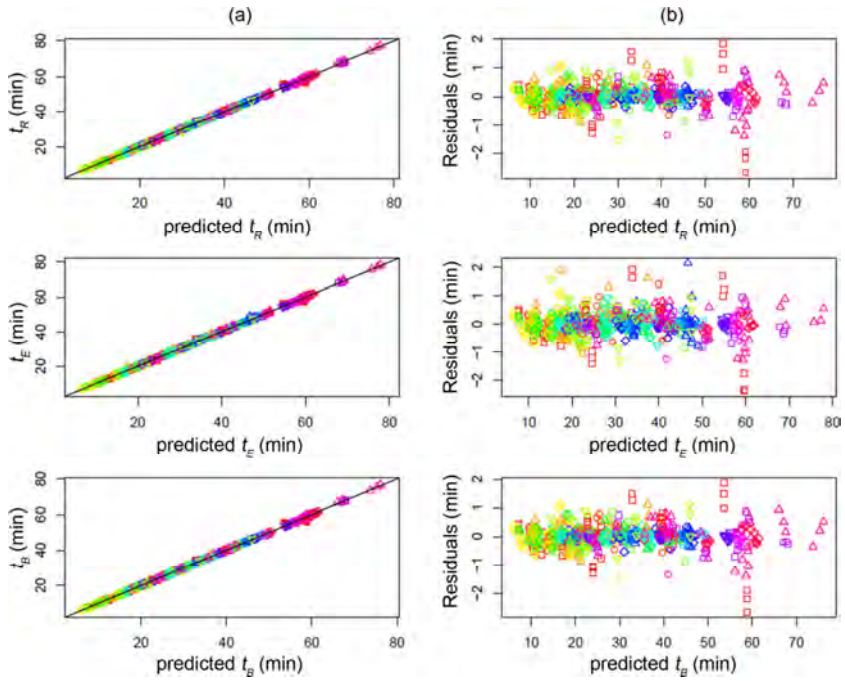


Fig. 2 Modelling results and corresponding residuals. (a) Predicted versus experimental values for t_R, t_E and t_B. (b) Corresponding residuals plots. Compound assignation: (red square) PPQ, (red circle) CC, (orange head up triangle) CQ, (yellow diamond) SL, (green head down triangle) AQ, (green square) QN, (green circle) SD, (green head up triangle) PM, (green diamond) PQ, (blue head down triangle) PG, (blue square) MQ, (blue circle) ART, (blue head up triangle) DHA, (blue diamond) AS, (blue head down triangle) AM, (purple square) HF, (purple circle) AE, (purple head up triangle) LF, (purple diamond) AT.

Table 4 Adjusted coefficients of determination (R^2_{adj}) of the stepwise multiple linear regressions for the 19 AAIs retention factors logarithm (log(k_{tR})), left half-width logarithm (log(w_l)) and right half-width logarithm (log(w_r)).

Compounds	R^2_{adj}			Compounds	R^2_{adj}		
	log(k_{tR})	log(w_l)	log(w_r)		log(k_{tR})	log(w_l)	log(w_r)
Piperaquine	0.997	0.813	0.491	Mefloquine	1.000	0.651	0.391
Cinchonine	0.997	0.656	0.456	Artemisinin	1.000	0.571	0.896
Chloroquine	0.995	0.483	0.667	Dihydroartemisinine	1.000	0.932	0.622
Sulfalene	0.991	0.701	0.373	Artesunate	1.000	0.963	0.904
Amodiaquine	0.996	0.484	0.433	Artemether	1.000	0.299	0.726
Quinine	0.998	0.586	0.762	Halofantrine	0.999	0.842	0.643
Sulfadoxine	0.995	0.870	0.618	Arteether	0.999	0.558	0.635
Pyrimethamine	0.999	0.534	0.053	Lumefantrine	0.998	0.975	0.889
Primaquine	0.999	0.737	0.360	Atovaquone	1.000	0.411	0.284
Proguanil	1.000	0.636	0.561	*Mean*	*0.998*	*0.669*	*0.567*

The temperature has the lowest effect (see Fig. 1c). It was selected as DoE factor in order to assess method robustness with respect to it. In other terms, these low t_R variations with respect to *T* underline the method robustness while changing *T* as expected routine use in Africa.

In order not to overload Fig. 1, only t_R are plotted. Nevertheless, predicted t_B and t_E were also computed to predict the selected criteria at given operating conditions. For instance, Fig. 1d, 1e and 1f show the behaviour of the predicted quality criteria S and R_S.

One can also observe from Fig. 1d, 1e and 1f that predicted S and R_S share some similarities. The main dissimilarity between these two critical quality attributes is that S only involves the difference between the peak limits (i.e. t_B and t_E) while R_S involves the difference between the peak apexes (i.e. t_R) and the width of both peaks. For instance, for 2 pairs of peak being equally separated (i.e. same $S > 0$), if the second peak width is larger in the first pair than in the second, it implies that R_S should be smaller for the first pair although the separation is acceptable for both pairs (i.e. same $S > 0$). In other terms, it seems to be risky to use R_S to evaluate the quality of a separation only, as it involves the peak width. Nevertheless, R_S is widely used and accepted by the scientific community to be the golden standard quality criterion in HPLC. Future works should compare R_S with the combination of S and peaks widths (i.e. multi-criteria approach) in order to determine which one could be the most appropriated in terms of method development and separation optimization. On Fig. 1d, 1e and 1f, the criteria curves break points occur when critical pair peaks change and discontinuities are resulting from distinct peak asymmetry. When 2 peaks interbred (i.e. selectivity change), criteria suddenly change discontinuously because the left and right half-widths of the critical pair peaks are different. These three figures (i.e. Fig. 1d, 1e and 1f) also clearly show why R_S or S should not be selected as multiple linear model responses but only as critical quality attributes.

3.3. Quality criterion and design space computation

The separation quality criterion (S) was computed over the whole experimental domain. A grid of 42875 points (i.e. 35×35×35) was defined and S was computed for each operating condition. The residuals distribution was used to generate a Gaussian predictive error (i.e. the Gaussian distribution standard deviation was set equal to the residuals standard deviation, see section 3.2.3.), which was used to propagate the predictive error to S using Monte-Carlo simulations [28,36]. 2500 Monte-Carlo simulations were thus carried out for each of the 42875 points. Then, the results were presented as probability surfaces (i.e. the probability for a criterion to lie in a range of values or to be higher or smaller than a threshold) rather than response surfaces.

The very similar chromatographic behaviour between ART and DHA prevented obtaining a separation of the 19 AAIs. Nevertheless, at some operating conditions, they were the only two coeluting peaks. Thus, two groups were formed. The first contained 17 AAIs and ART (group 1) and the second contained the same 17 AAIs and DHA (group 2). These two groups are justified from a therapeutic point of view because ART and DHA are never present in the same pharmaceutical formulation. The optimisation process was then repeated for each group independently. The probability surfaces for $P(S > 0)$ for group 1 and group 2 are presented in Fig. 3 and Fig. 4, respectively.

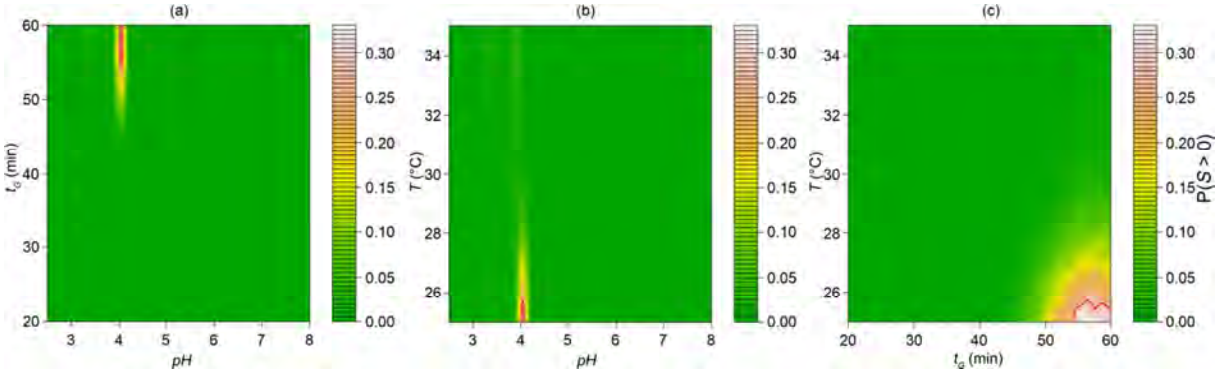


Fig. 3 Probability surfaces (i.e. P(S > 0)) for group 1 separation. (a) Gradient time (min) *vs* pH, (b) temperature (°C) *vs* pH and (c) temperature (°C) *vs* gradient time (min). The DS (π = 28%) is encircled by a red line.

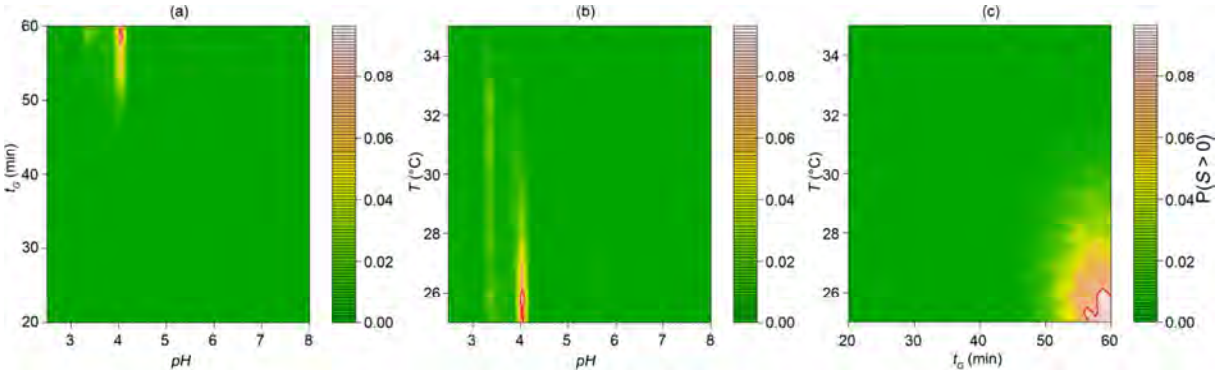


Fig. 4 Probability surfaces (i.e. P(S > 0)) for group 2 separation. (a) Gradient time (min) *vs* pH, (b) temperature (°C) *vs* pH and (c) temperature (°C) *vs* gradient time (min). The DS (π = 8%) is encircled by a red line.

For a given π, the DS shape is directly linked to the method robustness. Therefore, the DS shapes on Fig. 3 and Fig. 4 allowed concluding that the resulting screening methods are relatively highly robust with respect to modifications in t_G (from 54 to 60 min for group 1 and from 58 to 60 min for group 2) and T (from 25 to 25.8 °C for group 1 and from 25 to 26 °C for group 2). Conversely, these screening methods are far less robust with respect to *pH* (from pH 4 to pH 4.1 for both groups). However, the pH measurement variability can be estimated to 0.1% (see section 2.4.). Consequently, the relatively poor method robustness with respect to *pH* should therefore not be problematic when care is taken during the buffers preparation and pH measurements.

3.4. Optimal separation prediction

As shown on Fig. 3 and 4, the probability surface and the corresponding DS are very similar for the two groups of investigated compounds. Using group 1 or group 2 (i.e. using ART or DHA, respectively), the optimization process give quite identical optimal operating condition. The main difference is that the quality levels are different: π = 28% for group 1 and π = 8% for group 2. The DS shape similarity resulted from the analogous chromatographic behaviours between ART and DHA which were coeluting together. This

difference between quality levels for group 1 and 2 came from the shorter DHA retention times compared to ART. DHA slightly coeluted with MQ (Peak 12 on Fig. 5) explaining the much lower π for group 2.

Finally, only one operating condition was selected allowing the separation of the 18 AAIs of both groups independently. Obviously, the optimal separation was obtained for an operating condition located in the DS. The optimal operating condition giving the highest probability to have $S > 0$ (P = 33% for group 1 and P = 9.6% for group 2) was *pH* 4.05, t_G = 56.2 min and T = 25 °C. This probability could seem low. But it is important to keep in mind that, at this optimal condition, S depended on $t_{E,2}$ and $t_{B,3}$ of the two most proximate peaks or critical pair (Peak 2 and 3 in Fig. 5, SL and AQ, respectively). As mentioned in section 3.2., the predictive error was estimated at 0.8 min for both $t_{E,2}$ and $t_{B,3}$. As it can be seen from Fig. 5b, the predicted difference between $t_{E,2}$ and $t_{B,3}$ is smaller than the estimated predictive error. It resulted that for some of the 2500 simulations carried out at the optimal operating condition (33% of the 2500 simulations), $t_{E,2}$ was actually smaller than $t_{B,3}$ (i.e. P($S > 0$) = 33%). This does not mean that the probability to obtain a good separation at this operating condition is about 33% (i.e. suggesting a poor repeatability). It means that realizing 100 individual method developments following the same methodology as done in this work, 33 of them would have allowed obtaining good separations, i.e. $S > 0$. The 67 other method developments would have resulted in chromatograms with slightly coeluted peaks. The chromatogram recorded at the optimal condition is depicted on Fig. 5a for group 1 compounds.

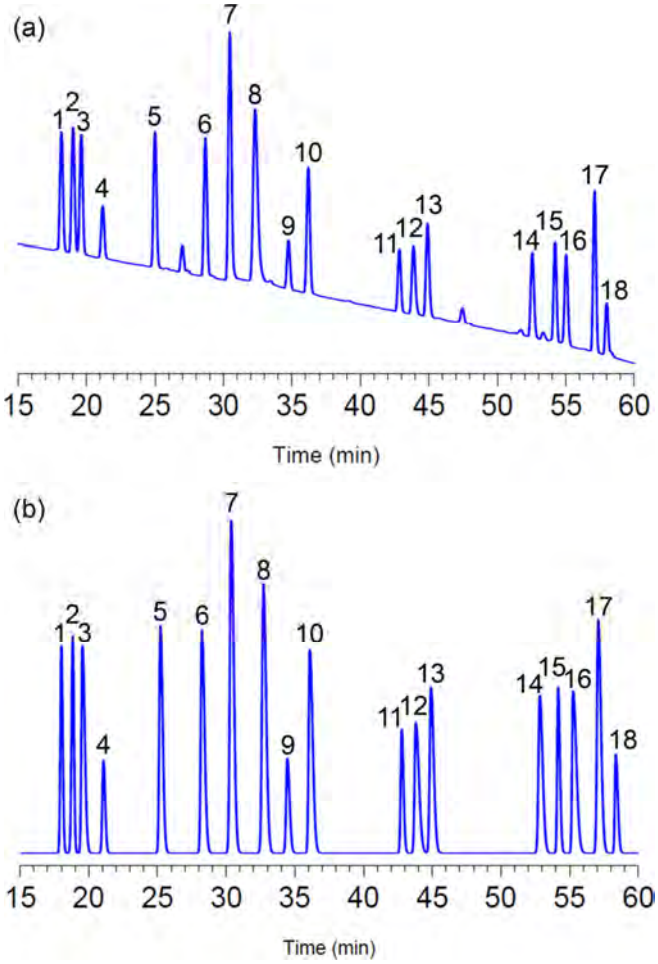


Fig. 5 (a) Experimental chromatogram recorded at *pH* = 4.05, t_G = 56.2 min and T = 25°C with group 1 compounds. (b) Respective predicted chromatogram. Peak numbering: 1 = CQ, 2 = SL, 3 = AQ, 4 = SD, 5 = CC, 6 = QN, 7 = PM, 8 = PP, 9 = PQ, 10 = PG, 11 = MQ, 12 = ART, 13 = AS, 14 = AM, 15 = HF, 16 = AE, 17 = LM and 18 = AT.

Chromatograms recorded for group 2 compounds at the optimal operating condition are very close to those displayed in Fig. 5a for group 1. Despite the inability to separate ART and DHA (inducing the creation of 2 groups), it was easy to identify them even if coeluting because their retention times were different enough allowing ART and DHA distinction (i.e.

t_R = 44.5 min for ART and t_R = 44.9 min for DHA). It is still possible to separately inject a reference solution of each compound in working conditions to confirm their identification.

3.5. Sub-mixture

One can observe that the screening method (Fig. 5) has a quite long analysis time. It can be also observed that the compounds eluted between 15 and 60 min. The first 15 min are not "used" to separate the compounds in a shorter time. It mainly resulted from the fact that the methanol proportion at the start of the gradient (*%MeOH*$_{ini}$) was low (i.e. *%MeOH*$_{ini}$ = 5%) and was not introduced as a DoE factor. *%MeOH*$_{ini}$ was not added as a DoE factor to keep the number of experiments to an acceptable value (i.e. 45 experiments). The introduction of *%MeOH*$_{ini}$ would have increased the number of experiments to 135 (i.e. 3×45 to estimate the quadratic effect of *%MeOH*$_{ini}$) for a full factorial design.

Nevertheless, the DoE-DS methodology can also be used to develop methods aiming at reducing the analysis time while optimizing the separation of specific sub-mixtures of compounds. Some compounds (related to a given pharmaceutical formulation) were therefore selected to test out this opportunity without performing any additional experiments. Indeed, once the compound chromatographic behaviours were modelled (see section 3.2.), the optimization process for the separation of any compound combinations could be carried out. The resulting DS would directly depend on the compound involved in the computation of *S* and P(*S* > 0). The selected sub-mixture contained AS, PM and SL. This combination is representative of a pharmaceutical formulation present on the Democratic Republic of the Congo market.

In order to minimize the time of analysis while simultaneously optimizing the separation of these 3 compounds, a multi-criteria optimization was carried out. The criteria were the separation criterion *S* and the time of analysis (i.e. the retention time at the end of the last peak). The thresholds were placed at 0.1 min for *S* and at 20 min for a more convenient analysis time. The DS thus defined the subspace where the probability to obtain *S* > 0.1 min concurrently with an analysis time < 20 min was higher than 97.5% as shown on Fig. 6.

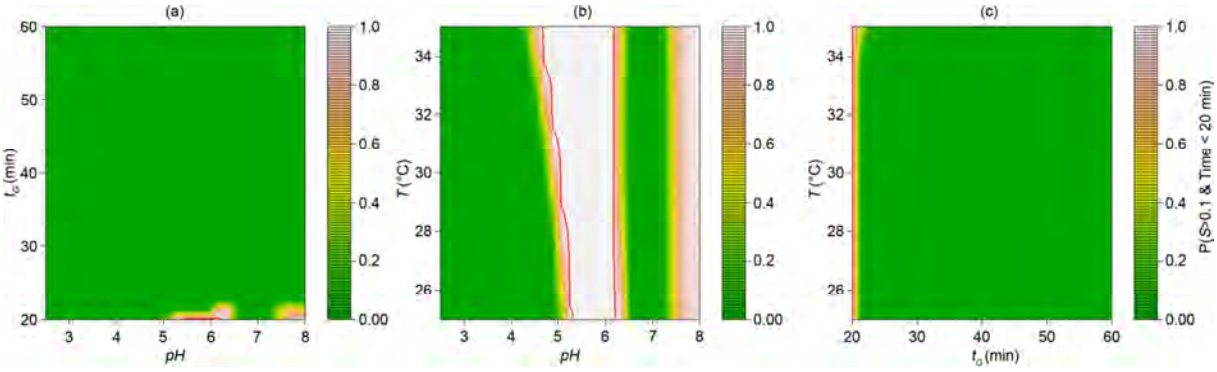


Fig. 6 Probability surfaces (i.e. P(*S* > 0.1 & analysis time < 20 min)) for sub-mixture separation. (a) Gradient time (min) *vs* pH, (b) temperature (°C) *vs* pH and (c) temperature (°C) *vs* gradient time (min). The DS (π = 97.5%) is encircled by a red line.

One can observes that the DS depicted in Fig. 6a and Fig. 6c is really small. The reason behind this observation is the closeness of the last peak with the analysis time threshold (i.e. < 20 min). Fig. 6b does not take t_G into account but only *T* vs. *pH*, the DS is therefore not narrowed due the last peak t_E (see Fig. 7). *T* and *pH* had a much lower effect on t_R of AS (see Fig. 2). Due to the huge DS size and the high π, the specific method is very robust in terms of separation. Concerning analysis time, the method is far less robust with respect to t_G which has the highest effect on compounds retention times. The optimal operating condition is *pH* 5.4, t_G = 20 min and *T* = 35 °C. The recorded chromatogram at the optimal condition is displayed in Fig. 7.

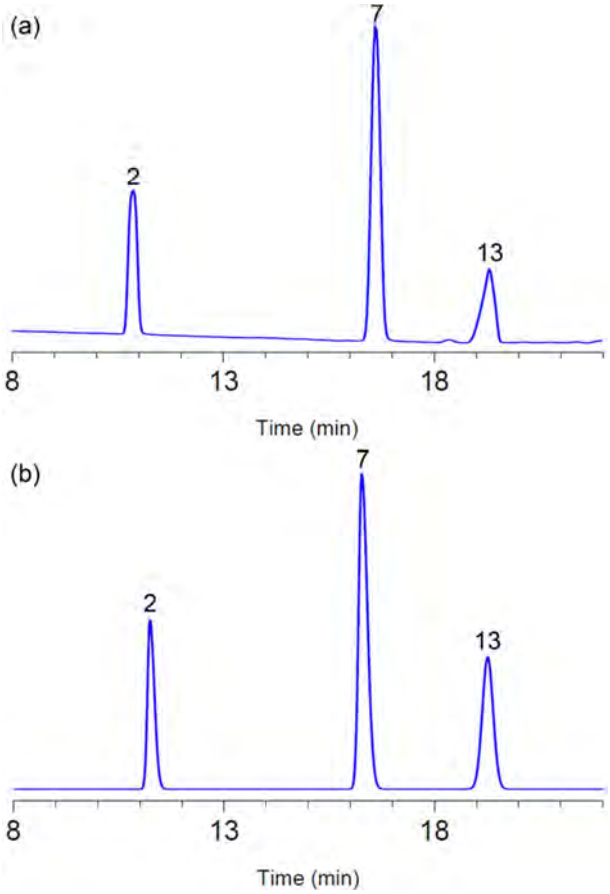


Fig. 7 (a) Experimental chromatogram recorded at *pH* = 5.4, T_G = 20 min and *T* = 35°C with sub-mixture compounds. (b) Corresponding predicted chromatogram. Peak numbering: 2 = SL, 7 = PM, 13 = AS.

Fig. 7 presents the predicted and experimental chromatograms. The adequacy between predicted and experimental t_R was very good. It is quite obvious that a method development (using already available commercial softwares or a 2^3 full factorial design with t_G and $\%MeOH_{ini}$ as factors) for the separation of these three compounds would be able to find operating condition giving a good separation in a shorter analysis time. However, this present method optimization dedicated to the pharmaceutical formulation (Arte-Plus) was carried out from the same data that those used for the optimization of the screening method. It underlined the fact that the DoE-DS methodology is generic. Here an innovative methodology was used to optimize the separation while simultaneously estimating the robustness of either a general screening method or specific methods. In this case, $\%MeOH_{ini}$ could also help reduce the time of analysis. For the reasons mentioned above, the resulting huge number of experiments hindered the use of 4 DoE factors. As mentioned earlier in the manuscript, temperature was preferred to $\%MeOH_{ini}$ in order to assess method robustness with respect to *T* rather than shortening the analysis while selecting $\%MeOH_{ini}$ as a DoE factor.

4. Conclusions

In the current trend of being QbD compliant, it is of first importance to develop methodologies that provide robust optimal separations defined by a DS. With regards to this objective, DoE-ICA-DS allowed obtaining optimal screening methods for the separation of 19 AAIs. The methods robustness was evaluated thanks to the DS quality level, DS shape and the assessment of the factors effects. It resulted that the obtained screening methods were very robust against temperature modification. This result is very important when one of the final aims is a method transfer to African laboratories where column ovens are not always available. Theses screening methods can be considered as a step forward for the fight against counterfeit medicines. The present work also allowed demonstrating the ability of the DoE-ICA-DS methodology to encounter optimal separation for complex mixtures (i.e. containing compounds with very similar structures and physico-chemicals properties). The present study also demonstrated the ability of fitted mathematical models to be used to identify and corroborate theoretical chromatographic behaviours of studied compounds. It highlighted the fact that DoE strategy can be a very useful tool for chromatographists in order to develop or refine the understanding of some chromatographic behaviour. Furthermore, the separation of a 3 compounds mixture was also carried out without performing any additional experiments. The resulting method was also very robust to temperature changes. Finally, the results presented in this manuscript strengthen the fact that DoE-ICA-DS can be considered as a generic QbD compliant methodology for the optimization and the robustness assessment of new chromatographic methods for the analysis of pharmaceutical formulations or more complex matrices such as plant or biological materials.

Acknowledgements

The authors would like to thank the Walloon Region of Belgium for the PPP convention funds N°917007 and Arlenda SA for partial funding of Optimal-DS project. A research grant from the Belgium National Fund for Scientific Research (F.R.S-FNRS) to E. Rozet is also gratefully acknowledged.

References

[1] World Malaria Report 2010, World Health Organization, (2010), http://whqlibdoc.who.int/publications/2010/9789241564106_eng.pdf

[2] M.K. Laufer, C.V. Plowe, Drug Resist. Update 7 (2004) 279.

[3] A.F. Cowman, S.J. Foote, Int. J. Parasitol. 95 (1990) 503.

[4] T.K. Mutabingwa, Acta Trop. 95 (2005) 305.

[5] C.O. Obonyo, E.A. Juma, B.R. Ogutu, J.M. Vulule, J. Lau, T. Roy. Soc. Trop. Med. H. 101 (2007) 117.

[6] N. Valecha, H. Joshi, P.K. Mallick, S.K. Sharma, A. Kumar, P.K. Tyagi, B. Shahi, M.K. Das, B.N. Nagpal, A.P. Dash, Acta Trop. 111 (2009) 21.

[7] N. Singh , M.M.Shukla, G. Chand, P.K. Bharti, M.P. Singh, M.K. Shukla, R.K. Mehra, R.K. Sharma, A.P. Dash, T. Roy. Soc. Trop. Med. H. In press (2011).

[8] R.D. Marini, J. Mbinze Kindenge, M.L.A. Montes, B. Debrus, P. Lebrun, J. Mantanus, E. Ziemons, S. Rudaz, Ph. Hubert, Chim. Oggi 28 (2010) 10. http://hdl.handle.net/2268/62379

[9] C. Horváth, W. Melander, I. Molnár, J. Chomatogr., A 125 (1976) 129.

[10] P.W. Carr, J. Li, A.J. Dallas, D.I. Eikens, L. Choo Tan, J. Chromatogr., A 656 (1993) 113.

[11] D. Nagrath, F. Xia, S.M. Cramer, J. Chromatogr., A In press (2011).

[12] L.R. Snyder, J.W. Dolan, J.R. Gant, J. Chromatogr. 165 (1979) 31.

[13] L.R. Snyder, J.W. Dolan, D.C. Lammen, J. Chromatogr. 485 (1989) 65.

[14] P. Nikitas, A. Pappa-Louisi, J. Chromatogr., A 1216 (2009) 1737.

[15] G. Vivó-Truyols, J.R. Torres-Lapasió, M.C. García-Alvarez-Coque, J. Chromatogr. A 1018 (2003) 169.

[16] B. Dejaegher, Y. Vander Heyden, J. Chromatogr., A 1158 (2007) 138.

[17] R. Ragonese, M. Mulholland, J. Kalman, J. Chromatogr., A 870 (2000) 45.

[18] E. Hund, Y. Vander Heyden, M. Haustein, D. L. Massart, J. Smeyers-Verbeke, Anal. Chim. Acta 404 (2000) 257.

[19] Guidelines of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), Pharmaceutical development, Q8(R2) Step 4, Geneva, 2009, pp. 1.

[20] E. Rozet, A. Ceccato, C. Hubert, E. Ziemons, R. Oprean, S. Rudaz, B. Boulanger, Ph. Hubert, J. Chromatogr. A 1158 (2007) 111.

[21] J.J. Peterson, J. Qual. Technol. 36 (2004) 139.

[22] J.J. Peterson, R.D. Snee, P.R. McAllister, T.L. Schofield, A.J. Carella, J. Qual. Technol. 41 (2009) 111.

[23] J.J. Peterson, J. Biopharm. Stat. 18 (2008) 959.

[24] G.W. Stockdale, A. Cheng, Qual. Technol. Quant. Manag. 6 (2009) 391.

[25] P. Lebrun, B. Govaerts, B. Debrus, A. Ceccato, G. Caliaro, Ph. Hubert, B. Boulanger, Chemom. Intell. Lab. Syst. 91 (2008) 4. http://hdl.handle.net/2268/1640

[26] B. De Backer, B. Debrus, P. Lebrun, L. Theunis, N. Dubois, L. Decock, A. Verstraete, Ph. Hubert, C. Charlier, J. Chromatogr., B 877 (2009) 4115. http://hdl.handle.net/2268/4442

[27] F. Krier, M. Brion, B. Debrus, P. Lebrun, A. Driesen, E. Ziemons, B. Evrard, Ph. Hubert, J. Pharm. Biomed. Anal. 54 (2011) 694. http://hdl.handle.net/2268/75222

[28] B. Debrus, P. Lebrun, A. Ceccato, G. Caliaro, E. Rozet, I. Nistor, R. Oprean, F.J. Rupérez, C. Barbas, B. Boulanger, Ph. Hubert, Anal. Chim. Acta *In press* (2011).

[29] B. Debrus, P. Lebrun, A. Ceccato, G. Caliaro, B. Govaerts, B.A. Olsen, E. Rozet, B. Boulanger, Ph. Hubert, Talanta 79 (2009) 77. http://hdl.handle.net/2268/13236

[30] A. Hyvärinen, J. Karhunen, E. Oja, Independent Component Analysis, Wiley, New York, 2001.

[31] W. Dewé, R.D. Marini, P. Chiap, Ph. Hubert, J. Crommen, B. Boulanger, Chemom. Intell. Lab. Syst. 74 (2004) 263. http://hdl.handle.net/2268/6036

[32] J.A. Nelder, R. Mead, Computer Journal 7 (1965) 308.

[33] The R Project for Statistical Computing, http://www.r-project.org/

[34] A. Hyvarinen, E. Oja, Neural Networks, 13 (2000) 411.

[35] N.R. Draper, H. Smith, Applied regression analysis, 3rd Ed., Wiley, New York, 1998.

[36] M.A. Herrador, A.G. Asuero, A.G. Gonzalez, Chemom. Intell. Lab. Syst. 79 (2005) 115.

Partie III.2

Application de la méthodologie

Section III.2.1

Application en Sciences médico-légales

Développement et validation d'une méthode CLHP pour la détermination qualitative et quantitative de cannabinoïdes

B. De Backer, B. Debrus, P. Lebrun, L. Theunis, N. Dubois, L. Decock, A. Verstraete, Ph. Hubert, C. Charlier, Innovative development and validation of an HPLC/DAD method for the qualitative and quantitative determination of cannabinoids in cannabis plant material, J. Chromatogr. B 877 (2009) 4115.

Résumé

La chromatographie gazeuse (CG) est généralement utilisée, par exemple en sciences médico-légale, pour l'analyse des échantillons de cannabis. Cependant, comme la CG nécessite l'échauffement des échantillons, les composés acides des cannabinoïdes sont décarboxylés pour donner les composés neutres respectifs. Inversement, la CLHP permet la détermination de la composition originelle des cannabinoïdes d'une plante par analyse directe. De nombreuses méthodes en CLHP ont été décrites dans la littérature. Néanmoins, la majorité d'entre elles ne permettent pas de séparer efficacement tous les cannabinoïdes ou n'ont pas été validées en accord avec les directives correspondantes. A l'aide d'une méthodologie innovante pour modéliser des réponses chromatographiques, une méthode CLHP-UV simple et précise a été développée pour la quantification des cannabinoïdes neutres et acides présents dans des plants de cannabis : le Δ9-tetrahydrocannabinol (THC), l'acide correspondant du THC (THCA), le cannabidiol (CBD), l'acide correspondant du CBD (CBDA), le cannabigerol (CBG), l'acide correspondant du CBG (CBGA) and le cannabinol (CBN). Le Δ8-Tetrahydrocannabinol (Δ8-THC) a été déterminé qualitativement. En se basant sur la théorie des plans d'expériences, des équations linéaires multiples ont été développées et utilisées pour identifier les conditions chromatographiques optimales. La méthode chromatographique a été validée selon les profils d'exactitude basés sur les intervalles de tolérance "β-expectation", l'estimation de l'erreur totale et l'évaluation de l'incertitude de mesure. Cette méthode analytique peut dès lors être utilisée pour diverses applications telles que la détermination du phénotype, l'évaluation de la teneur en composés psychoactifs et le contrôle de la qualité.

Summary

GC is commonly used for the analysis of cannabis samples, e.g. in forensic chemistry. However, as this method is based on heating of the sample, acidic forms of cannabinoids are decarboxylated into their neutral counterparts. Conversely, HPLC permits the determination of the original composition of plant cannabinoids by direct analysis. Several HPLC methods have been described in the literature, but most of them failed to separate efficiently all the cannabinoids or were not validated according to general guidelines. By use of an innovative methodology for modelling chromatographic responses, a simple and accurate HPLC-DAD method was developed for the quantification of major neutral and acidic cannabinoids present in cannabis plant material: Δ9-tetrahydrocannabinol (THC), THC acid (THCA), cannabidiol (CBD), CBD acid (CBDA), cannabigerol (CBG), CBG acid (CBGA) and cannabinol (CBN). Δ8-Tetrahydrocannabinol (Δ8-THC) was determined qualitatively. Following the practice of design of experiments, predictive multiple linear models were developed and used in order to find optimal chromatographic analytical conditions. The method was validated following an approach using accuracy profiles based on β-expectation tolerance intervals for the total error measurement, and assessing the measurements uncertainty. This analytical method can be used for diverse applications, e.g. plant phenotype determination, evaluation of psychoactive potency and control of material quality.

1. Introduction

Cannabis can be considered as the most controversial plant in our society: next to the important medical use, cannabis is also the most frequently consumed drug of abuse in Europe. It has been estimated that about four million European adults (∼1% of all 15- to 64-year-olds) are using cannabis each day or almost daily; and that around 23 million Europeans (∼7% of all 15- to 64-year-olds) have consumed cannabis at least one time during the past year [1]. The plant Cannabis sativa L. constitutes the basic material of all cannabis products. C. sativa L. belongs to the family of the Cannabinaceae. The current systematic classification of cannabis is listed in Table 1 [2,3].

Table 1 Current systematic classification of Cannabis sativa L. [2,3].

Division	Angiosperms
Class	Dicotyledon
Subclass	Archichlamydeae
Order	Urticales
Family	Cannabinaceae
Genus	*Cannabis*
Species	*sativa* L.

1.1. Cannabinoids

The chemistry of cannabis has been studied extensively: approximately 500 compounds have been identified. The most interesting among these constituents are the cannabinoids; terpenophenolic compounds unique to cannabis and concentrated in a resinous secretion produced by the trichomes of the plant. These trichomes are particularly concentrated at specific parts of the female inflorescence [2].

The cannabinoids form a group of related compounds of which about 70 are known [2,4]. Of the major cannabinoids in C. sativa L., ∆9-tetrahydrocannabinol (THC) is generally accepted to be the compound that possesses the psychoactive properties [5,6]. Inplant tissues, cannabinoids are biosynthesized in an acidic (carboxylated) form. The most common types of acidic cannabinoids found are ∆9-tetrahydrocannabinolic acid A (THCA-A), cannabidiolic acid (CBDA) and cannabigerolic acid (CBGA). THC acid exists under two forms: THCA-A and THCA-B. However, only traces of THCA-B can be detected in cannabis samples [3], THCA-A is the major form and will be further referred to as THCA. CBGA is the direct precursor of THCA, CBDA and cannabichromenic acid (CBCA) (Fig. 1). The carboxyl group is not very stable and is easily lost as CO_2 under influence of heat or light, resulting in the corresponding neutral cannabinoids: THC, cannabidiol (CBD) and cannabigerol (CBG) [2,7]. These are formed during heating and drying of harvested plant material, or during storage and when cannabis is smoked [6,8,9].

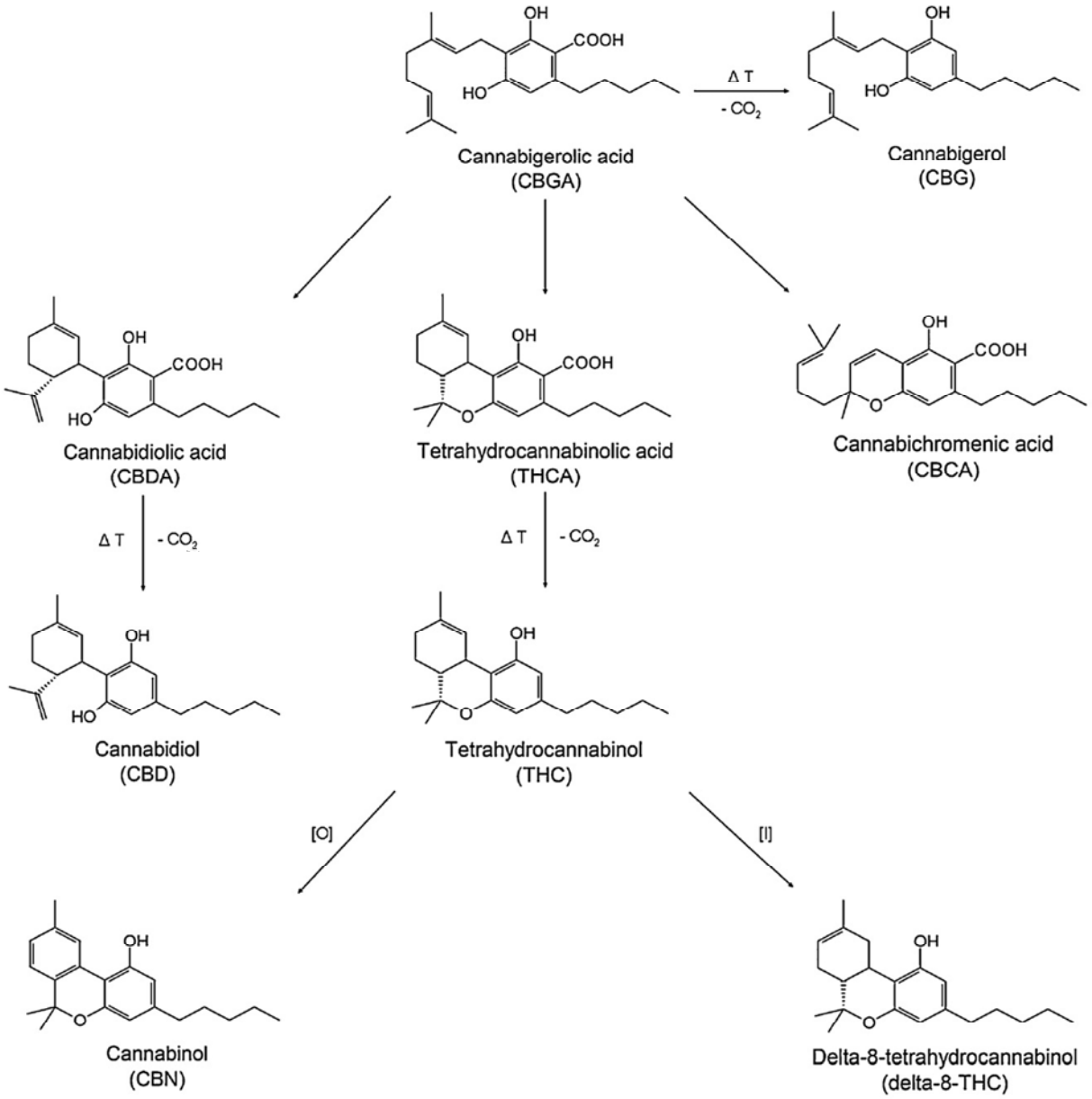


Fig. 1 Biosynthetic pathway for the production of cannabinoids and main breakdown products of THC (ΔT = heating, [O] = oxidation, [I] = isomerization).

The variable conditions during all stages of growing, harvesting, processing, storage and use also induce the presence of breakdown products of cannabinoids. The most commonly found degradation product in aged cannabis is cannabinol (CBN), produced by oxidative degradation of THC under the influence of heat and light [2,10]. THC can also be transformed by isomerization to Δ8-THC, which is an artefact. In order to quantify the "total THC content" once present in the fresh plant material, the concentrations of degradation products have to be added to THCA and THC contents.

1.2. Phenotypes

Hillig and Mahlberg [8] identified three chemotypes (chemical phenotypes) of cannabis: drug-type plants (chemotype I) show a high [total THC/total CBD] ratio (>> 1.0), intermediate type plants (chemotype II) have an intermediate ratio (close to 1.0), and fibretype plants (chemotype III) exhibit a low [total THC/total CBD] ratio (<< 1.0). For forensic and legal purposes, the most important classification of cannabis types is that into the drug-type and the fibre-type. The latter, usually called “hemp”, refers to varieties that have low THC concentrations but generally contain other non-psychoactive cannabinoids as major compounds, like CBD or CBG. In many countries, hemp cultivation is prohibited by legislation because of the presence of the psychoactive compounds. In countries where hemp cultivation is allowed, the cultivars are tested in order to verify that the psychoactive potency is below a minimum acceptable level [5]. In Europe, the maximum THC content allowed for the cultivation of hemp is either 0.2% or 0.3% of the weight of dry matter, in function of the country.

1.3. Analytical methods

The analysis of the original composition of plant material is necessary for diverse purposes as phenotype determination and quality control of medicinal cannabis used in therapeutic treatment. In addition, it has been repeatedly suggested that the effects of THC or other single cannabinoids are not equal to that of whole cannabis preparations [11,12]: some of the bio-activity observed for these preparations could be due to acidic cannabinoids [13]. That way, a method allowing the qualitative and quantitative determination of neutral as well as acidic cannabinoids in plant material must be available [2].

Gas Chromatography (GC) is the most commonly used method for the analysis of cannabis products [5,8,14–18], but it does not permit the determination of acidic cannabinoids due to decarboxylation into their neutral forms during analysis. Furthermore, this thermal conversion of acidic cannabinoids seems to be incomplete [19]. In order to quantify neutral cannabinoids by GC, a time-consuming derivatization step is mandatory. On the contrary, High Performance Liquid Chromatography (HPLC) allows the determination of the neutral forms since no heating occurs during separation. Use of HPLC is thereby the simplest method for the determination of the original composition in cannabinoids of plant material. Raharjo and Verpoorte [15] reviewed different HPLC methods for the analysis of cannabinoids. However, most of them were not validated according to the new guidelines using the total error approach, or were not able to separate efficiently all the major cannabinoids [2,3,20,21]. Because of the complex composition of plant material, the analysis of major cannabinoids is not easily achieved and overlap of peaks occurs (between CBD/CBG and CBN/CBGA) [2,20]. Consequently, Hazekamp et al. [2,20] had to combine HPLC with a secondary analysis by GC in order to identify and quantify all major cannabinoids. The use of mass spectrometry coupled to HPLC may be a solution in order to resolve all peaks in a single

analytical run [10,20]. However, this method is expensive and not routinely available to most laboratories.

The goal of the present study was therefore to develop and to validate a simple HPLC-DAD method, allowing a good separation followed by a qualitative and quantitative determination of major neutral and acidic cannabinoids present in plant material. Determination was performed on cannabinoids of potential interest for the medicinal research community and cannabinoids used for the classification of cannabis phenotypes and for monitoring of the psychotropic potency: THC, THCA, CBD, CBDA, CBG, CBGA, CBN and ∆8-THC. The method was validated within broad ranges of concentrations adapted to the levels found in the three cannabis plant phenotypes.

2. Materials and methods

2.1. Chemicals and reagents

Cannabinoid reference standards for THC, CBD, CBN and ∆8-THC were purchased from LGC Standards (Molsheim, France). Reference standards for THCA, CBDA, CBGA and CBG were purchased from Echo Pharmaceuticals BV (Weesp, The Netherlands). All standards had a purity of ≥98%. Prazepam was purchased from Certa (Braine-l'Alleud, Belgium). For extractions, HPLC grade methanol and chloroform were purchased from LabScan (Dublin, Ireland). For the mobile phase, HPLC quality methanol was purchased from Biosolve (Valkenswaard, The Netherlands); ultrapure distilled water and deionized water were prepared in-house and filtered prior to use; ammonium formate and formic acid were purchased from Sigma (Bornem, Belgium). All reagents were at least of analytical grade.

2.2. Cannabis samples

Eight samples of drug-type cannabis and one sample of non-psychotropic cannabis were provided by police (confiscated samples). Two other samples of fibre-type cannabis were generously provided by the laboratory of Ecophysiology and Plant Breeding of the Université catholique de Louvain.

2.3. Sample preparation

Plant material samples were dried for 24 h in a 35 °C forced ventilation oven. Crumbly samples were then grinded and mixed. 200mg of this fine powder were weighed in a flask and extracted with 20 mL of a mixture methanol/chloroform (v/v: 9/1) by agitation during 30 min. The extract was filtered and appropriately diluted in a small test tube. A 100 µL aliquot of the dilution was evaporated under a gentle stream of nitrogen and redissolved in 100µL of a mixture of water/methanol (v/v: 5/5). Prazepam (100 mg/L) was used as internal standard.

2.4. HPLC equipment and chromatographic conditions

All chromatographic runs were carried out using a Hewlett-Packard (HP) HPLC System (Agilent Technologies, Böblingen, Germany), consisting of a G1311A quaternary solvent pump (1200 series), a G1322A solvent degasser (1200 series), a G1313A autosampler (1100 series) and a G1316A column compartment (1100 series). A Waters (Zellik, Belgium) 2996 photodiode-array detector (DAD) was used for detection. Full spectra were recorded in the range 200–400 nm. Chromatographic separations were achieved using a Waters XTerra® MS C18 analytical column (5µm, 250mm×2.1 mm i.d.), protected by a Waters XTerra® MS C18 guard column (5µm, 10mm×2.1 mm i.d.). Equipment control, data acquisition and integration were performed with Empower Pro 2.0 software.

The mobile phase consisted of a mixture of methanol/water containing 50 mM of ammonium formate (adjusted to pH 5.19). Initial setting was 68% methanol (v/v), which was linearly increased to 90.5% methanol over 25 min, then increased to 95% in 1 min. After maintaining this condition for 3 min, the column was set to initial condition in 1min and re-equilibrated under this condition for 6 min. The total runtime was 36 min. Flow-rate was set to 0.3 mL/min, the injection volume was 30 µL. All experiments were carried out at 30 °C.

2.5. Method validation

In accordance to ISO 17025 and the guidelines of the French Society of Pharmaceutical Sciences and Techniques (SFSTP), the present method was fully validated using total error approach [22–25]. The e.noval software V2.0 (Arlenda, Liège, Belgium) was used to compute all validation results and build the accuracy profiles.

3. Results

3.1. Method optimization

3.1.1. Experimental design

Three HPLC factors have been investigated: the percentage of methanol at the beginning of the gradient (pcI), the pH of the aqueous part of the mobile phase (pH) and the gradient time to reach 95% of methanol (T_G). Table 2 shows the levels of these three factors. Design of experiments (DoE) methodology has been used and a full factorial design was selected, which is convenient to explore the space of factors. As such, a total of 45 experimental conditions were defined and a chromatogram was recorded at each of these. At the central point (pcI = 40%, pH 6.3 and T_G = 20 min), two independent repetitions (preparation of new buffer) were carried out to estimate the reproducibility of the system.

Table 2 Description of the three factors involved in the experimental design.

Factors	Levels				
pcI (%)	5	40	75		
pH	2.6	4.45	6.3	8.15	10.0
T_G (min)	10	20	30		

3.1.2. Statistical models

In the resulting chromatograms, the peaks were detected and indexed at their beginnings, apexes and ends. The retention factors (log(*k*)) have been used to create a multivariate responses surface model. Fig. 2 illustrates the fit of the observed retention times versus the predicted retention times using the statistical models. Residuals are mainly located into the interval [−2, 2] min. As the adjusted R^2 of each model were higher than 0.95, the overall quality of the fit is good although some outliers were observed.

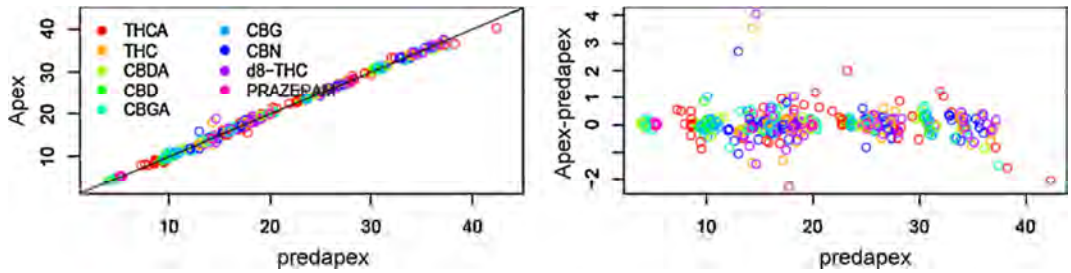


Fig. 2 Actual original responses (Apex) versus predicted one (predapex). (right) The residuals.

3.1.3. Optimization — Design Space

The minimal separation (separation of the critical pair) is optimized using the methodology presented by Lebrun et al. [26]. The separation is defined as the difference between the beginning of a peak and the end of the preceding peak. Consequently to the

response (retention times) modelling, the experimental domain is investigated to encounter an optimal separation. The propagation of the predictive error through the criterion (the separation, S) was analyzed to give confidence in this optimum. The Design Space (DS) is defined as the set of factor conditions that are likely to provide satisfactory results in the future use of the analytical method (e.g. routine). Mathematically, the DS applied in this case is defined as in Eq. (1),

$$\mathrm{DS} = \left\{ x_0 \in \chi | E_{\hat{\theta}} \left[\mathrm{P}(S > \lambda) | \hat{\theta} \right] \geq \pi \right\} \quad (1)$$

where x_0 is the set of factor conditions belonging to the experimental domain χ, for which the expected probability to have a separation (S) higher than λ is higher than π, given the uncertainty of the estimation of the parameters θ of the model. A separation of at least 0 min (λ = 0 min) should be obtained. Monte Carlo simulations are performed to propagate uncertainty from parameters to responses and criterion. A summary of the optimal values of factors (the best probabilities to achieve a minimal separation of at least 0 min; baseline resolved peaks) is shown in Table 3. Fig. 3 shows the probability surfaces in different directions of the space of factors, around the optimal solution and with, for each graph, two factors that are fixed at optimal values.

The chromatograms predicted at the conditions described in Table 3 can be seen in Fig. 4. Despite the poor DS probability (40%), a good agreement between the predicted chromatogram and the real processed chromatogram is observed. Separation of all of the compounds is well achieved within the DS. A chromatogram experimentally obtained with cannabinoid standards is shown in Fig. 5. Qualitative HPLC profiles of herbal cannabis and cannabis resin samples are provided in Figs. 6 and 7, respectively.

Table 3 Optimal factor setting maximizing the separation of the compounds.

Optima	pcI (%)	pH	T_G (min)
P(S > 0)>40%	68	5.2	30

3.2. Method Validation

3.2.1. Selectivity

The selectivity of detection of each compound was ensured by the determination of the retention times and the recording of the complete UV spectra of the cannabinoids. Spectra are shown in Fig. 8.

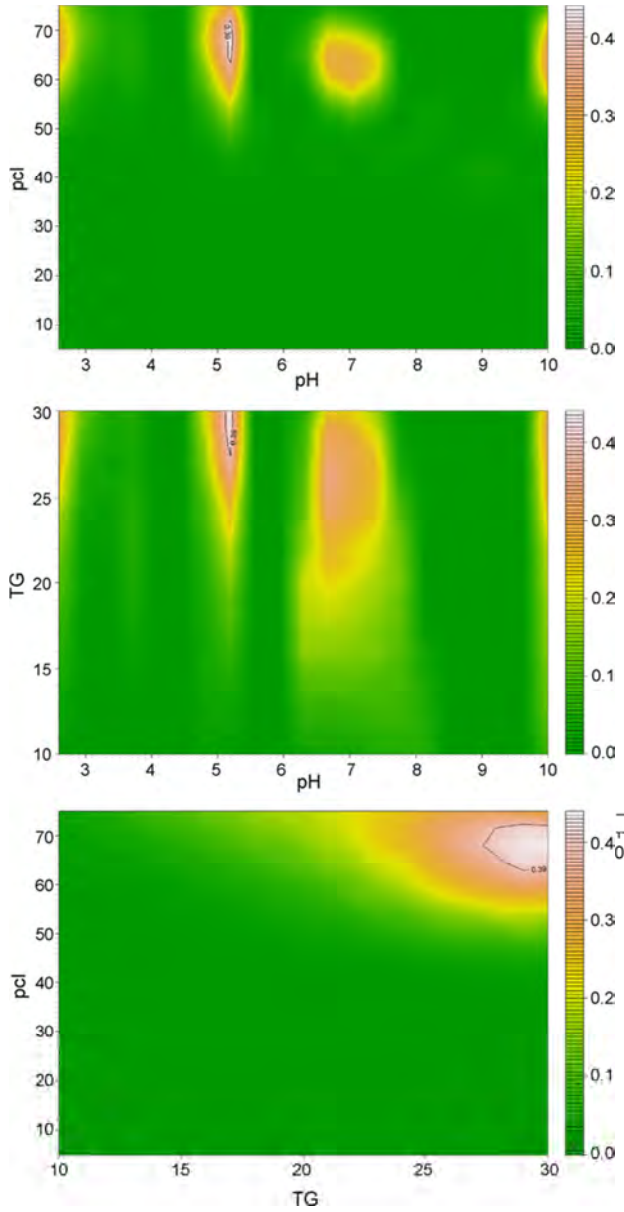


Fig. 3 Representation of the Design Space of the method on the experimental domain. Inside black lines, the expected probability to have well-separated peaks is at least 0.39.

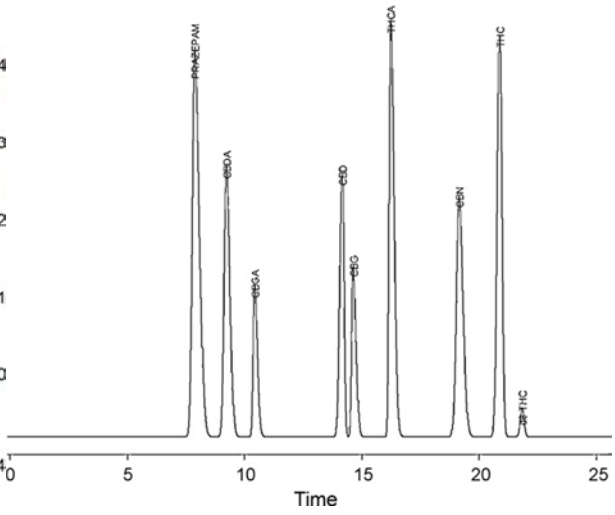


Fig. 4 Predicted chromatogram at optimal solution.

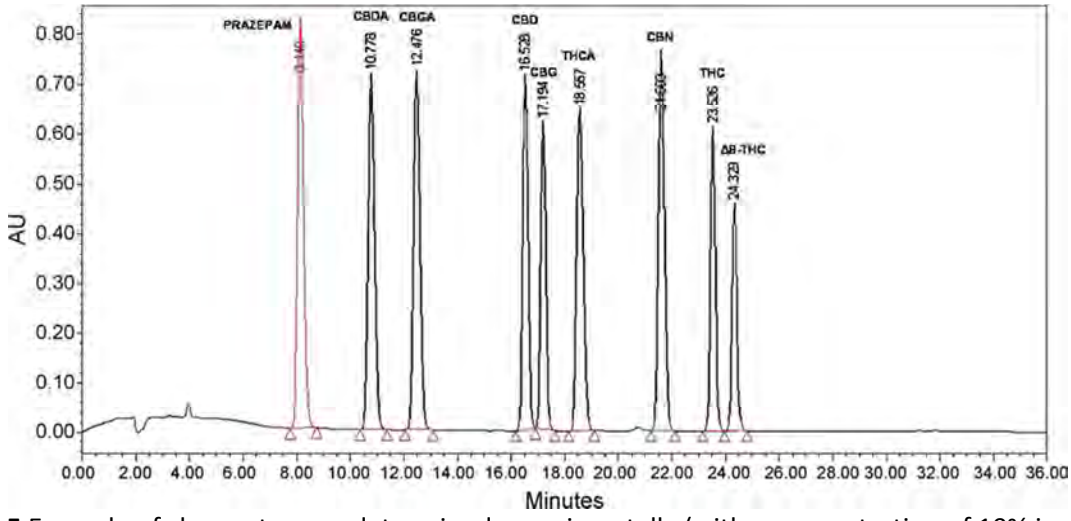


Fig. 5 Example of chromatogram determined experimentally (with a concentration of 10% in each cannabinoid) and retention times of the compounds (prazepam is used as internal standard).

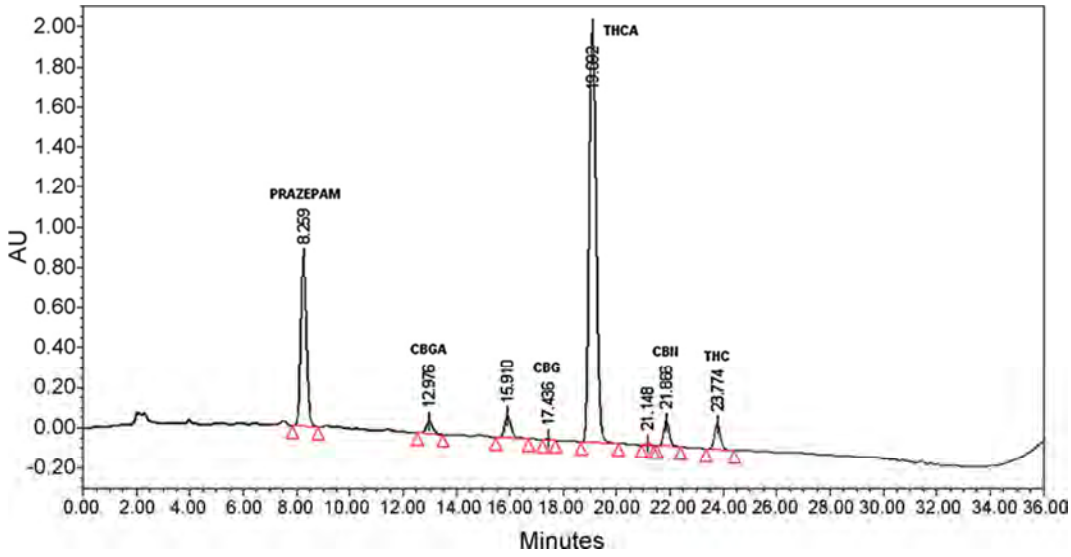


Fig. 6 Qualitative HPLC profile of herbal cannabis sample.

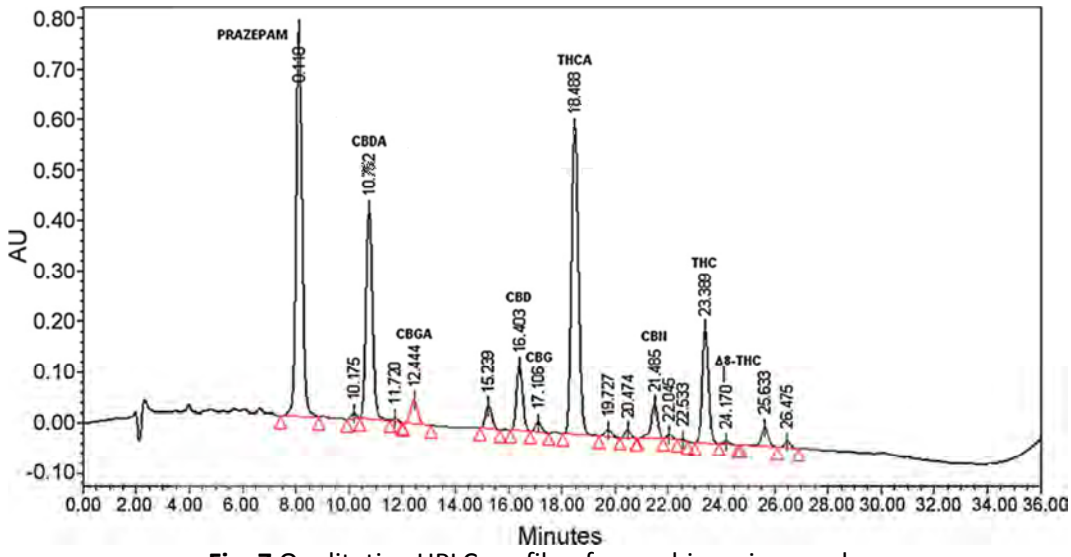


Fig. 7 Qualitative HPLC profile of cannabis resin sample.

3.2.2. Linearity

The response function is, within a certain range, the relationship between the response observed and the concentration of the analyte in the sample [27]. Calibration curves were obtained from standard solutions in methanol containing eight different concentrations for each cannabinoid from 0.15 to 20% (percentage of weight of dry plant material), corresponding to 0.375 to 50 µg/mL. The concentration levels were chosen in order to cover the different contents in cannabinoids in plant materials of diverse types. Each calibration point was analyzed in duplicate on two consecutive days. Calibration curves were calculated using unweighted linear regression analysis and linearity was expressed by the R^2-value. The calibration parameters were stable with regression coefficients always >0.99 for each cannabinoid studied. The regression coefficients for each analyte are listed in Table 4. The curves were linear in the concentration range studied for each analyte.

3.2.3. Limit of quantification (LOQ), limit of detection (LOD)

The LOQ were experimentally determined by analyzing standard solutions at 0.025%, 0.05%, 0.075% and 0.1% (corresponding to 0.0625, 0.125, 0.1875 and 0.25 µg/mL). The lower LOQ was determined as the concentration which provided measurements with an accuracy within the acceptance limits (±20%) from their nominal values. The LOD was determined as the smallest dilution that gave a good correlation between the compound UV–vis spectrum and the spectra library. LOQ and LOD for each cannabinoid are listed in Table 4.

Table 4 Linearity (expressed by R^2), limits of quantification (LOQ), limits of detection (LOD).

Compounds	R^2	LOQ (%)	LOD (%)
THCA	0.9969	0.05	0.025
THC	0.9940	0.05	0.025
CBDA	0.9939	0.05	0.05
CBD	0.9951	0.075	0.075
CBGA	0.9948	0.05	0.05
CBG	0.9959	0.15	0.1
CBN	0.9917	0.05	0.025

3.2.4. Trueness, precision and accuracy

A statistical approach based on the total error measurements including both bias and standard deviation was applied to validate the method.

Validation standards: because cannabis without major cannabinoids was not available, the validation standards were prepared by spiking samples of nettle (Urtica dioica, which belongs to the same order as C. sativa L.) with an extract of cannabis resin. This hashish extract contained all the cannabinoids of interest in significant amounts, except Δ8-THC. The method was therefore not validated for the quantification of Δ8-THC. The extract of hashish was quantified and used to prepare three validation standards. The volumes added

correspond to different concentrations for each cannabinoid, in function of the quantity initially present in the hashish sample.

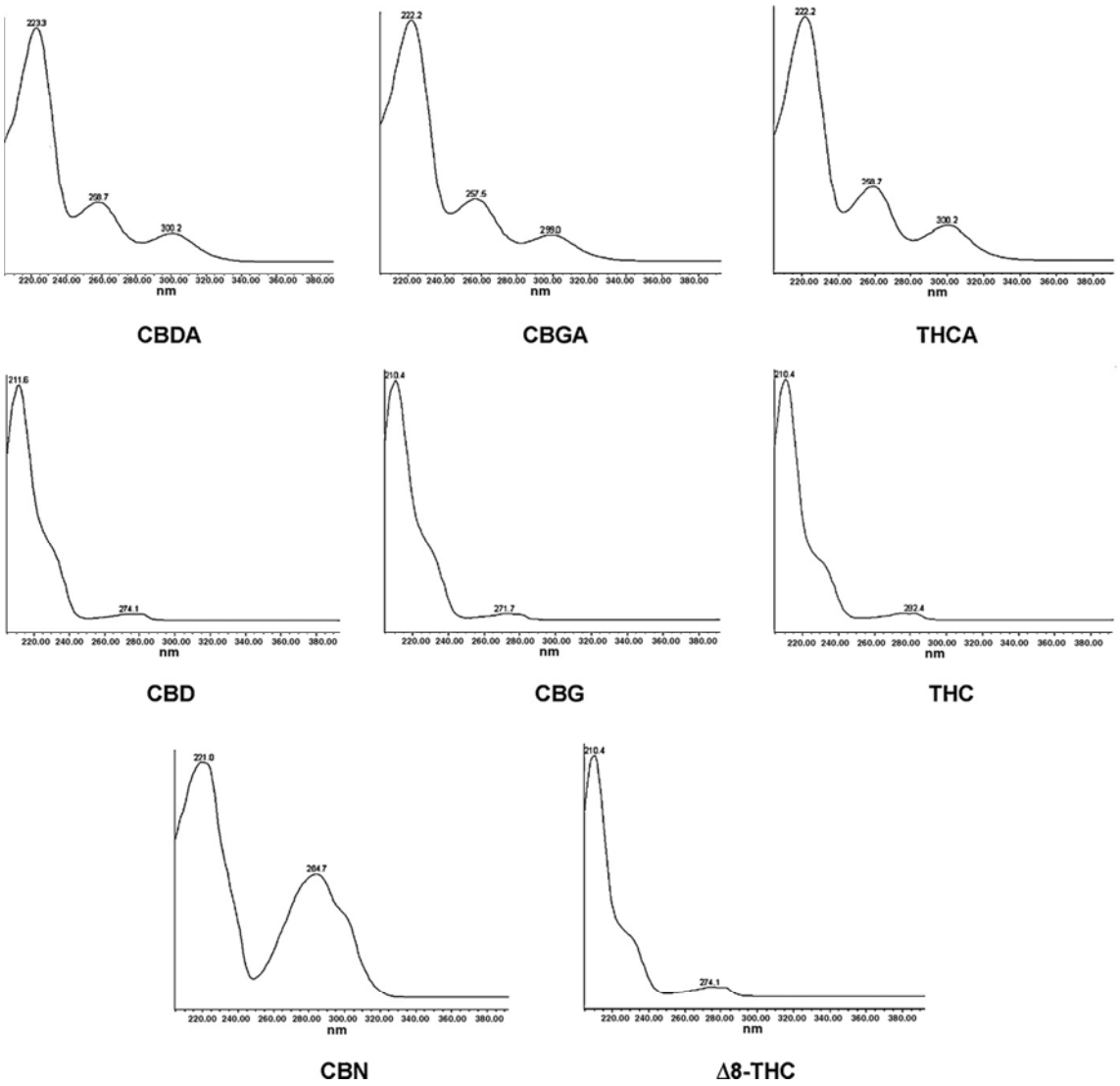


Fig. 8 UV spectra of studied cannabinoids.

The concentrations spiked for each cannabinoid are listed in Table 5. Each validation standard was analyzed in triplicate on three consecutive days. The concentrations of the validation standards were back-calculated from the obtained results to determine the mean relative bias, the standard deviation for intermediate precision and finally the upper and lower β-expectation tolerance limits at the 17.5% level.

Trueness and precision give information on respectively systematic and random errors. Trueness refers to the closeness of agreement between the exact concentration in spiked material and the obtained main results. Trueness is expressed in terms of relative bias (%) and was calculated from the validation standards for each compound [24,28]. Trueness was acceptable for all cannabinoids, since the relative bias (%) were always smaller than 10%. Results are presented in Table 6.

Table 5 Mean introduced concentrations (%).

Compounds	Concentration level 1 (%)	Concentration level 2 (%)	Concentration level 3 (%)
THCA	1.83	3.06	6.12
THC	1.15	1.92	3.84
CBDA	1.00	1.67	3.33
CBD	0.549	0.917	1.834
CBGA	0.134	0.217	0.434
CBG	0.092	0.154	0.308
CBN	0.158	0.264	0.528

The precision of the method was determined by computing the Relative Standard Deviations (RSDs) for repeatability and time different intermediate precision at each concentration level of the validation standards [24,25,28], and did not exceed 11% for all of the cannabinoids (Table 6).

Table 6 Method validation for seven cannabinoids in plant material. Trueness, precision, accuracy and uncertainty.

	Level	THCA	THC	CBDA	CBD	CBGA	CBG	CBN
Trueness								
Relative Bias (%)	1	-0.9	5.3	-2.3	6.6	-10.6	1.4	0.1
	2	3.1	5.1	3.9	9.5	-1.1	1.8	8.9
	3	4.0	1.6	6.1	11.1	3.4	9.1	8.8
Intra-assay precision								
Repetability (RSD%)	1	1.03	1.31	2.15	5.79	3.04	2.39	3.82
	2	1.01	2.21	2.30	4.71	3.23	3.35	3.98
	3	2.39	1.94	4.98	6.28	3.00	1.93	6.09
Between-assay precision								
Intermediate precision (RSD%)	1	3.97	3.80	3.89	7.52	3.64	5.34	4.46
	2	3.14	4.32	2.30	6.01	3.79	7.78	6.84
	3	3.26	2.12	5.69	6.28	10.94	1.93	6.09
Accuracy								
β-expectation tolerance limits	1	[1.64,1.99]	[1.11,1.32]	[0.90,1.05]	[0.51,0.66]	[0.11,0.13]	[0.08,0.10]	[0.14,0.17]
	2	[2.93,3.38]	[1.82,2.22]	[1.68,1.80]	[0.91,1.10]	[0.20,0.23]	[0.13,0.19]	[0.25,0.33]
	3	[5.91,6.82]	[3.66,4.15]	[3.22,3.85]	[1.86,2.22]	[0.34,0.56]	[0.32,0.35]	[0.50,0.64]
Uncertainty								
Relative expanded uncertainty (%)	1	9.1	8.7	8.8	16.5	7.9	12.1	9.6
	2	7.2	9.8	4.8	13.2	8.2	17.7	15.4
	3	7.2	4.6	12.3	13.2	25.1	4.1	12.8

Accuracy takes into account the total error (sum of the systematic and random errors) of the test results [24,25,28]. It refers to the closeness of agreement between the test results and the acceptance reference value. The acceptance limits were set at ±30% as recommended [22]. As shown in Fig. 9, the relative upper and lower β-expectation tolerance intervals (%) did not exceed the acceptance limits (±30%) for each cannabinoid and each concentration level. The β-expectation tolerance limits are listed in Table 6. The approach used guarantees that each further measurement of unknown samples will be included within the tolerance limits at the 17.5% level.

3.2.5. Recovery

The absolute recoveries of THCA, THC, CBDA, CBD, CBGA, CBG and CBN were determined at the three different concentrations listed in Table 5 [22,29]. The mean recoveries are shown in Table 7. Those absolute recoveries were calculated by comparing peak areas of each cannabinoid obtained from freshly prepared matrix samples treated according to the described procedure with those found after the direct injection on the analytical column of standard solutions at the same concentrations. All the recoveries were good demonstrating the high extraction efficiency of the method.

3.2.6. Uncertainty assessment

The uncertainty characterizes the dispersion of the values that could reasonably be attributed to the measurand. Several uncertainty results were generated. The expanded uncertainty represents an interval around the results where the unknown true value can be observed with a confidence level of 95%. The relative expanded uncertainties (%) are obtained by dividing the corresponding expanded uncertainties with the corresponding introduced concentrations. Values for each cannabinoid are presented in Table 6 and were between 4.1 and 25.1%.

3.3. Analysis

The present method was applied for the analysis of different cannabis products. The preparation of these samples was the same as described above. Samples 1–8 were police confiscates of drug-type cannabis in which THCA and THC are the main cannabinoids. Samples 9 and 10 are fibre-type cannabis, respectively Fedora 17 and Santhica 27 varieties. Fedora 17 is a “classical” fibre-type variety containing CBDA and CBD as major cannabinoids. Santhica 27 is a new variety of hemp in which the biogenesis of cannabinoids seems to have stopped precociously: CBGA and CBG are the main cannabinoids [17]. Sample 11 is a non-psychotropic cannabis which grew wild (seeds coming from bird food). Cannabinoid concentrations of these samples of herbal cannabis are listed in Table 8.

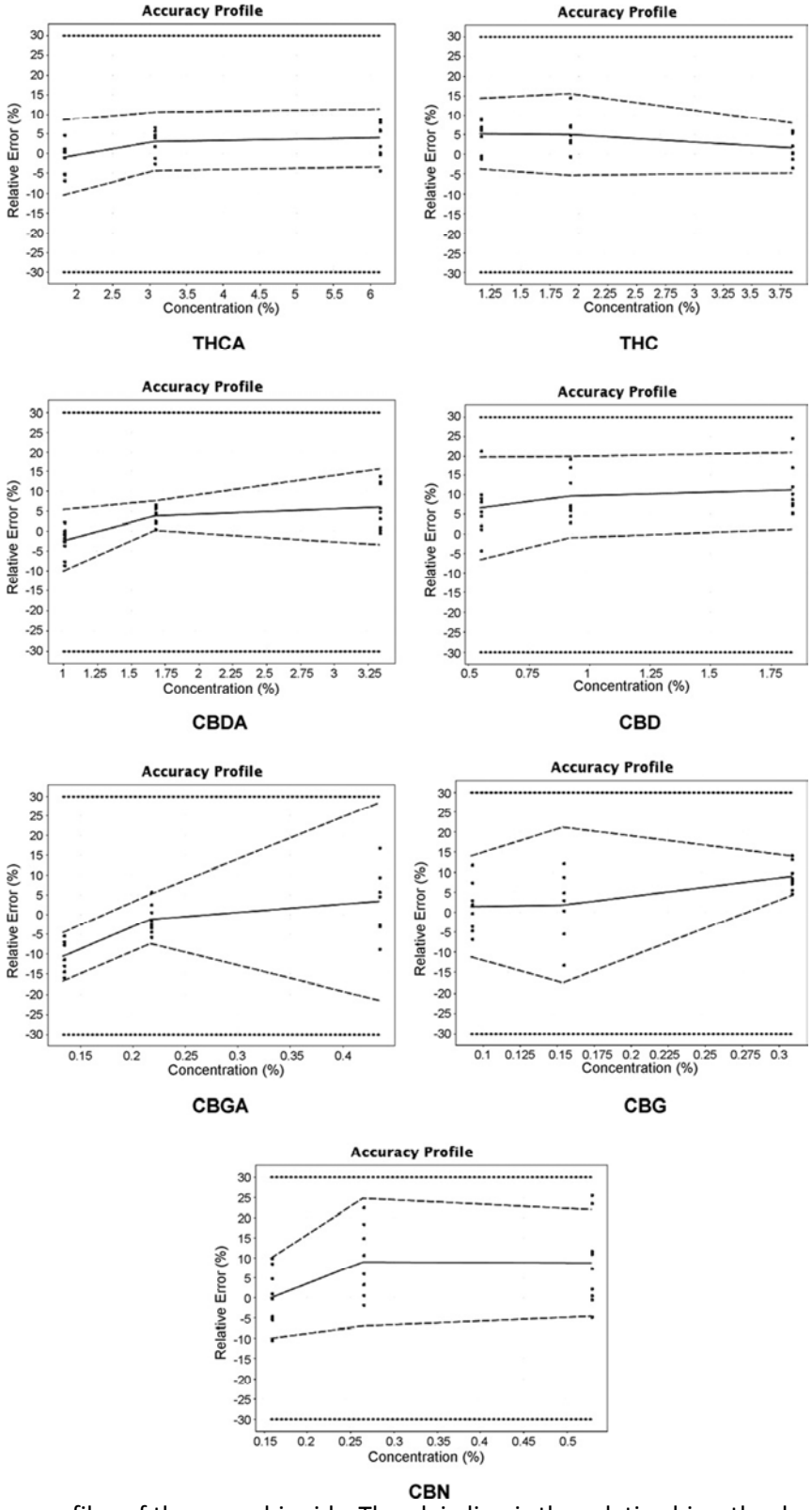


Fig. 9 Accuracy profiles of the cannabinoids. The plain line is the relative bias, the dashed lines are the β-expectation tolerance limits and the dotted lines represent the acceptance limits (30%). The dots represent the relative back-calculated concentrations and are plotted with respect to their targeted concentration.

Table 7 Mean recoveries of cannabinoids.

Compounds	Number of repetitions (n)	Recovery ± SD (%)
THCA	3	102.1 ± 2.6
THC	3	104.0 ± 2.1
CBDA	3	102.6 ± 4.3
CBD	3	109.1 ± 2.3
CBGA	3	97.2 ± 7.2
CBG	3	104.1 ± 4.3
CBN	3	105.9 ± 5.1

Table 8 Cannabinoid concentrations found in different types of herbal cannabis products.

Sample	Type	THCA (%)	THC (%)	CBN (%)	Total THC (%)	CBDA (%)	CBD (%)	CBGA (%)	CBG (%)
1	Drug	20.24	1.55	<LOD	21.79	<LOD	<LOD	0.72	<LOQ
2	Drug	5.40	3.23	0.06	8.69	<LOD	<LOD	0.12	<LOQ
3	Drug	22.92	2.58	<LOD	25.51	<LOD	<LOD	1.89	0.28
4	Drug	15.68	1.56	<LOD	17.24	<LOD	<LOD	0.30	<LOQ
5	Drug	15.81	1.21	<LOD	17.02	<LOD	<LOD	0.43	<LOD
6	Drug	15.33	1.29	<LOD	16.82	<LOD	<LOD	0.41	<LOD
7	Drug	8.18	4.05	0.33	12.56	<LOD	<LOD	0.17	<LOQ
8	Drug	10.79	3.20	0.08	14.08	<LOQ	<LOD	0.38	<LOQ
9	Fibre	0.09	<LOQ	<LOD	0.09	2.37	0.10	0.10	<LOQ
10	Fibre	<LOQ	<LOD	<LOD	<LOQ	<LOD	<LOD	2.58	0.16
11	Fibre	0.57	0.10	<LOD	0.67	<LOD	<LOD	<LOQ	<LOD

4. Discussion

Gas Chromatography (GC) is the most commonly used method for the analysis of cannabis products, e.g. in forensic chemistry [5,8,14–18]. However, as this method is based on heating the sample, thermal conversion occurs and the acidic forms of cannabinoids are converted into their decarboxylated counterparts. In order to determine neutral cannabinoids, a time-consuming derivatization step is mandatory. Conversely, High Performance Liquid Chromatography (HPLC) permits the determination of the original composition of the cannabinoids in the plant by direct analysis. In contrast to GC, no decomposition of the cannabinoids occurs during analysis by HPLC. Furthermore, THCA decarboxylation during GC analysis is often supposed to be complete [5,14]; but Dussy et al. [19] demonstrated in 2005 that this conversion is only partial. Various analytical conditions were studied and a maximal conversion of about 67% was obtained at an injector temperature of 220 °C. Laboratories quantifying total THC by HPLC, building the sum of THCA and the already present THC in the plant, get therefore a higher value than those who quantify THC by GC [19].

Several HPLC methods have been described in the literature [2,3,20,21], reviewed in 2004 by Raharjo and Verpoorte [15], but most of them failed to separate efficiently all the cannabinoids or were not validated according to the new guidelines using total error approach. Some methods were not validated for acidic cannabinoids as these were, until recently, not commercially available.

Hazekamp et al. [2,20] described two methods (acidic or basic eluent) who did not permit a full separation of peaks for either CBGA/CBN or CBD/CBG. A secondary analysis by GC was necessary to quantify those cannabinoids, causing a waste of time. Another solution is the coupling of the HPLC system with a mass spectrometer. However, mass spectrometry is not routinely available to most laboratories.

The selectivity of the compounds can be modified by adjusting the pH of the eluent. The relative retention times of the acidic cannabinoids are influenced by changing the pH, while the order of elution and the relative retention times for the neutral cannabinoids remain the same [2,20]. By adjusting the pH of our eluent (pH 5.19) and the gradient elution slope, thanks to the optimization method, we were able to modify precisely the relative retention times of the compounds in order to fully separate each of them (however, CBD and CBG may yet slightly overlap if present in high concentrations, >10%). Consequently, the method developed and validated allows a good separation of eight major cannabinoids of interest in a single run of 25 min (36 min with re-equilibration).

5. Conclusions

Using original tools, a simple and accurate HPLC method for the quantification of major cannabinoids in cannabis plant material has been developed and validated. This analytical method can be used for diverse applications, e.g. plant phenotype determination, evaluation of psychoactive potency and control of medicinal sample quality. It could also be an aid for checking the identity of cannabis specimen of different origin, next to other techniques as determination of microelements or stable isotopes of carbon and nitrogen [30]. In addition, quantification of total CBG can be useful for the identification of different types of fibre hemp analyzed [17].

Acknowledgements

The authors acknowledge the Belgian Science Policy (Belspo) for the financial support (GEOCAN project DR-00-48) and are very grateful to the reviewers for providing important comments that led to significant improvements of this article.

References

[1] European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), Annual Report 2008: The State of the Drug Problem in Europe, Office for Official Publications of the European Communities, Luxembourg, 2008.

[2] A. Hazekamp, Cannabis; extracting the medicine, PhD thesis, Universiteit Leiden, The Netherlands, 2007.

[3] T. Lehmann, R. Brenneisen, J. Liq. Chromatogr. 18 (1995) 689.

[4] M.A. Elsohly, D. Slade, Life Sci. 78 (2005) 539.

[5] A.K. Hewavitharana, G. Golding, G. Tempany, G. King, N. Holling, J. Anal. Toxicol. 29 (2005) 258.

[6] C. Giroud, Chimia 56 (2002) 80.

[7] G.A. Thakur, R.I. Duclos Jr., A. Makriyannis, Life Sci. 78 (2005) 454.

[8] K.W. Hillig, P.G. Mahlberg, Am. J. Bot. 91 (2004) 966.

[9] N.J. Doorenbos, P.S. Fetterman, M.W. Quimby, C.E. Turner, Ann. N. Y. Acad. Sci. 191 (1971) 3.

[10] A.A. Stolker, J. van Schoonhoven, A.J. de Vries, I. Bobeldijk-Pastorova, W.H. Vaes, R. van den Berg, J. Chromatogr. A 1058 (2004) 143.

[11] E.B. Russo, J.M. McPartland, Psychopharmacology (Berl.) 165 (2003) 431.

[12] E. Russo, G.W. Guy, Med. Hypotheses 66 (2006) 234.

[13] K.C. Verhoeckx, H.A. Korthout, A.P. van Meeteren-Kreikamp, K.A. Ehlert, M. Wang, J. van der Greef, R.J. Rodenburg, R.F. Witkamp, Int. Immunopharmacol. 6 (2006) 656.

[14] H. Stambouli, A. El Bouri, M.A. Bellimam, T. Bouayoun, N. El Karni, Ann. Toxicol. Anal. 17 (2005) 79.

[15] T.J. Raharjo, R. Verpoorte, Phytochem. Anal. 15 (2004) 79.

[16] D.W. Lachenmeier, L. Kroener, F. Musshoff, Anal. Bioanal. Chem. 378 (2004) 183.

[17] G. Fournier, O. Beherec, S. Bertucelli, Ann. Toxicol. Anal. 16 (2004) 128.

[18] N. Fucci, Forensic Sci. Int. 138 (2003) 91.

[19] F.E. Dussy, C. Hamberg, M. Luginbühl, T. Schwerzmann, T.A. Briellmann, Forensic Sci. Int. 149 (2005) 3.

[20] A. Hazekamp, C. Giroud, A. Peltenburg, R. Verpoorte, J. Liq. Chromatogr. Rel. Technol. 28 (2005) 2361.

[21] V. Gambaro, L. Dell'Acqua, F. Farè, R. Froldi, E. Saligari, G. Tassoni, Anal. Chim. Acta 468 (2002) 245.

[22] C.T. Viswanathan, S. Bansal, B. Booth, A.J. DeStefano, M.J. Rose, J. Sailstad, V.P. Shah, J.P. Skelly, P.G. Swann, R. Weiner, AAPS J. 9 (2007) E30.

[23] B. Boulanger, P. Chiap, W. Dewé, J. Crommen, P. Hubert, J. Pharm. Biomed. Anal. 32 (2003) 753.

[24] P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, J. Pharm. Biomed. Anal. 36 (2004) 579.

[25] P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, J. Pharm. Biomed. Anal. 45 (2007) 82.

[26] P. Lebrun, B. Govaerts, B. Debrus, A. Ceccato, G. Caliaro, P. Hubert, B. Boulanger, Chemometr. Intell. Lab. Syst. 91 (2008) 4.

[27] E. Rozet, C. Hubert, A. Ceccato, W. Dewé, E. Ziemons, F. Moonen, K. Michail, R. Wintersteiger, B. Streel, B. Boulanger, P. Hubert, J. Chromatogr. A 1158 (2007) 126.

[28] E. Rozet, A. Ceccato, C. Hubert, E. Ziemons, R. Oprean, S. Rudaz, B. Boulanger, P. Hubert, J. Chromatogr. A 1158 (2007) 111.

[29] Guidance for Industry: Bioanalytical Method Validation, US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), 2001, May.

[30] E.K. Shibuya, J.E.S. Sarkis, O. Negrini-Neto, J.P.H.B. Ometto, J. Braz. Chem. Soc. 18 (2007) 205.

Section III.2.2

Application à une monographie de la Pharmacopée Européenne

Optimisation et validation d'une méthode CLHP rapide pour la quantification du sulindac et de ses impuretés

F. Krier, M. Brion, B. Debrus, P. Lebrun, A. Driesen, E. Ziemons, B. Evrard, Ph. Hubert, Optimization and validation of a fast HPLC method for the quantification of sulindac and its related impurities, J. Pharm. Biomed. Anal. 54 (2011) 694.

Résumé

La Pharmacopée Européenne décrit une méthode de chromatographie liquide pour la quantification du sulindac. Cette méthode en phase normale utilise une phase mobile quaternaire incluant le chloroforme et souffre d'un temps d'analyse relativement long. Dans la présente étude, une nouvelle méthode en CLHP, utilisant une colonne courte remplie de particules de silice greffée d'un diamètre inférieur à 2 µm et pouvant être utilisée sur un équipement de CLHP standard, a été développée. Les conditions chromatographiques ont été optimisées à l'aide d'une nouvelle méthodologie basée sur les plans d'expériences afin d'obtenir une séparation optimale. Quatre facteurs ont été étudiés : la durée de l'étape isocratique initiale, la proportion de modificateur organique au début du gradient, la proportion de modificateur organique à la fin du gradient et la durée de ce dernier. La condition opératoire optimale a permis la séparation du sulindac et de 3 de ses impuretés apparentées en 6 minutes au lieu de 18 minutes pour la méthode de la Pharmacopée. Finalement, la méthode a été validée, avec succès, en utilisant l'approche des profils d'exactitude afin de démontrer son aptitude à quantifier les composés étudiés avec précision.

Summary

The European Pharmacopoeia describes a liquid chromatography (LC) method for the quantification of sulindac, using a quaternary mobile phase including chloroform and with a rather long run time. In the present study, a new method using a short sub-2 µm column, which can be used on a classical HPLC system, was developed. The new LC conditions (without chloroform) were optimised by means of a new methodology based on design of experiments in order to obtain an optimal separation. Four factors were studied: the duration of the initial isocratic step, the percentage of organic modifier at the beginning of the gradient, the percentage of organic modifier at the end of the gradient and the gradient time. The optimal condition allows the separation of sulindac and of its 3 related impurities in 6 min instead of 18 min. Finally, the method was successfully validated using an accuracy profile approach in order to demonstrate its ability to accurately quantify these compounds.

1. Introduction

Sulindac is a non-steroidal anti-inflammatory drug with a poor pharmacological activity. It can be metabolised by reversible reduction into a sulphide metabolite, with a high pharmacological activity, or by irreversible oxidation into a sulphone metabolite, which has no pharmacological activity [1]. E-sulindac is an impurity resulting from the synthesis of sulindac. Sulindac is used for the treatment of acute or chronic inflammatory conditions. High performance liquid chromatography (HPLC) is the most commonly used technique for the analysis of sulindac and its related impurities [2–6]. The European Pharmacopoeia (Eur. Ph.) [7] describes a normal phase HPLC method for this purpose. However, this method requires a quaternary mobile phase containing chloroform and a rather long analysis time. In pharmaceutical, biomedical and food analysis, the recent trend has been to develop fast analysis methods (fast-HPLC) [8–11].

Over the last decade, different strategies have been developed to reduce the analysis time while maintaining efficiency or to improve efficiency with a similar analysis time [12,13]. One of these strategies, ultra-performance liquid chromatography (UHPLC) uses a column packed with sub-2µm particles. However, such small particles generate high back pressure and necessitate the use of appropriate and quite expensive equipment able to withstand such an ultra–high pressure (up to 1200 bars). In order to reduce back pressure, some column manufacturers put forward the use of short sub-2µm columns with a higher internal diameter. The advantage of this type of column is that it can be used with conventional HPLC equipment (up to 400 bars) [14]. In this study, a Platinum C18 Rocket column (53×7mmi.d., 1.5µm particle size) was used. This column provides a shorter analysis time, a lower solvent, faster equilibration, accurate quantitation and high efficiency [15,16]. Some recent papers report methods based on this brand of column for the determination of pharmaceutical substances in plant tissues [17–19], for the monitoring of amiodarone [20], or for the separation of microcystins and nodularins on narrow-bore [21]. In this study, a new fast and easy HPLC method was developed without the use of chloroform for the separation and determination of sulindac and its related impurities. This method was optimised in terms of separation by using design of experiments (DoE) methodology [22–24] and the design space (DS) concept [25,26]. Finally, the method was validated using the accuracy profile approach based on a β-expectation tolerance interval [27–32].

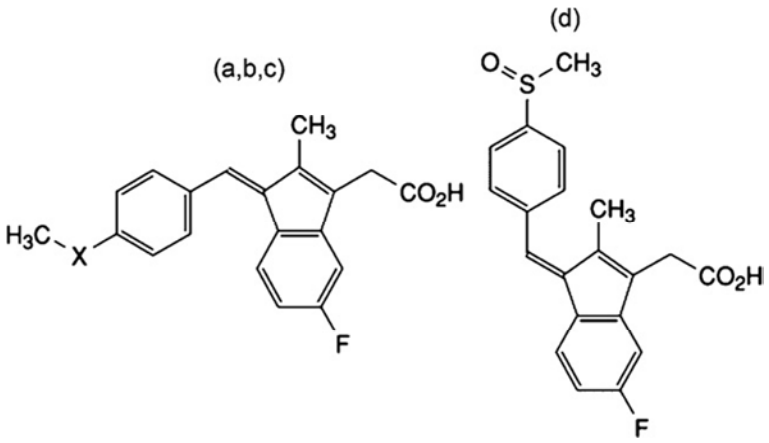


Fig. 1 Chemical structures of sulindac and its impurities. (a) Sulindac (X = SO), (b) sulindac sulphone (X = SO_2), (c) sulindac sulphide (X = S), (d) E-sulindac.

2. Materials and methods

2.1. Chemicals and solvents

Acetonitrile (LiChrosolv, 99.9%), acetic acid (99.8%) and ethyl acetate for analysis (99.5%) were purchased from Merck (Darmstadt, Germany). Absolute ethanol (99.99%) was purchased from Fisher Scientific (Loughborough, UK) and chloroform Chromasolv (99.8%) from Sigma–Aldrich (Buchs, Switzerland). Sulindac was supplied by EP (Strasbourg, France). Sulindac sulphone and sulindac sulphide were purchased from Sigma (Steinheim, Germany). E-sulindac was supplied by the United States Pharmacopeia (Rockville, USA). The chemical structures of sulindac and its related impurities chemical structures are shown in Fig. 1.Water was purified using a Millipore system (18.2 MΩ/cm resistivity, Milli-Q) and filtered through a 0.22 µm Millipore Millipak®-40 disposable filter unit (Millipore Corporation, USA). Phosphate buffer solution (pH 7.4; 50 mM) was prepared by dissolving 27.2 g of potassium dihydrogen phosphate (Merck), 6.3 g of sodium hydroxide (VWR, Leuven, Belgium) and 0.64 g of sodium azide (Merck) in 2573.6 mL deionised water. The pH value was adjusted to 7.4 with 0.1 M sodium hydroxide solution. Phosphate buffer solution (pH 2.0; 10 mM) was prepared by dissolving 4 mL phosphoric acid (Aldrich) and 12.12 g potassium dihydrogen phosphate (Merck) in 2000.0 mL deionised water. The pH value was then adjusted to 2.0 with a 1.2 M hydrochloric acid solution.

2.2. Apparatus

Analyses were performed on an Agilent technologies HPLC 1100 series system (Agilent Technologies, Santa Clara, CA, USA) equipped with a solvent delivery binary pump G1312A, an on-line degasser G1379A, a thermostatised autosampler G1328A, a column oven G1316A and a diode-array detector G1316A. Chemstation® (Rev.B.01.03[204]) was used to control the whole chromatographic system and to acquire, process and store all the data obtained.

A Mettler Toledo (Schwerzenbach, Switzerland) MX5 microbalance was used to weigh all the compounds (precision: 1 µg). A Seven Easy Mettler Toledo pH meter was used to adjust the pH value.

2.3. Chromatographic conditions

2.3.1. Reference method [7]

The reference method used is the one described for a related substance of sulindac in Eur. Ph. 6.7. This reference method is in the normal phase mode. Chromatographic analyses were performed on an Alltima Silica column (250×4.6 mm i.d., 10 µm particle size), which was kept at 20 °C. The mobile phase was prepared by mixing acetic acid, ethanol, ethyl acetate and chloroform (1:4:100:400 (v/v/v/v)) and was degassed before use. The HPLC system was operated in isocratic mode at a flow rate of 2.0 mL/min and the injection volume was 20 µL. UV detection was performed at 280 nm.

2.3.2. Optimized method

Sulindac is a weak acid and has a pKa of 4.7. Related impurities have a pKa near to 4.7. Then, it was decided to work with a mobile phase composed of acetonitrile and a phosphate buffer solution (pH 2.0; 10 mM). At the selected mobile phase pH, the compounds are almost totally non-ionised. Thus, to have good separation, a C18 silica packed column was chosen. The chromatographic analyses were performed on a Platinum C18 Rocket column (53×7 mm i.d., 1.5 µm particle size), which was kept at 35 °C. Despite the high internal diameter of the column, in order to avoid being concerned by too high pressures, acetonitrile was selected as an organic modifier as it has a lower viscosity than methanol. The mobile phase proportion and the gradient shape were optimised by means of DoE methodology in order to find the design space (DS) (see Section 3.4). The injection volume was 100 µL. UV detection was performed at 340 nm. The HPLC system was operated with a gradient mode at a flow rate of 3.0 mL/min. Following the gradient method transfer rules [9] with as a starting column geometry, a conventional C18 150×4.6 mm i.d. (5 µm) operated at 1 mL/min, the Platinum C18 Rocket column should be used with a 7.7 mL/min flow rate. Nevertheless, at this flow rate, the back pressure is higher than the maximal operating back pressure on an Agilent technologies HPLC 1100 series system (i.e. >350–400 bars). Therefore, to maintain a backpressure lower than 350 bars and not too much sacrificing performance, the flow rate was adjusted to the above mentioned value of 3.0 mL/min. Thanks to the high column internal diameter, the back pressure obtained at this flow rate was only about 280 bars.

2.4. Standard solutions

A stock solution of sulindac was prepared by dissolving 50.0 mg of sulindac in 50.0 mL of methanol (1 mg/mL). A stock solution of E-sulindac, sulindac sulphide and sulindac sulphone was prepared by dissolving 2.0 mg of each compound in 20.0 mL of methanol. The calibration and validation standards [27–29] were prepared by mixing and diluting the stock solutions with phosphate buffer solution (pH 7.4; 50mM) to reach the concentration levels: 100/10; 100/5; 100/1; 100/0.5; 50/0.25; 25/0.125; 1/0.005; 0.5/0.0025 µg/mL (sulindac concentration/concentrations of related impurities, respectively).

3. Theory/calculations

3.1. Computations

The experimental design results were processed using the software R version 2.7.2. The validation data were processed with the software e-noval version 2.0e (Arlenda s.a., Liège, Belgium).

3.2. Experimental design

Four HPLC factors were investigated using DoE methodology through a design matrix. All of the factors were quantitative:

- The percentage of acetonitrile at the beginning of the gradient (ACNlower),
- The percentage of acetonitrile at the end of the gradient (ACNupper),
- The gradient time (Gradient time),
- The duration of the initial isocratic step (Plateinit).

To assess the reproducibility of runs, three central points were repeated independently twice for each “Plateinit” level (see Table 1). In addition, quadratic effects were assumed for each factor except for “Plateinit” which was considered as linear. In order not to miss an effect or interaction which would not be assessed or modelled by the chromatographic theory, no a priori knowledge about the empiric compound behaviour was added in the multiple linear models. A full factorial design was chosen to explore the experimental domain. 58 experiments (3^3×2+4) were carried out in order to explore the experimental domain. It is important to bear in mind that these experiments were used to models the chromatographic behaviour of each peak and concurrently evaluate the method robustness. In practice, experiments were sufficiently rapid and easily carried out to overcome this high number of experiment. Levels of each factor are summarized in Table 1.

Table 1 Description of the levels of four factors involved in the experimental design.

	Plateinit (min)	ACNlower (%)	ACNupper (%)	Gradient time (min)
Levels	0 – 1	15 – 30 – 55	55 – 60 – 65	1 – 3 – 5
Central points	0 – 0.5 – 1	30	60	3
Maximum effects	Linear	Quadratic	Quadratic	Quadratic

3.3. Statistical models

In the resulting chromatograms, four peaks were detected and indexed at their beginnings, apexes (retention time) and ends as proposed by Lebrun et al. [26]. The logarithms of the retention factors (log(*k*)) were selected as responses of the multi-linear model following Eq. (1).

$$\begin{aligned}\log(k) = {} & \beta_0 + \beta_1 \times Plateinit + \beta_2 \times ACNlower + \beta_3 \times ACNlower^2 \\ & + \beta_4 \times ACNupper + \beta_5 \times ACNupper^2 + \beta_6 \times Gradient \\ & + \beta_7 \times Gradient^2 + \beta_8 \times ACNlower \times Plateinit + \beta_9 \\ & \times ACNupper \times Plateinit + \beta_{10} \times Gradient \times Plateinit \\ & + \beta_{11} \times ACNlower \times Plateinit + \beta_{12} \times ACNlower \\ & \times Gradient + \beta_{13} \times ACNupper \times Gradient + \varepsilon\end{aligned} \quad (1)$$

where β_0,. . ., β_{13} are the parameters to estimate and ε is the predictive error, for instance using the classical least squares regression.

3.4. Optimisation – design space

The separation of the critical pair (i.e. the two most proximate peaks) was optimised using the methodology presented in the publications of Lebrun et al. and Dewé et al. [26,33]. The design space (DS) is the zone of the experimental domain where the quality of the chromatogram is high. The separation criterion is defined as the difference between the beginning of a peak and the end of the preceding peak. This criterion was calculated from the predicted retention times. The analysis of the model's predictive error was also carried out to find the DS, i.e. the set of operating conditions that gives a high probability of obtaining a separation criterion higher than a given threshold. Mathematically, the DS is defined as in Eq. (2),

$$\mathrm{DS} = \left\{ x_0 \in \chi | E_{\hat{\theta}} \left[\mathrm{P}(S > \lambda) | \hat{\theta} \right] \geq \pi \right\} \quad (2)$$

where x_0 is the set of factor conditions belonging to the experimental domain χ, for which the expected probability of having a minimal separation (min(*S*)) higher than λ, is higher than π, given the uncertainty of the estimation of model parameters θ = (β_0, . . ., β_{13}, ε).

The objective of this study was to determine the optimal chromatographic conditions allowing us to obtain a separation criterion of at least λ = 0 min (i.e. baseline resolved peaks) with a probability of at least 90% (π = 0.9). Monte-Carlo simulations were then performed to propagate a predictive error from responses to the separation criterion. This step was

carried out to estimate the error affecting the criterion in order to compute the probability of reaching the separation time of 0 min [33].

4. Results and discussion

4.1. Reference method

As can be seen in Fig. 2, the reference method enabled the separation of all the compounds within 18 min and was completed within 25 min. Under these conditions, the separation of all compounds was acceptable. The Eur. Ph. method uses chloroform, which is now considered as a toxic organic solvent. Furthermore, the relatively long analysis time involved in that method, in combination with a high flow rate of 2 mL/min, implies a large consumption of chloroform. Consequently, it was of particular interest for us to develop a new method avoiding the use of this kind of solvent and presenting a shorter analysis time.

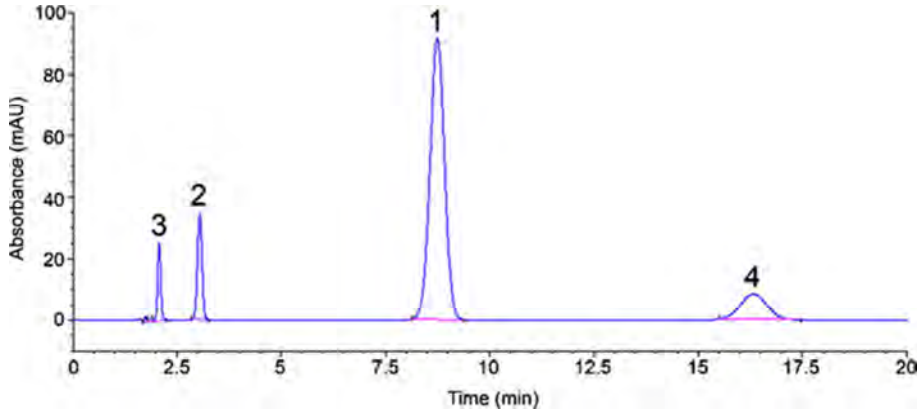


Fig. 2 Reference method chromatogram (European Pharmacopoeia 6.7 (1: sulphide, 2: sulphone, 3: sulindac, 4: E-sulindac)).

4.2. Optimized method

Taking into account the fact that the aim of this study was to develop a short run time method, “Plateinit” was limited to 1 min, gradient time to 5 min and after some preliminary screening experiments, “ACNlower” was set from 15 to 55% and “ACNupper” from 55 to 65%. Indeed, a percentage of “ACNlower” less than 5% is not recommended for the integrity of the column while a percentage higher than 55% would not be able to separate satisfactorily the different compounds. Meanwhile, a percentage of “ACNupper” higher than 65% increased drastically the slope of the gradient. A baseline drift could be prejudicial to the quality of the quantitative results. Fig. 3 illustrates the quality of the fit of the observed retention times versus the predicted retention times using the statistical models previously described. Most of the residuals were mainly located within the interval [−0.2 min, 0.2 min]. However, the overall quality of fit was considered as good as the adjusted R^2 values for each model were higher than 0.95. Fig. 4 shows the probability surfaces in different directions of

the space around the optimal solution (for each graph, two factors were fixed at their optimal values).

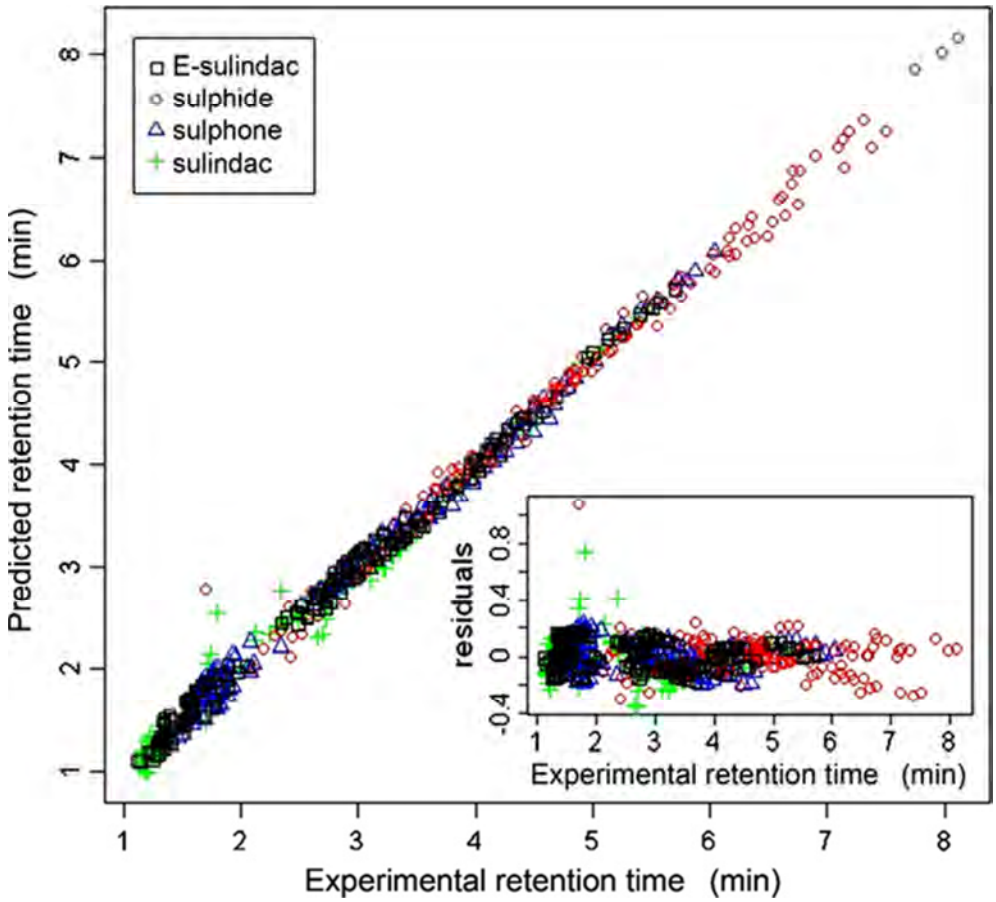


Fig. 3 Experimental retention times versus predicted ones. Residuals are depicted at the bottom right corner.

As we can see, the best probability surface was obtained when the duration of the initial isocratic plate was around 1 min. The probability of separating the critical pair is higher when the gradient time was around 3.6 min: when this time increases, the probability of separation decreases. We found a good probability of separation when the percentage of organic modifier at the beginning of the gradient was between 40 and 45% and when the percentage of organic modifier at the end of the gradient was 55%. The optimal values allowing the achievement of the higher probability (P(separation > 0) > 0.9) ensuring a separation of at least 0 min with baseline-resolved peaks were 1 min for the Plateinit, 40.5% for the ACNlower, 55% for the ACNupper and the gradient time was 3.6 min.

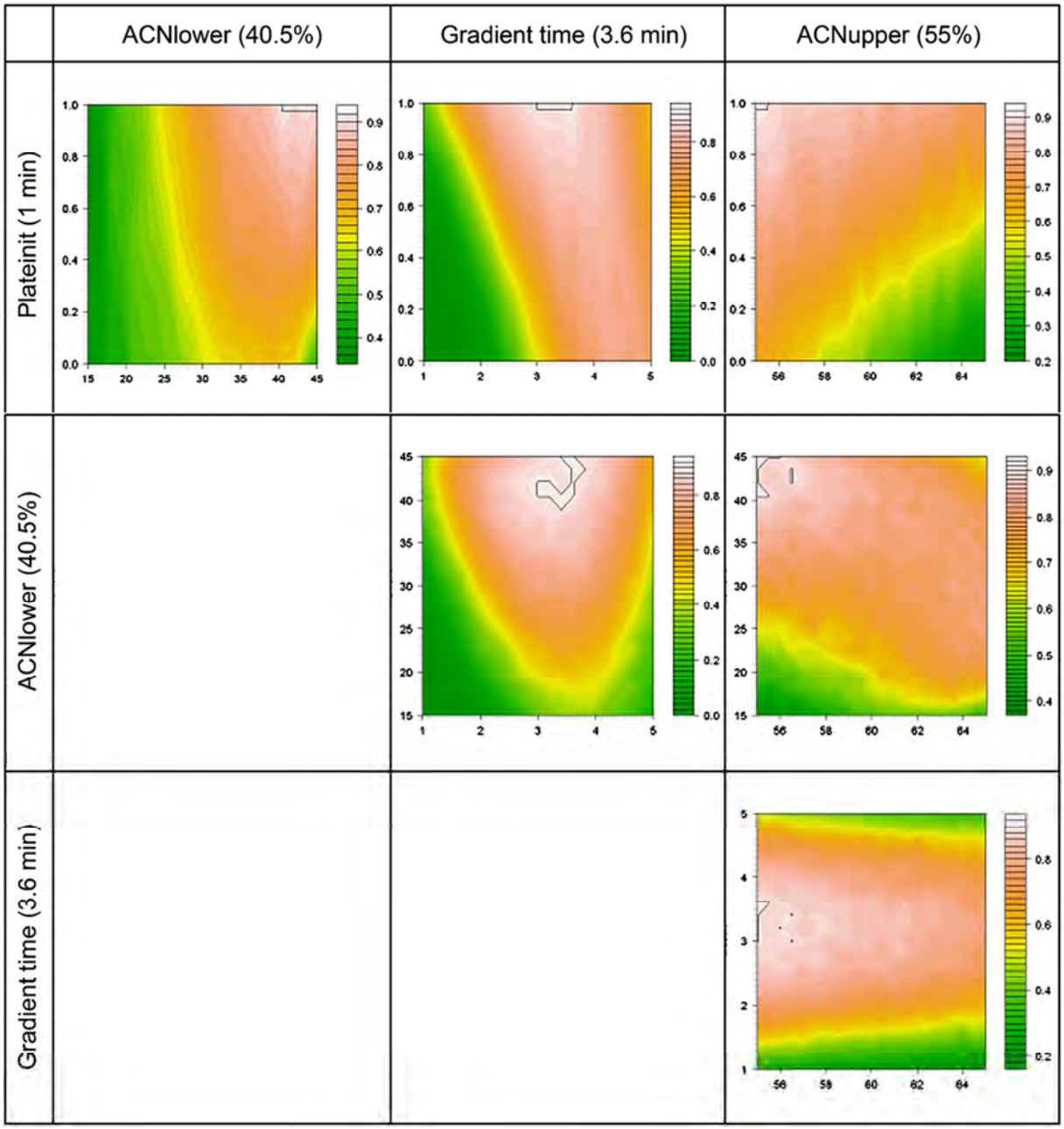


Fig. 4 Surface of probability to reach $S > 0$. The design space is surrounded by black lines for an expected probability to have well-separated peaks is 0.9. Factors optimal values are placed between parentheses.

The developed HPLC method for the quantification of sulindac and its related impurities divided the run time of analyses by three compared to the reference method. Fig. 5a and b shows the optimal predicted and experimental chromatograms. As can be seen, the predicted retention times were found to be very close to the experimental values and an acceptable separation was obtained within an analysis time of 6 min.

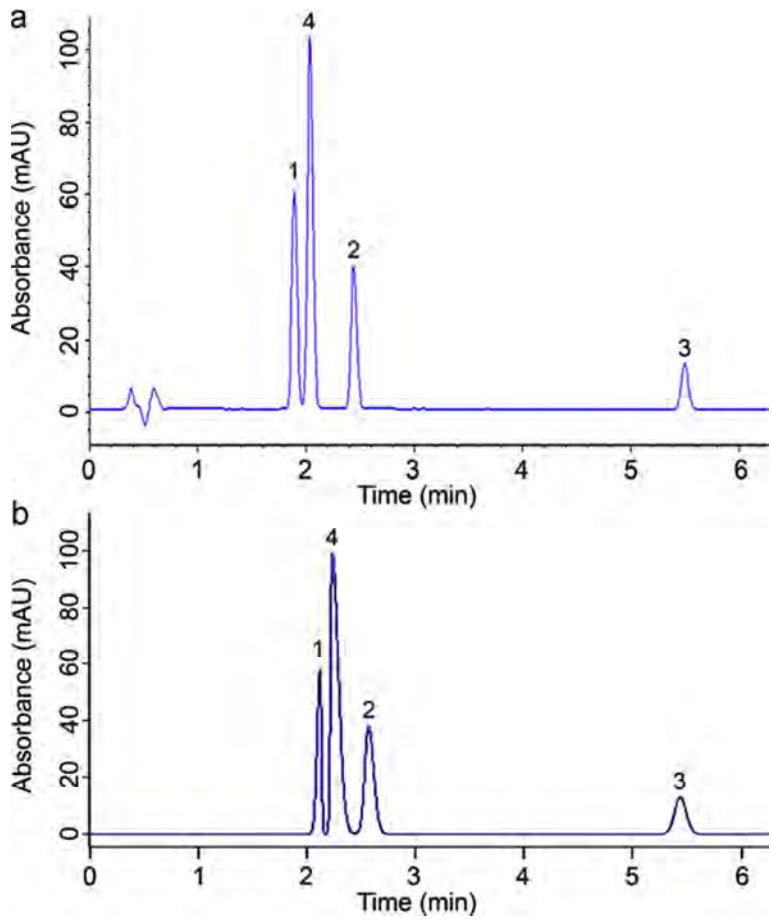


Fig. 5 (a) Experimental chromatogram recorded at optimal solution. (b) Predicted chromatogram at optimal condition (1: sulindac, 2: sulphide, 3: sulphone, 4: Esulindac).

4.3. Method validation

The aim of this validation was to establish that the analytical method was suitable for the quantitative determination of the 4 compounds described above. In other words, it was to demonstrate the method ability to quantify them. Several widely recognised validation criteria were evaluated (selectivity, trueness, precision, accuracy, and limits of quantification and detection) in accordance with ICH Q2(R1) guidelines [31]. An original approach using accuracy profiles based on tolerance intervals was applied to evaluate the reliability of the results. The tolerance interval used was a “β-expectation tolerance interval” defining an interval in which it is expected that each future result will fall with a defined probability (β). It is therefore a predictive methodology. This tolerance interval is computed for each validation standard concentration level, using their estimated intermediate precision standard deviation and bias. By joining together the upper tolerance limits on the one hand and the lower tolerance limits on the other hand, the method defines an accuracy profile. As long as this profile stays within the acceptance limits set according to the needs of the final user or to regulatory expectations the method can be considered as valid. Indeed, it guarantees that future results will be included in the a priori set acceptance limits with at least a probability of β (e.g. 0.95%).

4.3.1. Response function

The response function represents the relationship between the detector response and the concentration of the analyte in the sample. In order to determine the most appropriate calibration curve, different models were evaluated. Due to the wide concentrations range, the calibration curve for sulindac was built with 5 calibration standards, from 0.5 µg/mL to 100.0 µg/mL. Independent validation standards were also prepared (5 validation standards from 0.5 µg/mL to 100.0 µg/mL). For the related impurities (E-sulindac, sulindac sulphone and sulindac sulphide), the concentrations range of the calibration standards was reduced. Only 4 calibration standards from 10 µg/mL to 0.125 µg/mL were released. Each calibration and validation standard was analysed in triplicate and four consecutive series were performed in order to evaluate the intermediate precision. The concentration of the validation standards was then back-calculated to determine the mean relative bias as well as the standard deviation for repeatability and intermediate precision, in accordance with the work of Rozet et al. [34]. On this basis, different accuracy profiles were then plotted to select the most appropriate calibration model for the analytical method. The acceptance limits were set at ±5% for sulindac and ±15% for its related substances and the upper and lower β-expectation tolerance interval were set at 95%. The best response function was achieved with a weighted (1/X) quadratic regression transformation for sulindac, with linear regression after logarithm transformation for E-sulindac and sulphone and with a linear regression through 0 fitted with the higher level of concentration only for sulphide.

4.3.2. Trueness

Trueness is defined as the closeness of agreement between the average value obtained from a large series of test results and an accepted reference value [31]. From the results obtained and considering the appropriate model for each compound, the concentration of the validation standards was back-calculated to determine trueness expressed in terms of relative bias (%). As can be seen in Table 2, the relative biases never exceeded 1% for sulindac or 7% for the impurities.

4.3.3. Precision

Precision is defined as the closeness of agreement (degree of scatter) between a series of measurements obtained from multiple sampling of the same homogeneous sample under the prescribed conditions [31]. Precision was computed here using the relative standard deviations (RSDs) for repeatability and intermediate precision at each concentration level of the validation standards. The results are indexed in Table 2. It can be seen that RSD values never exceeded 2 or 4% for repeatability or intermediate precision, respectively, illustrating the good precision of the analytical method.

4.3.4. Accuracy

Following ICH Q2(R1) guidelines [31], accuracy is defined as the closeness of agreement between the value which is accepted either as a conventional true value or an accepted reference value and the value found. Upper and lower β-expectation tolerance intervals (%) were calculated and did not exceed the acceptance limits for each concentration level, as shown in Fig. 6. These observations proved that the HPLC method used in this study was accurate over the studied concentration range. In fact, β-expectation tolerance interval based accuracy profiles allow us to guarantee that a high percentage (e.g. 95%) of future results will lie within the acceptance limits [31].

4.3.5. LOQs and LODs

The limits of quantification (LOQs) were obtained by calculating the smallest and highest concentration level beyond which accuracy limits or β-expectation limits go outside the acceptance limits.

Table 2 Validation criterion for sulindac, E-sulindac, sulphone and sulphide.

Validation criteria	Sulindac	E-sulindac	Sulphide	Sulphone
Response function	Weighted (1/X) quadratic regression	Linear regression after logarithmic transformation	Linear regression through 0	Linear regression after logarithmic transformation
Trueness (concentration (µg/ml)/relative bias (%))				
Level 1	101.4/-0.01	10.25/3.39	9.74/-0.05	10.13/2.37
Level 2	50.68/0.43	5.12/-1.71	4.87/-0.45	5.06/-3.36
Level 3	25.32/0.04	1.025/-2.72	0.97/3.46	1.01/0.08
Level 4	1.01/0.01	0.13/-0.29	0.12/6.63	0.13/-0.11
Level 5	0.51/0.78	-	-	-
Precision (repeatability (RSD%)/intermediate precision (RSD%))				
Level 1	0.53/1.38	0.45/1.76	0.43/1.28	0.57/1.54
Level 2	0.53/0.60	0.58/0.98	0.45/0.96	0.57/1.08
Level 3	0.60/0.70	0.89/1.37	0.87/1.93	0.83/1.77
Level 4	0.67/0.67	1.61/3.93	0.71/2.71	0.92/2.70
Level 5	0.94/1.51	-	-	-
Accuracy (lower β-EL (%)/upper β-EL (%))				
Level 1	-4.22/4.19	-2.57/9.35	-4.23/4.14	-2.58/7.31
Level 2	-0.99/1.85	-4.46/1.03	-3.37/2.47	-6.53/0.19
Level 3	-1.65/1.74	-6.47/1.02	-2.47/9.40	-5.27/5.44
Level 4	-1.53/1.55	-12.63/12.05	-2.45/15.77	-8.9/8.69
Level 5	-3.39/4.96	-	-	-
Linearity				
Slope	0.9997	1.0330	0.9969	1.020
Intercept	0.0422	-0.0792	0.0159	-0.0659
R^2	0.9996	0.9989	0.9997	0.9988

The lower and upper limits of quantification were evaluated from the corresponding accuracy profile at 0.5063 and 101.4 µg/mL for sulindac, at 0.1281 and 10.25 µg/mL for E-sulindac, at 0.2244 and 9.738 µg/mL for sulphide and at 0.1266 and 10.13 µg/mL for sulphone. The limits of detection (LODs) were computed at 0.1380 µg/mL for sulindac, at 0.0388 µg/mL for E-sulindac, at 0.0680 µg/mL for sulphide and at 0.0383 µg/mL for sulphone.

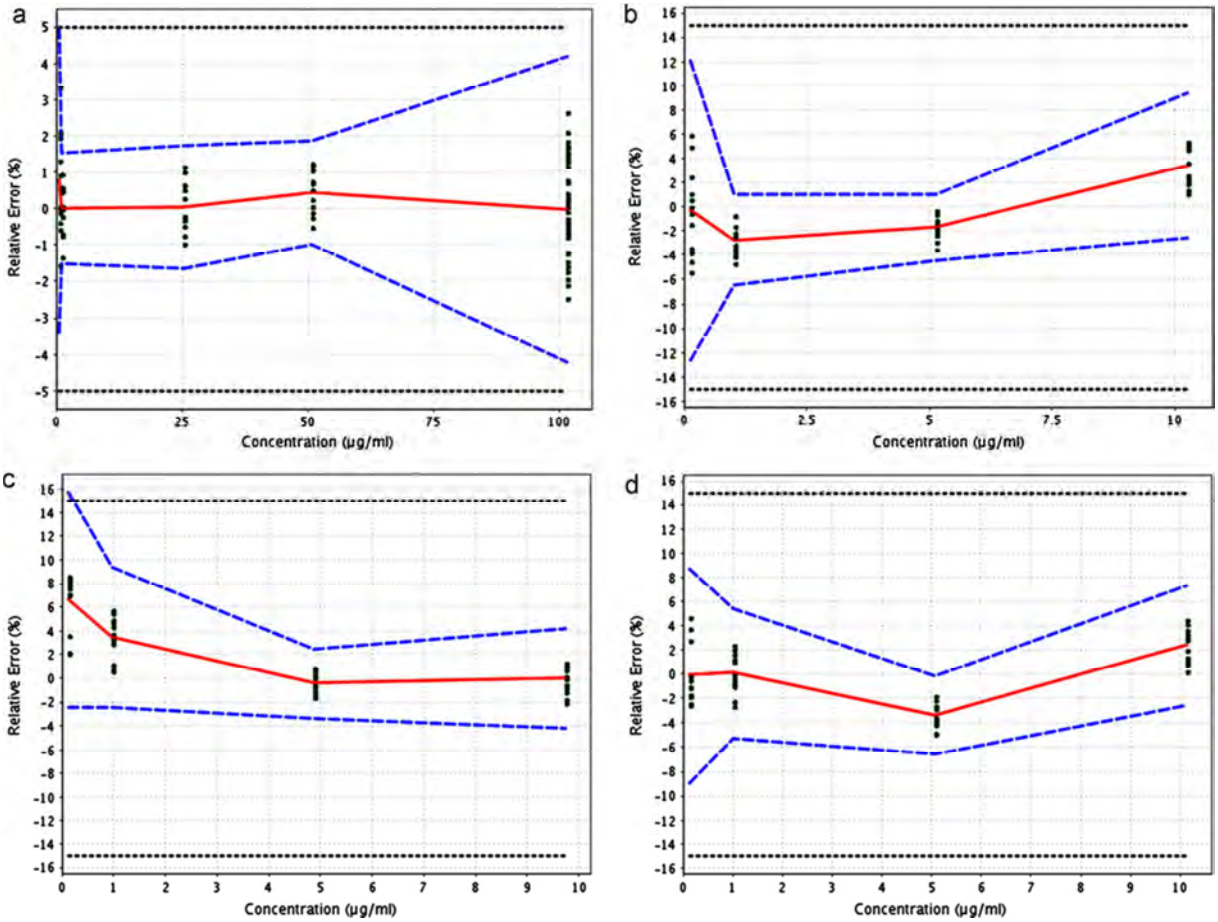


Fig. 6 Accuracy profiles of (a) sulindac, (b) E-sulindac, (c) sulphide and (d) sulphone. Relative bias (red line), ±5% acceptance limits (black dotted lines), 95% (sulindac) or 85% (related impurities) β-expectation tolerance limits (blue dashed line), and relative back-calculated concentrations (black points).

5. Conclusions

An analytical method for the quantification of sulindac and its related impurities was developed using a short column with sub-2 µm particles on a classical HPLC system. This method was optimised using DoE methodology and the DS concept. Under optimised conditions, the analysis time was considerably reduced (by about 3-fold). The solvent consumption was reduced by about 2-fold and the equilibration was reduced by about 1.5-fold. Furthermore, we did not use chloroform unlike in the Eur. Ph. Reference method. Finally, this particular method was validated successfully using an accuracy profiles approach for sulindac and its related substances. Several criteria were successfully tested (trueness, precision, accuracy, and limits of quantification and detection).

Acknowledgements

We would like to thank Stéphane Counerotte and Laurence Collard for their technical help. The author would also like to thank the Walloon Region of Belgium for the FIRST-DEI convention funds N°516130, 516138 and for the Waléo 2 convention funds N°516266.

References

[1] D.E. Duggan, K.F. Hooke, E.A. Risley, T.Y. Shen, C.G. Arman, J. Pharmacol. 201 (1977) 8.
[2] M. Siluveru, J.T. Stewart, J. Chromatogr. B 673 (1995) 91.
[3] C.R. Clark, C.L. McMillian, J.F. Hoke, K.D. Campagna, W.R. Ravis, J. Chromatogr. Sci. 25 (1987) 247.
[4] T. Hirai, S. Matsumoto, I. Kishi, J. Chromatogr. B 692 (1997) 375.
[5] D.G. Musson, W.C. Vincek, M.L. Constanzer, T.E. Detty, J. Pharm. Sci. 73 (1984) 1270.
[6] B.N. Swanson, V.K. Boppana, J. Chromatogr. 225 (1981) 123.
[7] Sulindac, European Pharmacopoeia, 6th ed., Council of Europe, Strasbourg, 2010.
[8] J.J. Destefano, T.J. Langlois, J.J. Kirkland, J. Chromatogr. Sci. 46 (2008) 254.
[9] D. Guillarme, D.T. Nguyen, S. Rudaz, J.L. Veuthey, Eur. J. Pharm. Biopharm. 68 (2008) 430.
[10] D. Guillarme, J.-L. Veuthey, V.R. Meyer, LC-GC Europe 21 (2008).
[11] Z.M. Qian, J. Lu, Q.P. Gao, S.P. Li, J. Chromatogr. A 1216 (2009) 3825.
[12] D. Guillarme, J. Ruta, S. Rudaz, J.-L. Veuthey, Anal. Bioanal. Chem. 397 (2010) 1069.
[13] D. Guillarme, J. Schappler, S. Rudaz, J.-L. Veuthey, TRAC – Trend. Anal. Chem. 29 (2010) 15.
[14] S. Fekete, K. Ganzler, J. Fekete, J. Pharm. Biomed. Anal. 51 (2010) 56.
[15] M. Dong, Today's chemist at work 9 (2000) 46.
[16] R. Dams, W.E. Lambert, K.M. Clauwaert, A.P. De Leenheer, J. Chromatogr. A 896 (2000) 311.

[17] M.W. Davey, E. Stals, B. Panis, J. Keulemans, R.L. Swennen, Anal. Biochem. 347 (2005) 201.
[18] M.W. Davey, E. Dekempeneer, J. Keulemans, Anal. Biochem. 316 (2003) 74.
[19] A.P. Griffith, M.W. Collison, J. Chromatogr. A 913 (2001) 397.
[20] J.M. Juenke, P.I. Brown, G.A. McMillin, F.M. Urry, J. Anal. Toxicol. 28 (2004) 63.
[21] L. Spoof, M.R. Neffling, J. Meriluoto, Toxicon 55 (2010) 954.
[22] G.E. Box, W.G. Hunter, J.S. Hunter, W.G. Hunter, Statistics for Experimenters: Design, Innovation, and Discovery, 2nd edition, Wiley, 2005.
[23] G. Cox, W. Cochran, Experimental Designs, 2nd edition, Wiley, 1992, 1957.
[24] R.H. Myers, D.C. Montgomery, C.M. Anderson-Cook, Response Surface Methodology: Process and Product Optimization Using Designed Experiments, 3rd edition, Wiley, 2009.
[25] J.J. Peterson, J. Biopharm. Stat. 18 (2008) 959.
[26] P. Lebrun, B. Govaerts, B. Debrus, A. Ceccato, G. Caliaro, Ph. Hubert, B. Boulanger, Chemometr. Intell. Lab. 91 (2008) 4.
[27] Ph. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, J. Pharm. Biomed. Anal. 36 (2004) 579.
[28] Ph. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, E. Rozet, J. Pharm. Biomed. Anal. 45 (2007) 70.
[29] Ph. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, J. Pharm. Biomed. Anal. 45 (2007) 82.
[30] Ph. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, N. Cohen, P.A. Compagnon, W. Dewé, M. Feinberg, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, J. Pharm. Biomed. Anal. 48 (2008) 760.
[31] International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use, Topic Q2 (R1): Validation of Analytical Procedures: Text and Methodology, Geneva, 2005.
[32] E. Rozet, C. Hubert, A. Ceccato, W. Dewé, E. Ziemons, F. Moonen, K. Michail, R. Wintersteiger, B. Streel, B. Boulanger, Ph. Hubert, J. Chromatogr. A 1158 (2007) 126.
[33] W. Dewé, R.D. Marini, P. Chiap, Ph. Hubert, J. Crommen, B. Boulanger, Chemom. Intell. Lab. Syst. 74 (2004) 263.
[34] E. Rozet, A. Ceccato, C. Hubert, E. Ziemons, R. Oprean, S. Rudaz, B. Boulanger, Ph. Hubert, J. Chromatogr. A 1158 (2007) 111.

Chapitre IV

Conclusions et perspectives

La chromatographie liquide est la technique analytique de choix dans le domaine pharmaceutique. Elle est majoritairement utilisée pour séparer, identifier et quantifier les composés d'intérêt et leurs substances apparentées. Que ce soit pour le contrôle de qualité des médicaments, l'étude de leur stabilité ou encore pour fournir des informations qualitatives et quantitatives aux chercheurs dans des disciplines aussi variées que la chimie de synthèse, la toxicologie ou la pharmacie galénique, le recours à la chromatographie liquide est aujourd'hui devenu une nécessité. La recherche visant à obtenir des stratégies innovantes pour le développement de méthodes chromatographiques prend donc une place de plus en plus importante dans le domaine analytique.

Dans le contexte actuel, dicté par les concepts de *quality by design* (QbD) et du *design space* (DS) préconisés par l'ICH, les méthodologies d'optimisation de méthodes chromatographiques doivent permettre de fournir aux analystes, en plus de conditions chromatographiques robustes, une bonne compréhension des processus d'optimisation et une crédibilité accrue sur les prédictions réalisées.

C'est pourquoi, dans le cadre de ce travail, une méthodologie combinant à la fois la modélisation du comportement chromatographique des composés et l'optimisation des conditions analytiques a été développée. Cette méthodologie, basée sur les plans d'expériences (*design of experiments* : DoE), la prise en compte de l'erreur de prédiction et la propagation de cette erreur, depuis les réponses modélisées jusqu'au critère de séparation, a permis de formaliser et d'identifier des DS.

Cependant, comme les plans d'expériences font drastiquement varier les conditions analytiques, la sélectivité observée peut être modifiée de façon importante. A ce stade du développement d'une méthode, l'étape de détection et d'appariement des pics chromatographiques est longue et fastidieuse, et ce d'autant plus que le nombre de composés impliqués est grand. Une nouvelle méthodologie a donc dû être développée pour permettre de "lire" automatiquement les chromatogrammes. L'analyse en composantes indépendantes (*independent component analysis* : ICA) a été utilisée à cette fin pour séparer numériquement les composés coéluants et pour estimer les limites d'intégration de ces derniers. En outre, l'analyse statistique des résultats a permis de classifier les composantes indépendantes résultantes en fonction de leur relevance. Les composantes du bruit ou des artefacts ont ainsi été écartées de celles correspondant aux pics chromatographiques, permettant *in fine* de reconstruire des chromatogrammes "épurés". Cette méthodologie a également démontré sa capacité à estimer le nombre de composés présents dans un mélange inconnu.

La stratégie de lecture automatique de chromatogrammes a ensuite été implémentée au processus d'optimisation de séparations chromatographiques afin d'obtenir une méthodologie globale, permettant le développement automatisé de méthodes chromato-

graphiques. A l'aide de deux échantillons, une formulation pharmaceutique et un mélange de composés inconnus, la méthodologie globale a permis de démontrer non seulement son aptitude à optimiser les séparations de composés étudiés mais aussi à accélérer notablement le processus de détection et d'appariement des pics. Cette étude a également confirmé que les DS identifiés correspondaient bien à des zones robustes.

Au moyen de la méthodologie développée, la séparation de 19 antipaludéens a été optimisée et a ainsi permis l'obtention d'une méthode de criblage dans le cadre de la lutte contre la contrefaçon des médicaments antipaludéens. En outre, cette étude s'est également focalisée sur la mise en évidence de la concordance existant entre les modèles mathématiques et la théorie de la chromatographie. En effet, il a été démontré que les effets prédits par les modèles s'accordaient très bien avec la théorie de la chromatographie. Sur cette base, Il a été envisagé que les paramètres résultant de la modélisation des temps de rétention pouvaient être utilisés pour identifier ou corroborer le comportement chromatographique de composés d'intérêt. Par ailleurs, le plan expérimental réalisé lors de cette étude, initialement prévu pour développer une méthode de criblage, a également permis d'optimiser la séparation de trois composés correspondant à une formulation pharmaceutique disponible sur le marché Congolais.

Lors, notamment de deux études portant sur l'optimisation de la séparation de cannabinoïdes, d'une part, et du sulindac et de trois de ses substances apparentées, d'autre part, la méthodologie DoE-ICA-DS a permis d'identifier les DS correspondants et de séparer ces composés avec succès. La fiabilité de conditions proposées a par ailleurs été largement confirmée par les excellentes performances de ces méthodes lors de leur validation.

Au vu des résultats obtenus dans le cadre de ce travail, il convient dès à présent de définir les objectifs futurs et les perspectives.

A ce stade du projet, il nous parait tout à fait raisonnable d'envisager la mise au point d'un programme informatique offrant la possibilité d'optimiser des méthodes analytiques en utilisant la méthodologie DoE-ICA-DS. Le fait que cette méthodologie est en parfait accord avec le concept de QbD laisse présager que la transposition de la technologie est pertinente. Cette méthodologie permettant d'identifier le DS pourrait également être appliquée à d'autres techniques analytiques telles que l'électrophorèse capillaire, la spectroscopie proche infrarouge ou la spectroscopie Raman.

De la même manière, il n'est pas déraisonnable de penser que cette méthodologie pourrait être utilisée pour l'optimisation de processus sortant du cadre analytique et plus particulièrement pour les procédés de production en technologie pharmaceutique.

Concernant les méthodes chromatographiques, et en se basant sur le fait que le but ultime de ces dernières réside dans le dosage de composés d'intérêt, il pourrait aussi être très intéressant d'identifier des DS calculés à l'aide de critères de quantification. Le DS ne définirait alors pas uniquement une zone de séparation optimale mais également une région où l'exactitude de dosage atteindrait des limites définies par l'utilisateur, tel que c'est le cas actuellement lors de la validation de méthodes analytiques utilisant le profil d'exactitude comme outil de décision. L'écart entre les bornes des intervalles du profil et les limites d'acceptation servirait alors de critère pour l'implémentation du DS.

L'utilisation des statistiques Bayésiennes (ou inférence Bayésienne) représente également une perspective intéressante. L'essor relativement récent des statistiques Bayésiennes est principalement dû aux capacités de calcul croissantes des ordinateurs. Malgré certaines divergences existant entre les statistiques Bayésiennes et les statistiques "fréquentistes", elles peuvent être considérées comme complémentaires. Pour rester concis, trois avantages notables des statistiques Bayésiennes peuvent être énoncés.

- Premièrement, Les statistiques Bayésiennes offrent la possibilité d'inclure, avant même de commencer les calculs statistiques, des informations initiales (des distributions *a priori*) qui synthétisent des informations ou des connaissances, autres que les données. Lors des calculs Bayésiens, ces distributions initiales s'ajustent aux données et aboutissent à des distributions finales représentatives de probabilités. Dans le cadre de la modélisation en CL, l'avantage capital est que l'analyste peut inclure aux calculs statistiques des informations chromatographiques tirées de son expérience ou de théories afférentes.
- Deuxièmement, en utilisant les statistiques fréquentistes, le DS est basé sur la prédiction de réponses moyennes. Dès lors, l'interprétation de ces prédictions peut être que, au futur, en moyenne, on observera une proportion (une fréquence) d'expériences donnant les résultats prédits. Au contraire, dans le cadre des statistiques Bayésiennes, les prédictions peuvent facilement être réalisées sur les réponses individuelles permettant de donner une garantie, une probabilité, telle que chaque future expérience individuelle donnera les résultats escomptés. A ce stade, notons l'importance de différencier fondamentalement les notions de fréquence et de probabilité.
- Troisièmement, en statistiques fréquentistes, la "probabilisation" se base uniquement sur les observations, en fonction de valeurs hypothétiques des paramètres ; il est d'ailleurs habituel de nommer celles-ci, valeurs estimées des paramètres. Les probabilités sont estimées à partir de fréquences qui peuvent se vérifier à long terme, c'est-à-dire lorsqu'on tend asymptotiquement vers un très grand nombre d'observations. En revanche, en statistiques Bayésiennes, la "probabilisation" se base aussi bien sur les observations que sur les paramètres qui sont exprimés sous forme de distributions. Appliquées à des régressions linéaires multiples, les statistiques Bayésiennes permettent d'estimer l'incertitude des modèles, c'est-à-dire la distribution des paramètres de ces

derniers. Les prédictions réalisées au sein du DS peuvent dès lors être considérées, à juste titre, comme des probabilités prédictives.

Enfin, les DS exposés dans ce travail offrent dans certains cas des niveaux de qualité (π) relativement faibles (par exemple π = 40% à la section III.1.3.*B*). Cependant, les expériences réalisées au sein du DS ont permis d'obtenir des séparations tout à fait acceptables et en bon accord avec les prédictions. Il conviendrait dès lors d'évaluer précisément l'impact du niveau de qualité obtenu sur la crédibilité des DS et également l'impact des données d'entrées, et de leur variabilité, sur le niveau de qualité. En utilisant les statistiques fréquentistes, comme ce fut le cas jusqu'à présent, l'estimation de la relation existant entre le niveau de qualité et la crédibilité du DS, ne pourrait être envisagée qu'après avoir répété la méthodologie DoE-ICA-DS aux mêmes échantillons et ce pour un plus grand nombre de mélanges que celui présenté dans ce travail. Les statistiques Bayésiennes pourraient à nouveau être utilisées pour estimer de manière plus précise la crédibilité donnée à un DS en fonction du niveau de qualité.

Chapitre V

Résumé de la thèse

L'objectif principal de ce travail était la mise au point d'une méthodologie innovante pour le développement de méthodes chromatographiques. Celle-ci a été pensée afin de répondre aux récents concepts de "*quality by design*" (QbD) et de "*design space*" (DS) préconisé par la Conférence internationale d'harmonisation (ICH). Cette méthodologie repose sur la modélisation des temps de rétention au début, au sommet et à la fin de chaque pic chromatographique des composés d'un mélange en utilisant le critère de séparation (*S*) plutôt que la résolution (R_S). A cette fin, des régressions linéaires multiples en mode pas à pas ont été utilisées.

Dans le cadre de cette modélisation, l'erreur estimée a ensuite été propagée au moyen d'une simulation de Monte Carlo afin d'estimer la distribution du critère de qualité prédéfini (*S*) et de calculer le DS afférent. Ce dernier peut donc être défini comme étant le sous-espace du domaine expérimental à l'intérieur duquel la probabilité que le critère de séparation soit supérieur à un seuil déterminé, soit plus grande qu'un niveau de qualité. En termes chromatographiques, le DS correspond donc à une zone où, en moyenne et avec un niveau de qualité donné, la séparation de l'ensemble des composés étudiés ne montre pas de phénomènes de coélution.

Parallèlement à ces travaux, nous avons démontré l'intérêt de l'analyse en composantes indépendantes (ACI) pour la lecture automatique des chromatogrammes. Ce traitement du signal a rendu possible la séparation numérique de composés coélués et l'estimation avec une assez grande fiabilité des temps au début et à la fin des pics chromatographiques correspondants.

La combinaison de cet algorithme numérique à la stratégie d'optimisation des séparations décrite précédemment nous a permis de proposer une méthodologie tout à fait originale ouvrant la voie vers le développement automatisé de méthodes chromatographiques.

Afin de valider la méthodologie globale, elle a d'abord été testée à la fois sur un mélange de composés inconnus et sur une formulation pharmaceutique. Elle a montré sa capacité à accélérer significativement le processus de développement d'une méthode chromatographique et à permettre l'obtention de séparations optimales. De plus, cet exercice nous a permis de démontrer formellement que le DS correspondait bien à une zone de robustesse.

Par la suite, nous nous sommes attachés à montrer la bonne concordance existant entre les modèles et la théorie de la chromatographie. Signalons qu'à cette occasion, une méthode de criblage de 19 antipaludéens a également été optimisée.

Enfin, la méthodologie a notamment été appliquée avec succès à l'optimisation de la séparation d'un mélange de cannabinoïdes, d'une part, et à celle du sulindac et de trois de ses impuretés, d'autre part.

En conclusion, la méthodologie globale développée a démontré ses aptitudes à détecter automatiquement et à séparer numériquement les pics chromatographiques de composés coélués, à optimiser la séparation de nombreux mélanges et à identifier le DS correspondant à une zone robuste, démontrant par là même sa puissance et sa flexibilité.

The main objective of the present work was to develop an innovative methodology for the development of chromatographic methods. This methodology was created in order to be compliant with the concepts of quality by design (QbD) and design space (DS) recommended by the International Conference on Harmonisation (ICH). This latter methodology is based on the modelling of the retention times at the beginning, the apex and the end of each peak corresponding to the compounds of a mixture and by using the separation criterion (*S*) rather than the resolution (R_S) as a critical quality attribute. Stepwise multiple linear regressions were used to create the models.

In the framework of these modelling, the estimated error was propagated from the modelled responses to the separation criterion. This was performed by using Monte Carlo simulations in order to estimate the distribution of the separation criterion and to compute the DS. This latter can thus be defined as the subspace, necessarily encompassed in the experimental domain (i.e. the knowledge space), within which the probability for the criterion to be higher than an advisedly selected threshold is higher than a minimum quality level. From a chromatographic point of view, the design space thus defines a region where, on average, with respect to a given quality level, peaks do not coelute.

Concurrently to these works, we demonstrated the usefulness of independent component analysis (ICA) to automatically read chromatograms. This statistical technique allowed to numerically separate coeluting compounds in order to estimate the retention times at their respective beginnings and ends.

The combination of the ICA-based methodology and the optimization one allowed proposing an innovative global methodology paving the way towards the automated development of chromatographic methods.

In order to validate this global methodology, a mixture of unknown compounds and a pharmaceutical formulation were first used. This study confirmed its ability to significantly speed up the development process and allowed obtaining optimal separations. It was also demonstrated that DS actually defines robust spaces.

Subsequently, mathematical equations resulting from retention times modelling were interpreted from a chromatographic point of view. During this other study, a method for the screening of 19 antimalarial drugs was also optimized.

Eventually, the methodology was successfully applied for the separation optimization of a mixture of cannabinoids and a mixture containing sulindac and 3 of its related impurities.

In conclusion, the global methodology developed during this work, demonstrated its abilities to automatically read chromatograms and to numerically separate coeluting peaks. It also confirmed its aptitudes to optimize the separation of numerous mixtures and to identify the DS corresponding to a zone of robustness. The proposed methodology finally proved its efficiency as well as its flexibility.

Chapitre VI

Liste des publications

1. E. Ziemons, G. Dive, B. Debrus, V. Barillaro, M. Frederich, R. Lejeune, L. Angenot, L. Delattre, L. Thunus, Ph. Hubert., Study of the physicochemical properties in aqueous medium and molecular modeling of tagitinin C/cyclodextrin complexes, J. Pharm. Biomed. Anal. 43 (**2007**) 910.

2. P. Lebrun, B. Govaerts, B. Debrus, A. Ceccato, G. Caliaro, Ph. Hubert, B. Boulanger, Development of a new predictive modelling technique to find with confidence equivalence zone and design space of chromatographic analytical methods, Chemom. Intell. Lab. Syst. 91 (**2008**) 4.

3. B. Debrus, P. Lebrun, A. Ceccato, G. Caliaro, B. Govaerts, B. A. Olsen, E. Rozet, B. Boulanger, Ph. Hubert, A new statistical method for the automated detection of peaks in UV-DAD chromatograms of a sample mixture, Talanta 79 (**2009**) 77.

4. B. Debrus, P. Lebrun, E. Rozet, I. Nistor, A. Ceccato, G. Caliaro, R. Oprean, B. Boulanger, Ph. Hubert, Nouvelle méthodologie pour le développement automatisé de méthodes analytiques en chromatographie liquide pour l'analyse de mélanges de composés inconnus, Spectra Analyse 268 (**2009**) 28.

5. G. Martin, F. Mansion, A.-C. Servais, B. Debrus, E. Rozet, Ph. Hubert, J. Crommen, M. Fillet, CE-MS method development for peptides analysis, especially hepcidin, an iron metabolism marker, Electrophoresis 30 (**2009**) 2624.

6. B. De Backer, B. Debrus, P. Lebrun, L. Theunis, N. Dubois, L. Decock, A. Verstraete, Ph. Hubert, C. Charlier, Innovative development and validation of an HPLC/DAD method for the qualitative and quantitative determination of major cannabinoids in cannabis plant material, J. Chromatogr. B 877 (**2009**) 4115.

7. N. Dubois, B. Debrus, Ph. Hubert, C. Charlier, Validated quantitative simultaneous determination of cocaine opiates and amphetamines in serum by UHPLC coupled to tandem mass spectrometry, Acta Clin. Belg. 65 *Supplement* 1 (**2010**) 75.

8. R. Marini Djang'Eing'A, J. Mbinze Kindenge, M. L. A. Montes, B. Debrus, P. Lebrun, J. Mantanus, E. Ziemons, C. Rohrbasser, S. Rudaz, Ph. Hubert, Analytical tools to fight against counterfeit medicines, Chim. Oggi 28 (**2010**) 10.

9. F. Krier, M. Brion, B. Debrus, P. Lebrun, A. Driesen, E. Ziemons, B. Evrard, Ph. Hubert, Optimisation and validation of a fast HPLC method for the quantification of sulindac and its related impurities." J. Pharm. Biomed. Anal. 54 (**2011**) 694.

10. B. Debrus, J. Broséus, D. Guillarme, P. Lebrun, Ph. Hubert, J.-L. Veuthey, P. Esseiva, S. Rudaz, Innovative methodology to transfer conventional GC-MS heroin profiling to UHPLC-MS/MS, Anal. Bioanal. Chem. 399 (**2011**) 2719.

11. B. Debrus, P. Lebrun, A. Ceccato, G. Caliaro, E. Rozet, I. Nistor, R. Oprean, F.J. Rupérez, C. Barbas, B. Boulanger, Ph. Hubert, Application of new methodologies based on design of experiments, independent component analysis and design space for robust optimization in liquid chromatography, Anal. Chim. Acta *In press* (**2011**). doi:10.1016/j.aca.2011.02.035

12. B. Debrus, P. Lebrun, J. Mbinze Kindenge, F. Lecomte, A. Ceccato, G. Caliaro, J. Mavar, R. Marini , B. Boulanger, E. Rozet, Ph. Hubert, Innovative HPLC method development for the screening of 19 antimalarial drugs based on generic approach, using design of experiments and design space, J. Chromatogr. A *Submitted for publication* (**2011**).

13. I. Nistor, M. Cao, B. Debrus, P. Lebrun, F. Lecomte, E. Rozet, L. Angenot, M. Frederich, R. Oprean, Ph. Hubert, Application of a new optimization strategy for the separation of tertiary alkaloids extracted from *Strychnos usambarensis* leaves, J. Pharm. Biomed. Anal. *Submitted for publication* (**2011**).

14. H. Zimé Diawara, B. Debrus, F. Gbaguidi, P. Lebrun, R. Semdé, M. Moudachirou, B. Evrard, J. Quetin-Leclercq, Ph. Hubert, Use of experimental design to obtain rich artemisinin aqueous and hydro-alcoholic extracts of *Artemisia annua* for the development of standardized treatments, J. Ethnopharmacol. *Submitted for publication* (**2011**).

15. M.H. Rafamantanana, B. Debrus, G.E. Raoelisonb, S. Uverg-Ratsimamanga,, E. Rozet, Ph. Hubert, J. Quetin-Leclercq, Application of design of experiments for the HPLC-UV-MS separation of aporphine alkaloids from leaves of S*pirospermum penduliflorum* thouars, J. Pharm. Biomed. Anal. *Submitted for publication* (**2011**).