# UNIVERSIDAD DE BURGOS

## CHEMISTRY DEPARTMENT – ANALYTICAL CHEMISTRY AREA


# DOCTORAL THESIS

## ENSURING ANALYTICAL QUALITY IN REGULATED MARKETS THROUGH MULTIVARIATE, MULTIWAY AND DOE STRATEGIES


**MARÍA LETICIA OCA CASADO**

**DIRECTORS: MARÍA DE LA CRUZ ORTIZ FERNÁNDEZ**

**LUIS ANTONIO SARABIA PEINADOR**



**BURGOS, NOVEMBER 2021**

*A mis padres y mi hermana.*
*Por todos los duros momentos que han quedado grabados en nuestras neuronas, aunque no podamos o no queramos acceder a ellos.*

*Empieza haciendo* **lo necesario,**
*después* **lo posible,**
*y de repente te encontrarás haciendo* **lo imposible.**

*(San Francisco de Asís)*

## AGRADECIMIENTOS

Antes de iniciar la exposición escrita de esta tesis doctoral, me gustaría agradecer el apoyo financiero de las siguientes instituciones, sin el cual la realización de este trabajo no habría sido posible:

- Universidad de Burgos, por la concesión de la beca predoctoral enmarcada en la convocatoria del año 2011 de las Ayudas para la Formación de Personal Investigador de la UBU (Resolución del 03 de noviembre de 2011).

- Ministerio de Ciencia e Innovación, por la financiación otorgada a través del proyecto de investigación CTQ2011-26022.

- Ministerio de Economía y Competitividad, por la financiación otorgada a través del proyecto de investigación CTQ2014-53157-R.

- Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación (AEI) y Fondo Europeo de Desarrollo Regional de la Unión Europea (FEDER, UE), por la financiación otorgada a través del proyecto de investigación CTQ2017-88894-R, cofinanciado con fondos europeos FEDER.

- Junta de Castilla y León, por la financiación otorgada a través de los proyectos de investigación BU108A11-2 y BU012P17, cofinanciados con fondos europeos FEDER.

Si hay algo que debo reconocer al mirar atrás desde aquí arriba es que, a diferencia de la caída, el ascenso ha sido difícil. En estas situaciones suele creerse que nunca llegará el momento del "por fin", pero, tarde o temprano, todo termina pasando una vez que se encuentra el camino de vuelta. Y ahora que sale nuevamente el sol y voy cerrando etapas, le ha llegado el turno a cuatro años de aprendizaje, experimentación (con sus inseparables baches, los consiguientes quebraderos de cabeza y los sufridos *eurekas* finales), peleas incesantes con los equipos de laboratorio para descubrir que la inteligencia artificial existe desde hace mucho más tiempo del que creemos, artículos y pósters mil veces rehechos, presentaciones en congresos y *workshops*, discusiones (amistosas) sobre lo sagrado y lo profano, reuniones y consejos de representación estudiantil, cafés "electrónicos" en familia y, sobre todo, alegría e ilusión que abrían cada día la puerta del laboratorio. Este cóctel perfecto no lo puede preparar una sola: hacen falta manos que acompañen a corazones de calidad, así que, en la hora de poner la rúbrica a esa etapa con una tesis que tiene justamente como leitmotiv la búsqueda de la calidad, no puedo por menos que agradecer sinceramente a todos los que me han servido de inspiración en esa cruzada:

- A **Cruz** y **Luis**, por sus enseñanzas, su pasión (contagiosa) por ir más allá, su guía, su ayuda cuando todo parecía perdido (para encontrar oráculos no hace falta remontarse a la Antigua Grecia) y, sobre todo, su enorme paciencia en estos últimos años de espera: aunque parezca increíble, ya estamos viendo los tres la luz al final del túnel.

- A **Ana**, **Sagrario**, **Silvia** y **Celia**, miembros del grupo de investigación *Quimiometría y Cualimetría* de la Universidad de Burgos, con las que compartí espacio y pensamientos en voz alta en el laboratorio (impensable en los tiempos que corren el estar todas juntitas allí) o detrás del ordenador: millones de gracias, Sagrario, por solucionarnos la vida con tus consejos, tus archivos *.mat* y esa ironía tuya tan chispeante con la que me siento tan identificada.

- Al **área de Química Analítica** y muy especialmente, a **Susana** y a **Marisol**, por su vitalidad incansable que despejaba los nubarrones en un día gris y su ayuda de valor incalculable en lo que fuese y para lo que fuese: no cambiéis nunca. ¿Qué habríamos hecho todos los que hemos pasado por esa área si tú no hubieras estado ahí, Marisol?

- A **David**, por iniciarnos a todas en el apasionante viaje de la cromatografía teórica a la cromatografía práctica, y a **Rocío**, la sevillana con acento burgalés tan especial con la que compartí buenos y malos momentos durante los primeros años de esta aventura.

- A **Laura R.**, por TANTO durante tanto. Más que una compañera de viaje durante la licenciatura, el máster y el doctorado. Más que una amiga única dentro y fuera del recinto universitario. No existe una palabra en la lengua castellana que pueda describir

lo que eres ni una que describa lo que significas para mí. GRACIAS por todo y PERDONA por alejarme en los últimos tiempos, pero hay heridas que hay que lamerse en soledad y proteger así a los que se quiere.

- A todos los **"electrones"**: **Laura R.**, **Borja**, **Cris**, **Erica**, **Sandra**, **Laura H.** y **Edna**, por tantísimas cosas vividas juntos entre los laboratorios de Química Analítica y Química Orgánica, episodios de llanto incluidos. Estáis todos en mi orbital *1s*, diga lo que diga la Química Cuántica: en el mío caben más de dos, algunos de ellos bien apareados desde hace tiempo.

- A los **técnicos de *Agilent***, **Felipe** y **Ángel**, por ser nuestros hombros para llorar cuando las máquinas se rebelaban, pero sobre todo, por devolverlas al redil en un periquete. mi hipótesis de que estos seres (porque lo son: tienen vida propia) reconocen la mano del amo quedó perfectamente con vosotros.

- A toda la gente que me rodeó en las distintas estancias predoctorales en la **Estación Tecnológica de la Leche**, en el **Laboratorio de Salud Pública de Burgos** y en el *Centre Wallon de Recherches Agronomiques*: muchísimas gracias por vuestro tiempo, vuestra dedicación y, en particular, vuestro buen humor hispano-belga, **Juan Antonio** y **Damien**: gracias a vuestra manera de ser, aprendí a adorar a ese pequeño país que hace saltar en mi interior las ganas de volar allí cada vez que me acuerdo del verano de 2015 o veo una foto de mi añorada Bruselas.

- A mis compañeros y amigos de **GSK Aranda**, especialmente al Área Técnica (agradecimiento extensivo tanto a los que siguen ahí como a los que no), por su apoyo y aliento en la distancia durante la fase de escritura. Mención especial para **Luis M.**, por el siete que le supuso mi decisión de parar para poder cerrar: nunca he otorgado reconocimientos en balde, y tú te mereces más de los que te he dado hasta ahora.

- A **Luisber**, por su compañía y aliento durante los años de experimentación, especialmente, en las eternas tardes de escritura y revisión que terminaban convirtiéndose en noches. *Sit tibi terra levis*, aunque yo siempre te llevaré en mi corazón.

- A **Juanjo**, por su infinita paciencia cuando, sin previo aviso, los planes tenían que ser aplazados o cancelados para poder terminar de dar a luz a esta tesis. Eres mi loco cuerdo favorito.

- Y, para concluir como corresponde, esto es, a lo grande, a **mis padres**, sin cuyo apoyo no habría podido llegar hasta aquí. GRACIAS: la mente es maravillosa, pero también puede ser muy cruel.

# INDEX

# CHAPTER 1

## CONTEXT, LEGISLATIVE FRAMEWORK AND OBJECTIVES OF THIS DOCTORAL THESIS

## 1.1. Context of this doctoral thesis

The doctoral thesis set out below centres around the vast concept of Food Quality and how to better ensure it by presenting analytical methodologies that mean a clear improvement in terms of Analytical Quality beside most of the current approaches typically applied in such a regulated field as the food sector.

Controlling and improving quality has become an important business strategy for many organizations, since quality is a competitive factor that can make a business either dominate its competitors or succumb to them.

Most people have a conceptual understanding of quality as relating to one or more desirable characteristics that a product or service should possess to meet the requirements of those who use it or even linked to the absence of a defect, fraud and adulteration in that product or service. Although this conceptual understanding is certainly a useful starting point, this historical view is extremely limited and could even lead to erroneous conclusions on the product or service being assessed. Nowadays, the quality of a product is described and evaluated in several ways, the ***dimensions of quality*** [1].

As regards food products, that multivariate decomposition stays, so the notion of food quality rests on a complex and multi-dimensional concept which is influenced by a wide range of situational and contextual factors [2], as it is depicted in Figure 1.1.


**Figure 1.1**. Range of dimensions that make up the holistic concept of Food Quality (from *https://knowledge4policy.ec.europa.eu/food-fraud-quality/topic/food-quality_en*).

In order to ensure the desired quality level of any food product, first of all, a prospective summary of properties or characteristics that ideally will be achieved and that are linked to one or several of those dimensions must be set. Some of those characteristics are

measurable, so a method for the analysis and determination of the value of at least each one of these characteristics for every product tested must be developed and validated, where, according to the U.S. Food and Drug Administration (FDA), *validation* is "the collection and evaluation of data, from the process design stage through commercial production, which establishes scientific evidence that a process is capable of consistently delivering quality products" [3].

Among the characteristics to be determined for every product, those physical, chemical, biological, or microbiological that should be within an appropriate limit, range, or distribution to attain product quality are named ***Critical Quality Attributes*** (CQAs) [4]. So, in this context, the definition of quality given by both the International Council for Harmonisation (ICH) of Technical Requirements for Pharmaceuticals for Human Use and the FDA as "the degree to which a set of inherent properties of a product, system, or process fulfils requirements" matches perfectly [3,7].

Most of those quality attributes, either CQAs or not, are typically determined on the final product; however, just testing quality in the end by following the traditional retrospective Quality by Testing (QbT) path is most unlikely to be capable of supporting quality, even if the testing level is significantly high: in this manufacturing scenario, at the moment of testing, products have already been made; therefore, if any defect is detected, it cannot be easily rectified, and so a batch may need to be reworked or even rejected [5], which does not agree with the expectations of the regulatory authorities for a validated process [3]. On the contrary, a prospective science- and risk-based approach to develop process understanding known as ***Quality by Design*** (QbD) is a fairly better option to be chosen for quality assurance. With respect to pharmaceutical development, the term "*Quality by Design*" was first defined in ICH guideline Q8(R2) "*Pharmaceutical Development*" [4] in 2009 as "a systematic approach to development that begins with predefined objectives, emphasizes product, process understanding and process control, based on sound science and quality risk management". This was a watershed for industry: from that moment, the key principle turned out to be that quality should be built into a product with a thorough understanding of both the product and the manufacturing process. This included establishing a knowledge of the risks involved in manufacturing the product and how best to mitigate those risks. This new paradigm would result in better control of parameters and variables, and reduce the emphasis on end-product testing [5]. To do so, QbD involves many quality and statistical tools and methods (e.g., Design of Experiments (DoE), multivariate statistics, Six-Sigma methodologies, and statistical quality control), since it requires understanding of how variables involved in formulation and manufacturing processes influence the quality of the final product [6].

However, although QbD has been widely adopted in the global industry, especially in the pharmaceutical world through a number of initiatives and regulatory documents [4,7,8,9,10,11,12], it can be applied far deeper. If a close look into the analytical environment is taken, it can be easily concluded that analytical procedures are also processes, so QbD could and should be implemented for the development of analytical methods: they are non-negligible components of any manufacturing process, since most of the quality attributes of a product are eventually determined from them. Not for nothing have there been many discussions [6,13,14] on the opportunity for analytical method developments to follow a similar approach to the QbD path for pharmaceutical processes as laid down in the ICH guidelines [4,7,8,9,10,15]: as part of the assessment of the quality of a product, the outcome from the testing procedures followed for the determination of its quality attributes must be trustworthy. That is to say, product quality assurance unfailingly needs analytical quality assurance to come first, a belief that has been the main aim to pursue throughout the entire carrying out of this thesis.

A key component of the development of the QbD paradigm is what is known as the ***Design Space***, which is defined in [4] 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". In a pharmaceutical context, a design space can be described in terms of ranges of material attributes and process parameters, or through more complex mathematical relationships, e.g. as a time-dependent function, or as a combination of variables such as components of a multivariate model [4]. The design space is thus a subspace that is necessarily encompassed within the experimental domain, which is the multidimensional space formed by the factor ranges used during the development stage. Bearing in mind the two main types of causes affecting the variability of a process (namely, natural *vs* special causes of variability), the design space can be considered as a zone of theoretical robustness as only natural causes of variability occur, so no drastic changes in the levels of the CQAs of the method should be observed [6].

To develop a QbD-compliant analytical method and finally reach the definition of its analytical design space (or *Method Operable Design Region* (MODR) so as to distinguish it from the process design space [6,16,17,18]), the intended purpose of the analytical method must be defined, that is to say, the scenario that depicts what will be measured, in which matrix, over which concentration range, and the CQAs of the method, together with their specifications. All this is known as the ***Analytical Target Profile***, similar to the Quality Target Product Profile in pharmaceutical development [6]. Apart from the CQAs, the method parameters that could theoretically impact on the analytical technique need to be set through a risk assessment.

- As for the former, the required performance criteria, that is, the responses that are measured to judge the quality of the analytical method (e.g., the resolution criteria, the run time of the analysis, the precision of the method, the lower limit of quantification or the dosing range of the method), together with their specifications, can be considered as the CQAs of the analytical procedure [6].

- A wisely selected number of high-risk instrumental or operating parameters (e.g., gradient time and/or pH of the mobile phase in chromatography and concentration of reagents in immunoassays) that impact on the analytical technique under development are the ***Critical Process Parameters*** (CPPs) of the analytical method, and they are usually obtained from a prior risk analysis exercise and a prioritization strategy.

Then, to define the analytical design space, the factors identified as CPPs have to be studied simultaneously. The best choice to do that is through the DoE methodology, whose usefulness in the QbD approach applied to several chromatographic procedures can be seen in [18,19,20]. The juxtaposition of univariate studies of the involved factors cannot determine an analytical design space properly, as potential interactions of the factors involved cannot be studied; these interactions may result in synergic or antagonistic effects on the responses measured and therefore modify the delimitation of the design space. This approach eventually leads to a multivariate domain of input factors ensuring that critically chosen responses lie within predefined limits with an acceptable level of probability [13]. The analytical method under development is therefore no longer defined by a single point in the space of its operating parameters (e.g., one value of wavelength, one value of proportion of organic modifier, or one value of pH of the mobile phase), but by a range of operating conditions that are defined by the analytical design space, where the analytical method provides quality outputs with adequate probability [6,18].

It is clear from this exposition on the QbD paradigm that one of the strengths of this approach in order to succeed in attaining robust procedures over time comes from **viewing the world as multivariate instead of univariate**: while studying variance is useful, understanding covariance can lead one to move rapidly from data to information to knowledge [21], which represents the core of the scientific foundations of the QbD methodology. As a way of example, the traditional approach to develop an analytical procedure in liquid chromatography (LC) is a trail-and-error approach, such as varying one-factor-at-a-time (OFAT). This leads to a large number of experiments and fails to analyze the interaction between factors. LC analytical procedures developed by this approach often result in a non-robust performance [22]. Furthermore, the lack of guaranteed robust and high-quality analytical methods has always been a serious concern of regulatory agencies, to such an extent that the ICH has stepped in and the upcoming ICH Q14 *"Analytical Procedure*

*Development and Revision of Q2(R1) "Analytical Validation""* will address this issue [15]. Similar to Q8(R2) and Q11, Q14 will provide a lifecycle management framework around analytical procedures and encourage the use of enhanced approaches, such as QbD, instead of traditional approaches [14].

However, one step forward can still be taken and thus move beyond multivariate analysis to **multi-way (or multi-mode) data analysis**. This was first accomplished in psychometric studies involving multiple subjects (people) given several tests over periods of time, which lead to data structures that fit into three-dimensional computer arrays represented by blocks of data or three-way arrays. Of course each subject's data (tests × time) could be analyzed separately by removing a matrix or slice from a block of data, but this approach would lose the covariance among subjects, so data models and methodologies to analyze entire blocks of data or three-way arrays at one time have been developed [21]. Most of these methods have a strongly exploratory character, which means that complex patterns of dependencies among the elements of the three ways can be found, without postulating specific configurations *a priori* and without applying tests to these patterns. This is partly because it is difficult to specify such patterns beforehand, and partly because hypothesis testing supposes that something is known about the distributions of the scores, which for multi-way data is only very rarely the case [23].

Among multi-way data, the easiest type for its structure to be visualized is the three-way one. To do so, let's start with two-way data and then extend the concepts to three-way. Two-way data are represented in a two-way data matrix typically with columns as variables and rows as objects. Three-way data are represented as a three-way array (box) where the columns and rows of the previous two-way array have now turned into slices. Each horizontal slice holds the data for one object in a three-way array, as can be seen in Figure 1.2.a. Each vertical slice holds data of one specific variable of one type (for example, in GC-MS data, counts at different $m/z$ diagnostic ions or retention times) and the back-to-front slices the variables of the other type [21]. As a consequence, for a three-way array of GC-MS data, the samples are the objects whereas the counts at different retention times (of the chromatography mode) corresponding at different $m/z$ ratios (of the mass spectrometry mode) are the two variable modes. Hence, ion counts are measured as a function of two properties: retention time and $m/z$ ratios, as depicted in Figures 1.2.a and 1.2.b. Other types of arrangements can exist as well, since some three-way arrays have two types of object modes and only one type of variable mode, or even three objects modes and no variable mode. An example is in multivariate image analysis where the data are three-way arrays with $x$-coordinate and $y$-coordinate of the pixels as object ways and wavelength as the variable way.

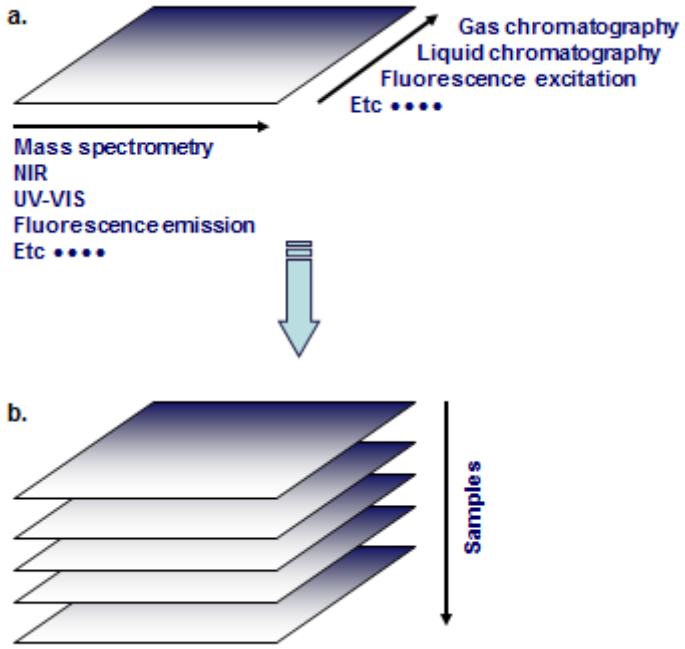


**Figure 1.2.** Example of arrangement of data in a three-way tensor. **a.** Two-way matrix (slice) for one sample. **b.** Stacking of all the sample matrices to build the final tensor.

If a block (tensor or box) of three-way data is available, exploring the interrelations in those data is possible [21]. An example is a three-way environmental data set consisting of several physicochemical analyses on several locations in a geographical area at several points in time, so a three-way analysis can help in distinguishing patterns, e.g., temporal and spatial behavior of the different chemical compounds [24]. Another application area is image analysis, where two-dimensional images of an object taken at different wavelengths are stacked on top of each other and analyzed with three-way tools [25]. Nevertheless, one of the most interesting applications of three-way analysis on a single block of data in Analytical Chemistry and Quality Assurance is in second-order calibration [21], where an instrument (typically, a hyphenated system like those required in confirmatory methods, e.g., in a GC-MS analysis) is used, so a matrix of measurements (a second-order array) for a single chemical analysis is generated. If the standard $\mathbf{X}_1$ contains the pure response of the measured analyte and $\mathbf{X}_2$ is the measurement of the mixture containing that analyte, then under certain conditions it is possible to quantify the analyte in the mixture, even if this mixture contains unknown interferents. This is done by stacking $\mathbf{X}_1$ and $\mathbf{X}_2$ on top of each other and building a three-way model of that stacked array. Not only the concentration estimate of the analyte is obtained but also those of the pure response profiles, e.g., pure mass spectra and chromatograms of the analyte [26,27,28,29].

With data coming from hyphenated instruments, low-rank (that is, one component per analyte) bilinearity of a single-sample measurement can often be assumed. At this point, it must be stated that the terms "*bilinearity*" or "*trilinearity*" are used to mean "*low-rank bilinearity*" or "*low-rank trilinearity*", respectively, as the former in themselves are meaningless from a data analysis point of view [21]. So, if the underlying model is (low-rank) bilinear/trilinear, the measured data are approximately (low-rank) bilinear/trilinear, and a (low-rank) bilinear/trilinear model can be meaningfully fitted in the ideal case. For example, for a GC-MS measurement, a hypothesized model of the data $\mathbf{x_k}$ from one sample (let's name it *k*) can then be written as

$$\mathbf{x_k} = c_1\mathbf{a_1}\mathbf{b_1}' + c_2\mathbf{a_2}\mathbf{b_2}' + \ldots + \mathbf{E} \tag{1.1}$$

where $\mathbf{a_1}$, $\mathbf{a_2}$, $\mathbf{b_1}$, $\mathbf{b_2}$ etc. ($\mathbf{a_r}$ = unit chromatograms of the compounds present in the sample, $\mathbf{b_r}$ = unit spectra of the compounds present in the sample, $c_r$ = concentrations of the compounds present in the sample) follow from the underlying model. So when the underlying model is low-rank bilinear, the data matrices produced are often close to this bilinear model, except for noise and nonlinearities.

With a stack of low-rank bilinear data matrices such as $\mathbf{x_k}$, a three-way array $\underline{\mathbf{X}}$ can be built. In this case, modeling this data amounts to estimate the following hypothesized expression:

$$\underline{\mathbf{X}} = x_{ijk} = \sum_{f=1}^{F} a_{if}\, b_{jf}\, c_{kf} + e_{ijk} \tag{1.2}$$

where the three loading matrices **A** (size $I \times F$, containing all $a_{if}$ values), **B** (size $J \times F$, containing all $b_{jf}$ values), and **C** (size $K \times F$, containing all $c_{kf}$ values) can be reliably obtained through the decomposition of $\underline{\mathbf{X}}$ by minimizing the error $e_{ijk}$.

The underlying model of $\underline{\mathbf{X}}$ is not necessarily low-rank trilinear if, for example, there are retention time shifts from sample to sample. However, the data can still be fitted by a low-rank trilinear model. If the deviations from the ideal trilinearity (low-rank) are small, directly interpretable components can be found. If not, components will still be found in the same sense as in PCA, where the components together describe the data, although no single component can necessarily be related to a single analyte [21]. For example, trilinearity in the underlying model can be assumed when data from a set of samples within a concentration range are obtained from a gas chromatograph coupled to a mass spectrometer. In this case, the data produced for one sample are bilinear, the underlying model is often low-rank trilinear, and the three-way technique separates $\underline{\mathbf{X}}$ into unit chromatograms, unit mass spectra and unit concentration profile as **A**-, **B**- and **C**-loadings.

One of such methods for three-way data is ***Parallel Factor Analysis (PARAFAC)***. It was developed in 1970 by Harshman [30] and, independently in that same year, by Carrol and Chang [31], following the concept of proportional parallel profiles by Cattell [32] to obtain an unique decomposition of a three-way data tensor to find the values of $a_{if}$, $b_{jf}$, and $c_{kf}$ in Eq. (1.2) by minimizing the sum of the squared errors (*SSE*):

$$SSE = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} e_{ijk}^2 \tag{1.3}$$

As in other least-squares procedures, PARAFAC starts from initial estimations of $a_{if}$, $b_{jf}$ and $c_{kf}$ and then refines these initial values iteratively until *SSE* is minimum. Since there are many parameters to retrieve ($I \times F$ values of $a_{if}$, $J \times F$ values of $b_{jf}$, and $K \times F$ values of $c_{kf}$), they are not all varied simultaneously in each iterative cycle of the algorithm, but in an alternating manner: in the first cycle only $a_{if}$ are allowed to vary (the remaining parameters are kept fixed); in the next one only $b_{jf}$ are varied; just $c_{kf}$ can change in the third try, and so on. This is why the entire procedure is known as alternating least-squares (ALS) [33]. The iterations will take place until convergence of fit, which does not necessarily imply convergence of parameters, but in practical situations this is usually the case. Therefore, the stopping criterion for the algorithm is usually a small relative change in the fit of the model [21].

A major practical obstacle in the use of the PARAFAC model is how to determine the appropriate number of components, that is, the number $F$ of factors in Eq. (1.2). Fitting a single model can be time-consuming, especially because refitting from different starting points is usually essential for assuring convergence to the global minimum. Several procedures have been devised [34], where the core consistency diagnostic (CORCONDIA), a certain Tucker3 core for assessing the appropriateness of a PARAFAC model [34], has proved to be specifically useful for this task.

On the other hand, one extremely powerful feature of PARAFAC is that the solutions yielded by this algorithm are often unique [35]. This means that a single solution exists for all parameters $a_{if}$, $b_{jf}$ and $c_{kf}$ when decomposing a low-rank trilinear array with this algorithm through Eq. (1.2), no matter which the initial estimations from which the PARAFAC-ALS procedure starts are. This uniqueness property has important consequences in analytical calibration from second-order instruments, because it allows the analytes present in a sample to be accurately quantified even if that sample contains unexpected constituents or interferents. The link between the PARAFAC uniqueness and the main advantage provided by second-order instrumentation (known as *second-order advantage*) can be stated as follows: if there is a single unique solution to Eq. (1.2), then the retrieved profiles **a** and **b** should be the true constituent profiles describing the chemical phenomena along each of the

corresponding data modes, while the **c** profile describes the concentration changes across the samples in the array [33].

- The values of $c_{kf}$ are directly proportional to the concentration of each pure constituent in each sample, particularly of the analyte of interest in the calibration and in the test samples. This fact turns a multivariate calibration into a virtual univariate (often called pseudo-univariate) one [33] to which the classic Inferential Statistics can be fully applied to complete the analytical method validation.

- By retrieving **a** and **b**, PARAFAC shows its ability to extract qualitative physicochemical information (spectra, chromatograms, time evolutions, pH profiles, etc.) on the behaviour of the sample constituents, as if they had been physically separated from the rest of the mixture. This digital separation of the constituents is sometimes named *mathematical chromatography* or *virtual chromatography* [33], and clearly strengthens the way to fulfil the legislative requirements related to unequivocal identification later detailed in Section 1.2 even if interferents are present in the sample, as it will be set out in the following chapters and has already been in [26,27,28,29].

In order to profit from the second-order advantage, the analytical instrument has to operate in a very reproducible way, which is not always the case. For example, one of the problems in using a hyphenated technique such as chromatography–spectroscopy is the poor reproducibility of the retention time axis: shifts in the chromatographic mode often occur from sample to sample; this destroys the trilinearity of the calibration model [21]. If these deviations are small, this *a priori* drawback can be counteracted by using different types of three-way models such as PARAFAC2 [36].

A further detailed explanation of the features of PARAFAC and other multi-way strategies and the advantages achieved from their use on second-order signals for the quantitative determination of substances can be found in [26]. This doctoral thesis, with the research work set out in the chapters to follow, is intended to display those advantages and thus reveal the goodness of that approach coupled to quality assurance from a QbD perspective by applying it to real problems framed in the Food Quality sector but, fundamentally, under the banner of Analytical Quality.

## 1.2. Legislative framework of this doctoral thesis

On the contextual basis just expounded in Section 1.1, the European Union (EU) food quality and safety policies, which are aimed at ensuring that European citizens have access to safe and wholesome food of highest standards, have been observed during the performance of the research works detailed in the following chapters. So, in this sense, the analytical and chemometric methodologies derived from this research are in full compliance with the EU food legislation in force at the moment of their development.

### 1.2.1. Food Quality from a Food Origin perspective

EU quality policy aims to protect the names of specific products to promote their unique characteristics, linked to their geographical origin as well as traditional know-how.

Products can be granted a "*geographical indication*" (GI) if they have a specific link to the place where they are made. The GI recognition enables consumers to trust and distinguish quality products while also helping producers to market their products better [37].

Geographical indications establish intellectual property rights for specific products, whose qualities are specifically linked to their area of production. Geographical indications in the EU comprise:

a. PDO – Protected Designation of Origin (for food, agricultural products and wines)

   Every part of the production, processing and preparation process must take place in the specific region. For wines, this means that the grapes have to come exclusively from the geographical area where the wine is made.

   Food and agricultural products granted with a PDO must be labelled as such according to Figure 1.3.a. That label is optional for PDO wines.

b. PGI – Protected Geographical Indication (for food, agricultural products and wines)

   For most products, at least one of the stages of production, processing or preparation takes place in the region. In the case of wine, this means that at least 85% of the grapes used have to come exclusively from the geographical area where the wine is actually made.

   Food and agricultural products granted with a PGI must be labelled as such according to Figure 1.3.b. That label is optional for PGI wines.

c. GI – Geographical Indication (for spirit drinks and aromatised wines)

For most products, at least one of the stages of distillation or preparation takes place in the region. However, raw products do not need to come from the region.

Label in Figure 1.3.b is optional for all GI products.

**a.** **b.**


**Figure 1.3.** EU quality labels for GI quality schemes: **(a)** PDO; **(b)** PGI and GI.

Products that are under consideration or have been granted GI recognition are listed in quality products registers (*eAmbrosia* for food and agricultural products, wine, spirits and aromatised wine; *GIview* for all geographical indications protected at European Union level). These registers also include information on the geographical and production specifications for each product [37].

EU legislation on agricultural quality schemes for agricultural products and foodstuffs identifies products and foodstuffs farmed and produced to exact specifications whilst encouraging diverse agricultural production, protecting product names and informing consumers about the specific character of agricultural products and foodstuffs [38]. In this sense, the European Commission has adopted several regulations on:

- The application of EU quality schemes for the agricultural and food sector:
  - *Regulation (EU) No 1151/2012* [39], which has replaced *Council Regulation (EC) No 510/2006* [40], whose principal provisions are maintained in the former.
  - *Commission Regulation (EC) No. 1107/96* [41].
- How to use the logos in relation to each quality scheme:
  - *Commission Delegated Regulation (EU) No 664/2014* [42].
- How the quality schemes should be applied:
  - *Commission Implementing Regulation (EU) No 668/2014* [43].

Additionally, the EU has covered the labelling guidelines for agri-food products which use PDOs or PGIs as ingredients in *Commission Communication 2010/C 341/03* [44].

As part of the EU's system of intellectual property rights, names of products registered as GIs (PDO, PGI and GI) are legally protected against imitation and misuse within the EU and in non-EU countries where a specific protection agreement has been signed. For all quality schemes, each EU country's competent national authorities must take the necessary measures to protect the registered names within their territory. They should also prevent and stop the unlawful production or marketing of products using such a name [37].

Particularly, the query to *eAmbrosia* about the total of products from Spain registered as GIs yielded the following results on 22 September 2021:

- 103 foodstuffs and/or agricultural and 97 wines registered as PDO.
- 96 foodstuffs and/or agricultural and 42 wines registered as PGI.
- 19 spirit drinks and 1 aromatised wine registered as GI.

One of these protected products is *Queso Zamorano*, which was granted a PDO on a nationwide basis on 6 May 1993 [45] and registered as such in the EU's quality products registers on 21 June 1996 (file number PDO-ES-0089 in *eAmbrosia*). The Annex to [45] details the regulation of both this PDO and its Regulatory Council (*Consejo Regulador, C.R.D.O. "Queso Zamorano"*), which is the ultimate institution in charge of defending, supervising, protecting, promoting and controlling the quality of this registered foodstuff within Spain [45]. The specific dimensions of quality for *Queso Zamorano* are detailed in [45] and reviewed and updated in [46], and comprise aspects related to milk production and origin, cheese manufacturing and ripening and precise sensory, physical and physicochemical characteristics of the final product. As stated in [46], the *Instituto Tecnológico Agrario de Castilla y León* (ITACyL), through the *Subdirección de Calidad y Promoción Alimentaria de la Junta de Castilla y León*, is the competent authority to carry out the verification of the compliance of that product with the specifications, as well as the application of the sanctioning regime included in Title III of Law 24/2003, of July 10, on Vine and Wine. The physicochemical tests to assess the quality of *Queso Zamorano* against the specifications collected in [46] are specifically performed in the *Estación Tecnológica de la Leche*, ISO 17025 accredited laboratory belonging to the ITACyL network. **This is the scenario where the research work described in Chapter 2 of this doctoral thesis was conducted**, where a fast non-destructive method was designed for the determination of the fat, dry matter and protein contents in *Queso Zamorano* from near infrared (NIR) signals and partial least-

squares regression (PLS) against the corresponding results yielded by ISO reference methodologies, namely:

- *Cheese and processed cheese products – Determination of fat content – Gravimetric method: ISO 1735:2004 (IDF 5:2004)*, which will be withdrawn and replaced by PNE-prEN ISO 23319 (ISO/DIS 23319).
- *Cheese and processed cheese products – Determination of the total solids content – Gravimetric method: ISO 5534:2004 (IDF 4:2004).*
- *Milk – Determination of nitrogen content – Part 1: Kjeldahl method: ISO 8968–1:2001 (IDF 20–1:2001)*. This standard was reviewed after the consecution of the study detailed in Chapter 2 of this doctoral thesis, and a new version was released in 2014 (ISO 8968–1:2014 (IDF 20–1:2014)).

The guidance on the use of NIR spectrometry in the determination of the total solids, fat and protein contents in cheese provided by *ISO 21546:2006* [47] (recently updated to *ISO 21546:2020* [48]) was followed to develop the intended PLS (one of the multivariate algorithms recommended by [47]) calibration model. The main changes in [48] compared with the previous edition in [47] are as follows:

*i)* The measurement principles "*Transmittance*" and "*Transflectance*" have been added and defined.

*ii)* All sample types have been covered: liquids, solids and semi-solids.

*iii)* The calibration and validation sections have been reviewed and updated.

*iv)* The outlier section has been revised and the plots renewed.

*v)* The procedures for sample handling and measurement have been expanded to include liquid samples and other examples.

*vi)* Annex A has been expanded to include raw milk analysis references.

Anyway, the key points on sample preparation, instrumental, analysis and calibration model estimation stay unaltered from the 2006 edition.

As a consequence, once both legislation frameworks (the one in effect during the research period for this work and the current one, respectively) have been collated, it can be concluded that the criteria followed and decisions taken back then throughout the analyses described in Chapter 2 of this doctoral thesis are completely supported by the legislation currently in force.

### 1.2.2. Food Quality from a Food Safety perspective

In the wake of a series of food incidents in late 1990s such as the bovine spongiform encephalopathy outbreak and the dioxin scare, attention was drawn to the need to establish general principles and requirements concerning food and feed, so European food safety policy underwent substantial reforms in the early 2000s. The "*Farm to Fork*" integrated approach was thus defined [49], primarily set out in its White Paper on Food Safety and then more formally laid down in *Regulation (EC) No 178/2002 of the European Parliament and of the Council* [50], where, broadly speaking:

- The General Food Law Regulation (Chapter II of [50]) ensures a high level of protection of human life and consumers' interests in relation to food, while ensuring the effective functioning of the internal market. This body of legislation sets out an overarching and coherent framework for the development of food and feed legislation at both Union and national levels by laying down general principles, requirements and procedures that underpin decision making in matters of food and feed safety based on a risk assessment approach, covering all stages of food and feed production and distribution.

- An independent agency responsible for scientific advice and support to help to protect consumers, animals and the environment from food-related risks, the European Food Safety Authority (EFSA), is set up (Chapter III of [50]).

- The Rapid Alert System for Food and Feed (RASFF) is introduced (Chapter IV of [50]), allowing Member States and the Commission to exchange information rapidly and to coordinate their responses to health threats caused by food or feed.

This integrated food safety policy in the EU [49] has enabled to:

- Assess compliance with EU standards in food and feed safety and quality within the EU and in non-EU countries in relation to their exports into the EU.

- Set up effective control systems based on the basic Union rules with regard to food and feed law laid down in [50] and in more specific legislative frameworks that cover several fronts, such as:

    - Hygiene of foodstuffs.

    - Food contamination. This point will be thoroughly tackled throughout Section 1.2.2.1.

    - Food labelling.

- Substances added to food, such as food additives, food enzymes, food flavourings or food supplements.
- Animal and plant health.
- Animal feed and feed labelling.
- Novel foods (namely, foods not consumed within the EU to a significant degree before May 1997).
- Genetically modified organisms (GMOs: organisms, with the exception of human beings, in which the genetic material has been altered in a way that does not occur naturally by mating and/or natural recombination [51]).

#### *1.2.2.1. Food contamination*

According to *Council Regulation (EEC) No 315/93* [52], "*contaminant*" means any substance not intentionally added to food which is present in such food as a result of the production (including operations carried out in crop husbandry, animal husbandry and veterinary medicine), manufacture, processing, preparation, treatment, packing, packaging, transport or holding of such food, or as a result of environmental contamination. Extraneous matter, such as, for example, insect fragments, animal hair, etc., is not covered by this definition.

Since contamination generally has a negative impact on the quality of food and may imply a risk to human health, the EU has taken measures to minimise contaminants in foodstuffs. In this sense, food containing a contaminant in an amount which is unacceptable from the public health viewpoint and in particular at a toxicological level shall not be placed on the market. Furthermore, contaminant levels shall be kept as low as can reasonably be achieved by following good practices at all the stages listed above. However, in order to protect public health, maximum levels for specific contaminants may be established where necessary. These levels shall be adopted in the form of a non-exhaustive list and may include:

- Limits for the same contaminant in different foods.
- Analytical detection limits.
- A reference to the sampling and analysis methods to be used.

Official controls on relevant substances including substances to be used in food contact materials, contaminants, non-authorised, prohibited and undesirable substances whose use or presence on crops or animals or to produce or process food or feed may result in residues of those substances in food or feed [49] must be performed to verify compliance

of foodstuffs with the rules in the areas of *i)* food and food safety, integrity and wholesomeness at any stage of production, processing and distribution of food, and *ii)* feed and feed safety at any stage of production, processing and distribution of feed and the use of feed [38].

1.2.2.1.1. Natural food contamination or resulting from cultivation or production procedures

Maximum levels for contaminants that may occur naturally or result from cultivation practices or production processes have been established and are regularly reviewed [53,54,55]. The development of sampling and analytical methods for the control of some specific substances, such as mycotoxins in foodstuffs, dioxins, dioxin-like polychlorinated biphenyls (PCBs) and non-dioxin-like PCBs in foodstuffs and lead, cadmium, mercury, benzo(a)pyrene and other trace elements and processing contaminant in foodstuffs, must be expressly developed according to [56,57,58].

1.2.2.1.2. Food contamination from residues of veterinary medicinal drugs, pesticides or biocidal products

Residues in foodstuffs might also originate from food-producing animals that have been treated with veterinary medicines or from plants exposed to pesticides or biocidal products, so EU countries must monitor food of animal and/or plant origin for the presence of residues. Maximum residue limits (MRLs) have been set and are updated periodically, so that no foodstuffs containing unacceptable quantities of contaminant substances may be marketed in the EU. **This is the scenario where the research work described in Chapter 3 of this doctoral thesis was conducted**, where the determination of six tranquillisers and a beta-blocker in pig muscle tissue by LC-MS/MS was optimized and ruggedness-assessed once the most appropriate experimental response had been selected. In this case, the legislation governing the design of residue monitoring methodologies is as follows:

a. *Regulation (EU) 2017/625 on official controls and other official activities performed to ensure the application of food and feed law, rules on animal health and welfare, plant health and plant protection products* [38].

   When the research work depicted in Chapter 3 of this doctoral thesis was performed, *Council Directive 96/23/EC on measures to monitor certain substances and residues thereof in live animals and animal products* [59] was in force, in whose Annex I substances of that kind appeared sorted in two groups, as reproduced in Table 1.1:

**Table 1.1.** Classification of certain substances and residues thereof to be detected in live animals, their excrement and body fluids and in tissue, animal products, animal feed and/or drinking water, as stated in Annexes I and II to Council Directive 96/23/EC.

| **GROUP A — Substances having anabolic effect and unauthorized substances** | **GROUP B — Veterinary drugs and contaminants** |
|---|---|
| (1) Stilbenes, stilbene derivatives, and their salts and esters.<br>(2) Antithyroid agents.<br>(3) Steroids.<br>(4) Resorcylic acid lactones including zeranol.<br>(5) Beta-agonists.<br>(6) Compounds included in Annex IV to Council Regulation (EEC) No 2377/90 of 26 June 1990. | (1) Antibacterial substances, including sulphonomides, quinolones.<br>(2) Other veterinary drugs.<br>(a) Anthelmintics.<br>(b) Anticoccidials, including nitroimidazoles.<br>(c) Carbamates and pyrethroids.<br>(d) Sedatives.<br>(e) Non-steroidal anti-inflammatory drugs (NSAIDs).<br>(f) Other pharmacologically active substances<br>(3) Other substances and environmental contaminants<br>(a) Organochlorine compounds including PcBs.<br>(b) Organophosphorus compounds.<br>(c) Chemical elements.<br>(d) Mycotoxins.<br>(e) Dyes.<br>(f) Others. |

- For Group A substances, surveillance should be aimed at detecting the illegal administration of prohibited substances and the abusive administration of approved substances, as applicable.
- For Group B substances, surveillance should be aimed particularly at controlling the compliance with MRLs for residues of veterinary medicinal products and with the maximum levels for pesticides, and monitoring the concentration of environmental contaminants.

Along with other Union acts, Council Directive 96/23/EC has been repealed and replaced by Regulation (EU) 2017/625 in an effort to rationalise and simplify the overall legislative framework and thus establish a harmonised Union framework for the organisation of official controls and official activities other than official controls along the entire agri-food chain. However, in Article 150 of Regulation (EU) 2017/625 transitional measures related to the repeal of Council Directive 96/23/EC are taken, so that competent authorities shall continue to perform the official controls necessary to detect the presence of the substances and groups of residues listed in Annex I to Council Directive 96/23/EC, in accordance with Annexes II, III and IV to that Directive, instead of the corresponding provisions of this Regulation, until 14 December 2022 or an earlier date to be determined in any delegated act eventually adopted by the Commission. So the list of substances and residues in Table 1.1 remains valid so far.

b. *Regulation (EC) No 470/2009 of the European Parliament and of the Council of 6 May 2009 laying down Community procedures for the establishment of residue limits of pharmacologically active substances in foodstuffs of animal origin* [60].

c. *Commission Regulation (EU) No 37/2010 on pharmacologically active substances and their classification regarding maximum residue limits in foodstuffs of animal origin* [61].

   As laid down in [60], the Annex to this Regulation sets out the list of pharmacologically active substances and their classification into allowed substances (in Table 1 of that Annex) and prohibited substances (in Table 2 of that Annex) regarding MRLs in food of animal origin.

d. *Commission Implementing Regulation (EU) 2021/808 on the performance of analytical methods for residues of pharmacologically active substances used in food-producing animals and on the interpretation of results as well as on the methods to be used for sampling* [62].

   When the studies collected in Chapter 3 of this doctoral thesis were being carried out, *Commission Decision 2002/657/EC implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results* [63] was in force, so the identification step conceived as part of the analytical methodologies there designed was by following the performance requirements set out in this act, namely:

   *i)* As stated in Section 2.3 of the Annex to [63], confirmatory methods (those that provide full or complementary information enabling the substance to be unequivocally identified and if necessary quantified at the level of interest) for organic residues or contaminants shall provide information on the chemical structure of the analyte. Consequently, methods based only on chromatographic

analysis without the use of spectrometric detection are not suitable on their own for use as confirmatory methods [63]. Therefore, when dealing with the issues discussed in Chapter 3, LC-MS/MS was selected as the measuring technique for the corresponding compounds of interest. This methodology is specifically declared in Table 1 of the Annex to [63] as a suitable confirmatory method for the substances included in both Group A and Group B of Annex I of Council Directive 96/23/EC (see Table 1.1).

*ii)* As stated in Section 2.3.3.2 of the Annex to [63], when mass spectrometric determination is performed by fragmentography, such as Selected Ion Monitoring (SIM), a system of identification points shall be used to interpret the data. For the confirmation of substances listed in Group A of Annex I of Council Directive 96/23/EC, a minimum of 4 identification points shall be required, while for those in Group B of Annex I of Council Directive 96/23/EC, a minimum of 3 identification points are required [63]. The amount of identification points earned per every diagnostic ion selected for the analytical detection step will depend on the chosen mass spectrometric technique, as depicted in Table 1.2.

**Table 1.2.** Relationship between a range of classes of mass fragment and identification points earned (taken from Table 5 in the Annex to [63]).

| Mass spectrometric (MS) technique | Identification points earned per ion |
|---|---|
| Low resolution mass spectrometry (LRMS) | 1.0 |
| LR-$MS^n$ precursor ion | 1.0 |
| LR-$MS^n$ transition products | 1.5 |
| High resolution mass spectrometry (HRMS) | 2.0 |
| HR- $MS^n$ precursor ion | 2.0 |
| HR-$MS^n$ transition products | 2.5 |

So, for example, if *N* ions were selected for a GC-MS (SIM) analysis (either by Electron Ionization (EI) or by Chemical Ionization (CI)) of this kind of substances, *N* identification points would result.

*iii)* Pursuant to d.*i)* above, the identification of every compound from any chromatographic technique coupled to mass spectrometric detection must take place at both levels:

- Chromatographic level: according to Section 2.3.3.1 of the Annex to [63], the retention time (or relative retention time) of the analyte in the test portion

shall match that of the calibration standard within a specified retention time window. The ratio of the chromatographic retention time of the analyte to that of the internal standard, i.e. the relative retention time of the analyte, shall correspond to that of the calibration solution at a tolerance of ± 0.5% for GC and ± 2.5% for LC [63].

- Spectral level: as for Section 2.3.3.2 of the Annex to [63], either for full scan or SIM detection, the relative intensities of the detected ions, expressed as a percentage of the intensity of the most intense ion or transition (base peak), shall correspond to those of the calibration standard, either from calibration standard solutions or from spiked samples, at comparable concentrations, measured under the same conditions, within the following tolerances:

**Table 1.3.** Maximum permitted tolerances for relative ion intensities (taken from Table 4 in the Annex to [63]).

| Relative intensity (% of base peak) | EI-GC-MS (relative) | CI-GC-MS, GC-MS$^n$, LC-MS, LC-MS$^n$ (relative) |
|---|---|---|
| > 50% | ± 10% | ± 20% |
| > 20% to 50% | ± 15% | ± 25% |
| > 10% to 20% | ± 20% | ± 30% |
| ≤ 10% | ± 50% | ± 50% |

Eventually, as a consequence of Regulation (EU) 2017/625 [38] coming into effect, Commission Decision 2002/657/EC is repealed from the date of entry into force of Commission Implementing Regulation (EU) 2021/808. However, in Article 7 of [62] it is also stated that until 10 June 2026, the requirements laid down in points 2 ("*Performance criteria and other requirements for analytical methods*") and 3 ("*Validation*") of the Annex to Decision 2002/657/EC shall continue to apply to methods which have been validated before the date of entry into force of [62], thus enabling them to remain in use for a limited period while their re-validation according to the requirements laid down in [62] takes place. Nevertheless, some comments must be made after confronting the content of [62] and [63]:

- In contrast to point d.*i)* above, paragraph 4 within Section 1.2.1 of Chapter I ("*Performance criteria and other requirements for analytical methods*") in Annex I to [62] sets out that analytical methods based only on chromatographic analysis without the use of mass spectrometric detection are not suitable on their own for use as confirmatory methods for prohibited or unauthorised pharmacologically active substances, not specifically including the allowed type within this rule. In the case

of mass spectrometry not being suitable for authorised substances, only HPLC-DAD and HPLC-FLD, or a combination of them, can be used as alternative for mass spectrometry based methods, provided that the relevant criteria for these techniques in Sections 1.2.5.1 and 1.2.5.2, respectively, of Chapter I in Annex I to [62] are fulfilled. So, analytical techniques such as LC-UV/VIS (single wavelength) are no longer suitable on their own for use as a confirmatory method.

- As described in point d.*ii)* above, a system of identification points shall still be used to select adequate acquisition modes and evaluation criteria for analyte identification, according to Section 1.2.4.2 of Chapter I in Annex I to [62]. Regarding this criterion, for confirmation of the identity of substances in a matrix for which an MRL is established (authorised use), a minimum of 4 identification points is required in [62], whereas 5 identification points are at least needed for unauthorised or prohibited substances. However, unlike [63], one point can originate from the chromatographic separation. Table 1.4 shows the number identification points that each of the techniques yields, bearing in mind that all mass spectrometric analyses shall be combined with a separation technique that shows sufficient separation power and selectivity for the specific application, such as liquid and gas chromatography, capillary electrophoresis (CE) and supercritical fluid chromatography (SFC).

**Table 1.4.** Identification points per technique (taken from Table 3 in Annex I to [62]).

| Analytical technique | Identification points |
|---|---|
| Separation (mode GC, LC, SFC, CE) | 1 |
| LR-MS ion | 1 |
| Precursor ion selection at < ± 0.5 Da mass range | 1 (indirect) |
| LR-$MS^n$ product ion | 1.5 |
| HR-MS ion | 1.5 |
| HR-$MS^n$ product ions | 2.5 |

So, for example:

- If *N* ions were selected for a GC-MS (SIM) analysis (either by Electron Ionization (EI) or by Chemical Ionization (CI)), 1+*N* identification points would result.

- If 1 precursor ion and two product ions for a LC-MS/MS analysis were selected, 1+1+(2·1.5) = 5 identification points would result. This would have been the outcome if [62] had been followed for the identification of the seven analytes monitored in the studies described in Chapter 3 of this doctoral thesis bearing in mind the MS/MS transitions chosen there. In conclusion, the identification schema designed back then for those analytes also fulfils the new criteria.

- As for point d.*iii)* above:

  - At a chromatographic level, [62] sets out specific criteria for both absolute and relative retention times, namely:

    - The retention time of the analyte in the extract shall correspond to that of the calibration standard, a matrix-matched standard or a matrix-fortified standard with a tolerance of ± 0.1 minute.

    - In case an internal standard is used, the ratio of the chromatographic retention time of the analyte to that of the internal standard, that means the relative retention time of the analyte, shall correspond to that of the calibration standard, matrix-matched standard or matrix-fortified standard with a maximum deviation of 0.5% for gas chromatography (same requirement as in [63]) and of 1% for liquid chromatography for methods validated from the date of entry into force of [62].

  - At a spectral level, the relative intensities of the diagnostic ions (ion ratio), expressed as a percentage of the intensity of the most abundant ion or transition, shall correspond to those of the matrix-matched standards, matrix-fortified standards or standard solutions at comparable concentrations, measured under the same conditions, within ± 40% relative deviation. Comparing this value with those collected in Table 1.3 and in [63], it can be concluded that the former means a quite more relaxed performance criterion, especially for those less abundant ions.

Consequently, after having compared the legislation frameworks in force during the research period for this work and at present, it can be concluded that the criteria followed and decisions taken back then throughout the analyses embedded in Chapter 3 of this doctoral thesis are fully backed up by the legislation currently in force.

1.2.2.1.3. Food contamination from food contact materials (FCMs)

Food comes into contact with many materials and articles during its production, processing, storage, preparation and serving, before its eventual consumption. Such materials and articles are generically called *Food Contact Materials* (FCMs). Either directly or indirectly, FCMs are either intended to be brought into contact with food, are already in contact with food, or can reasonably be brought into contact with food or transfer their constituents to food under normal or foreseeable use. As a consequence, the EU has defined several binding rules pertaining to FCMs such as materials for transporting or processing food, as well as packaging materials and kitchen or tableware. EU law may be complemented with Member States national legislation if specific EU rules do not exist [64].

*Regulation (EC) No 1935/2004* [65] provides a harmonised legal EU framework by setting out the general principles of safety and inertness for all FCMs. It requires that materials do not:

- Release their constituents into food at levels harmful to human health
- Change food composition, taste and odour in an unacceptable way.

Moreover, this framework provides:

- For special rules on active and intelligent materials (they are not inert by their design).
- Powers to enact additional EU measures for specific materials (e.g. for plastics).
- The procedure to perform safety assessments of substances used to manufacture FCMs involving the EFSA.
- Rules on labelling including an indication for use or by reproducing the appropriate symbol.
- For compliance documentation and traceability.

In addition to this general legislation, certain FCMs (ceramic materials, regenerated cellulose film, plastics (including recycled plastic), as well as active and intelligent materials) are covered by specific EU measures [66]. The most comprehensive one is *Commission Regulation (EU) No 10/2011* [67] on plastic materials and articles in view of the undeniable leading role that at the present moment plastics play in food packaging for producers and retailers from all over the world: compared with other materials such as cardboard, cans or glass, plastic is light, flexible and inexpensive, and it also promotes shelf life and facilitates the transport and use of products.

[67] sets out rules on the composition of plastic FCMs. Annex I to this regulation establishes a Union list of substances that are permitted for use in the manufacture of plastic FCMs. [67] also specifies restrictions on the use of these substances and sets out rules to determine the compliance of plastic materials and articles. In connection with this, to ensure the safety of plastic materials, migration limits, which specify the maximum amount of substances allowed to migrate into food in mg of substance per kg of food, are fixed. For the substances on the Union list, [67] sets out *Specific Migration Limits* (SML) established by EFSA on the basis of toxicity data of each specific substance. On the other hand, [67] lays down detailed migration testing rules as well. Although migration testing in food prevails, migration is usually tested using "*simulants*", which are representative for a food category. The migration testing is done under standardised time/temperature conditions, representative for a certain food use, and covers the maximum shelf life of packed food.

[67] is regularly amended due to the unceasing innovation and research on plastics and their toxicity, as the following list proves:

- *Commission Regulation (EU) 2020/1245*.
- *Commission Regulation (EU) 2019/37*.
- *Commission Regulation (EU) 2018/831*.
- *Commission Regulation (EU) 2017/752*.
- *Commission Regulation (EU) 2016/1416*.
- Specific amendments of the Union list of authorised substances (in Annex I to [67]).
    - *Commission Regulation (EU) 321/2011* [68], prohibiting the use of 2,2-bis(4-hydroxyphenyl)propane (bisphenol A, BPA) for the manufacture of polycarbonate infant feeding bottles, where "*infant*" means a child under the age of 12 months [69].
    - *Commission Regulation (EU) 1282/2011*.
    - *Commission Regulation (EU) 1183/2012*.
    - *Commission Regulation (EU) 202/2014*.
    - *Commission Regulation (EU) 2015/174*.
    - *Commission Regulation (EU) 2018/79*.
    - *Commission Regulation (EU) 2018/213* [70] on the use of BPA in varnishes and coatings intended to come into contact with food given the extensive use of this

compound in epoxy resins, particularly for application on the interior of food cans, so:

- The migration into or onto food of BPA from varnishes or coatings applied to materials and articles shall not exceed a specific migration limit of 0.05 mg of BPA per kg of food, where "*varnishes*" or "*coatings*" means materials or articles composed of one or more non-self-supporting layer or layers manufactured using BPA, applied on a material or article in order to impart special properties on it or to improve its technical performance. As a consequence, the SML for BPA is reduced from 0.6 mg per kg of food, in force since Regulation (EU) No 10/2011 was initially issued, to 0.05 mg per kg of food.
- No migration of BPA shall be permitted from varnishes or coatings applied to materials and articles specifically intended to come into contact with infant formula, follow-on formula, processed cereal-based food, baby food, food for special medical purposes developed to satisfy the nutritional requirements of infants and young children or milk-based drinks and similar products specifically intended for young children, which are specifically those aged between one and three years, as referred to in [69].
- BPA cannot be used for the manufacture of polycarbonate drinking cups or bottles which, due to their spill proof characteristics, are intended for infant and young children.

o *Commission Regulation (EU) 2019/1338*.

On the other hand, EUR 24105 EN [71] illustrates the required performance criteria for the analytical methods applied in the laboratories for FCMs and provide procedures for method validation in order to estimate their performance characteristics. Annex 2 to [71] collects the confirmatory criteria for chromatographic methods (GC, LC) needed for assuring the unequivocal identification of migrants from food contact materials. Those requirements match exactly those set out in the Annex to [63], which have been commented previously in Section 1.2.2.1.2.

**This is the scenario where the research works described in Chapters 4 and 5 of this doctoral thesis were conducted**, where the unequivocal determination of several plastic migrants was tackled:

- An analytical method to determine bisphenol F (BPF), bisphenol A (BPA), and their corresponding diglycidyl ether derivatives (BFDGE and BADGE, respectively) by GC-MS was developed and optimized. As can be inferred from the paragraphs above,

BPA is an industrial organic chemical that has been subjected to several reviews and assessments by the European authorities since 2008 due to its potential hazardous properties: BPA may damage fertility and has been identified as a substance affecting the hormonal systems of humans and animals, so, in 2017, BPA was included in the Candidate List of Substances of Very High Concern (SVHC) under category Repr. 1B due to its toxic for reproduction and endocrine-disrupting properties. In addition, BPA damages eyes and may cause allergic skin reactions and respiratory irritation. That is why the use of this substance is being continuously questioned and limited in the EU to protect people's health and the environment; in this sense, and apart from the restrictions in [67,68,70] concerning the food and beverage market:

- REACH (Registration, Evaluation, Authorization, and Restriction of Chemicals), an EU regulation that limits the use of certain chemical substances in consumer products, mandates an additional restriction on the use of BPA in thermal paper: after 2 January 2020, BPA shall not be placed on the EU market in thermal paper in a concentration equal to or greater than 0.02% by weight [72]. The main reason for this decision is that cashiers are exposed to BPA as they must often handle a great number of thermal paper receipts. The category most at risk is represented by unborn children of pregnant workers.

  Due to this restriction, BPA has been increasingly replaced with BPS in thermal paper: by 2022, it is expected that 61% (or 307 kilotonnes) of all thermal paper in the EU will be BPS-based [73]. This raises concern again, as BPS is suspected to affect human reproductive and hormonal systems since it is analogous to BPA.

- The *Toy Safety Directive* [74] sets up the permissible migration limit values for chemicals used in chewable toys by children under the age of 3 years old. As for BPA, *Commission Directive (EU) 2017/898* [75], by amending Appendix C to Annex II to [74], requires that the migration limit for BPA should be less than 0.04 mg/L in accordance with the methods laid down in EN 71-10:2005 and EN 71-11:2005.

Regarding the study detailed in Chapter 4 of this thesis, together with the already-in-force [65,67,68,71], *Commission Regulation (EC) 1895/2005* [76] was also considered for its development and performance, where, among other provisions, the use and/or presence of BFDGE in the manufacture of materials and articles intended to come into contact with food are prohibited. The restrictions laid down in [70], which came into effect once this research had been concluded, do not alter either the applicability of the designed methodology or its agreement with the current legal framework on this matter.

- The multiresidue determination of diisobutyl phthalate (DiBP), diisononyl phthalate (DiNP), 2,6-di-*tert*-butyl-4-methyl-phenol (BHT), bis(2-ethylhexyl) adipate (DEHA) and benzophenone (BP) by GC-MS is presented in Chapter 5, where [65,67,71] were borne in mind for its development and consecution.

## 1.3. Content of this doctoral thesis

The research work presented in this doctoral thesis and developed within the quality, chemometric and legal frameworks detailed in Sections 1.1 and 1.2 has been arranged as follows:

- Chapter 2 describes the design of a non-destructive analytical method to determine the total solids, fat and protein contents model on representative *Queso Zamorano* samples from NIR signals and PLS regression after having measured their absorbance at wavelengths in the NIR region and transformed this preprocessed spectral data into constituent concentrations by considering the results yielded by the reference analytical methods mentioned in Section 1.2.1 once the outliers had been removed from the calibration set.

  The methodology achieved has several advantageous features to be highlighted, namely:

  - It saves time of analysis significantly, in comparison with the internationally accepted but highly time-consuming reference methodologies that were routinely being followed.
  - It enables to assess the probability of false compliance when dealing with the minimum permitted levels of those constituents stated in [45,46], so the quality of *Queso Zamorano* in terms of its physicochemical composition can be easily verified.

- Chapter 3 collects two consecutive studies:

  - As part of an optimization stage, Section 3.2 shows the effect analysis of four experimental factors related to the extraction and purification steps of six tranquillisers and an anti-adrenergic in pig muscle tissue analysed by LC-MS/MS following the criteria in [63]. Since sample pretreatment is thought to affect the quantification of each analyte and its internal standard differently as the latter must be added to every sample at the beginning of the extraction procedure, three different response variables have been evaluated:

    a. An overall desirability function born from the combination of individual desirability functions for the absolute peak areas of every analyte and its internal standard.

    b. The absolute peak area of every analyte.

    c. The standardized peak area of every analyte to that of its corresponding internal standard.

The fit-for-purpose strategy designed here from the blending of desirability functions together with D-optimal designs proves to be particularly useful to optimize chromatographic methods by adapting the methodology of experimental designs exactly to the analytical problem in question.

- After being optimized throughout Section 3.2, the robustness of the sample preparation stage of the resulting analytical method is verified in Section 3.3 according to the guidelines in [63], where the introduction and testing of several variations at once is recommended instead of studying one alteration at a time by means of a Youden fractional factorial design. In this case, the conclusions on the influence of seven factors of the sample pretreatment phase on the final concentration of the seven sedatives in pig muscle samples have been drawn from three different approaches:

  a. Hypothesis testing using an external variance previously estimated.
  b. Lenth's method.
  c. Bayesian analysis.

  Since the same outcome is attained from these three approaches, both b and c options reveal themselves as advantageous strategies to determine the effect of every factor when there are no degrees of freedom left and no estimate of the experimental variability is available.

- A GC-MS method to determine several bisphenols (BPF, BPA) and their corresponding diglycidyl ether derivatives (BFDGE, BADGE) is developed and optimized throughout Chapter 4 by considering:

  - Some figures of merit that account for the quality of an analytical procedure, as appears in Section 4.2. In this context, an experimental design to evaluate how changes in the oven temperature program affect the quality of both determinations and identifications has been carried out by taking advantage of the chromatographic and spectral specificity provided by PARAFAC decompositions.

    The release of BPA from polycarbonate tableware into food simulant *D1* (ethanol 50% (v/v)) is assessed by using the analytical procedure optimized before.

  - The physicochemical properties of all the analytes so as to avoid choosing experimental conditions that could be unfavourable for some of them, in view of the poor recovery rates of both diglycidyl ether derivatives after the first optimization. The pH to which samples are adjusted during the initial preparation step strongly affects the ionisation degree of every compound, which is directly

dependent on the nature of the substituents of the aromatic ring. As can be seen in Section 4.3, this fact is taken into account for the evaluation of the influence of pH on the final recoveries in order to reach a compromise solution that enables the determination of the four analytes under study.

Once a better quantification of all the target compounds has been achieved through the newly optimized method, a kinetics study of the release of BPA from polycarbonate glasses into food simulant *D1* is conducted: two sets of seven migration tests are performed at two different incubation temperatures are performed. Even though a general migration law cannot be postulated due to structural differences in the evaluated glasses, it can be concluded that the relative migration rate of BPA changes with the heating temperature significantly.

It is worth mentioning that, in the studies depicted throughout both Sections 4.2 and 4.3, coelution of interferents has been overcome using PARAFAC decomposition; otherwise, the mandatory unequivocal identification of some analytes could not have been ensured according to the EU regulations currently in force and listed in Section 1.2, which accounts for the strength of this kind of approaches in terms of analytical quality and eventually food safety.

- Chapter 5 presents the multiresidue determination of several plasticizers by GC-MS followed by PARAFAC/PARAFAC2 decompositions in the midst of a "hostile" laboratory environment because of the ubiquitous presence of phthalates, which hinders the clear discrimination between the amount of phthalates coming from the test samples and that from blank ones; this drawback has led to the assessment of this baseline level of phthalates using $\alpha$ and $\beta$ error probabilities.

  Additionally, another difficulty has been overcome thanks to the devised statistical strategy: the chromatogram of DiNP appears as a multiple finger-peak signal because of an array of possible $C_9$ isomers. PARAFAC offers a more trustworthy approach to deal with these complex signals and has succeeded in the determination of this analyte since no peak areas are considered but its chromatographic and spectral loadings

  Once again and as commented for Chapter 4, both the identification and the quantification of all the analytes are achieved even if unknown compounds interfere.

  Finally, the presence of the plasticizers under study in a dummy intended for infants has been checked by means of the designed analytical methodology.

## 1.4. Objectives of this doctoral thesis

### 1.4.1. General objectives

The general objectives to be fulfilled during the execution of this thesis are presented below:

G1. Develop analytical methodologies that enable the unequivocal identification and/or quantification of all the target substances or constituents according to the requirements in both the EU and Spanish legislation currently in force and applicable international guidelines.

G2. Attain a higher degree of analytical quality than the traditional approaches by means of advantageous strategies that consider not only the whole analytical information from the instrumental signals thanks to multivariate or multi-way designs, but also several method performance characteristics and the resulting figures of merit jointly.

G3. Assess food quality on real samples in terms of food safety and geographical and traditional origin according to the requirements in both the EU and Spanish legislation currently in force and applicable international guidelines.

### 1.4.2. Specific objectives

Specific objectives can also be defined, bearing in mind the analytical purposes of each research, the different target chemical families assessed and their role in the food market.

**a.** Regarding the study described in ***Chapter 2, Section 2.2 – Prediction of Zamorano cheese quality by near-infrared spectroscopy assessing false non-compliance and false compliance at minimum permitted limits stated by Designation of Origin regulations***:

S1. Develop a fast analytical method to determine the fat, dry matter and protein contents in *Queso Zamorano* from NIR signals and PLS regression in accordance with [45,46,47].

S2. Determine the detection capability of the percentages in weight of total protein, dry matter and fat-to-dry matter ratio in regard to the minimum permitted limits set out in [45,46] by evaluating the probability of false compliance from the multivariate model achieved at S1.

**b.** As for the research work explained in ***Chapter 3, Section 3.2 – Desirability functions as response in a D-optimal design for evaluating the extraction and purification steps of six tranquillizers and an anti-adrenergic by liquid chromatography-tandem mass spectrometry***:

S3. Through the experimental design methodology, assess the effect of four factors belonging to the sample preparation stage in the analytical method aimed at the determination of six tranquillisers and a beta-blocker agent in pig muscle tissue by LC-MS/MS.

S4. Find the optimum solution for the levels those four factors should be at in order to achieve the highest sensitivity possible for every analyte without putting its identification at risk. The chemical behaviour of both the analyte and its internal standard throughout the entire analytical process should be carefully considered, since internal standards must be added to any test sample at the beginning of the extraction procedure, as laid down in [63].

**c.** In relation to the study detailed in ***Chapter 3, Section 3.3 – Robustness testing in the determination of seven drugs in animal muscle by liquid chromatography–tandem mass spectrometry***:

S5. Assess the robustness of the sample preparation stage in the analytical method considered in point **b** once it has been optimized according to S4 by following the guidelines in [63] for the robustness evaluation using the Youden approach.

S6. Given the saturated nature of fractional factorial designs, interpret the outcome from the resulting experimental plan by using both known-residual-error and unknown-residual-error methodologies.

**d.** With regard to the experimentation collected in ***Chapter 4, Section 4.2 – Optimization of a GC-MS procedure that uses parallel factor analysis for the determination of bisphenols and their diglycidyl ethers after migration from polycarbonate tableware***:

S7. Develop a GC-MS method to determine BPA, BPF, BADGE and BFDGE that fulfils the requirements in EU food migrant policies [65,67,68,76].

S8. Taking the matrix structure of the signal provided by GC-MS instrumentation into account, use of PARAFAC decompositions on the experimental data obtained to fully ensure the unequivocal determination (both qualitative and quantitative) of every compound according to the analytical performance criteria in [71].

S9. Through the experimental design methodology, assess the effect caused by changes in some parameters of the oven temperature program on the analytical quality of the determination of the target compounds.

S10. Validate the analytical method optimized at S9 by the evaluation of specificity, sensitivity, accuracy, recovery and detection capability with probabilities of false positive and false negative fixed at 0.05.

S11. Apply the designed analytical method to the study of the release of BPA from polycarbonate tableware into food simulant *D1*.

**e.** As for the research work explained in ***Chapter 4, Section 4.3 – Optimum pH for the determination of bisphenols and their corresponding diglycidyl ethers by gas chromatography-mass spectrometry. Migration kinetics of bisphenol A from polycarbonate glasses***:

S12. Given the low recoveries of the diglycidyl ether compounds achieved by the GC-MS/PARAFAC procedure resulting from previous point **d**, optimize the sample pretreatment step of the analytical method by considering the relationship between pH and the chemical structure of every compound while keeping ensuring its unequivocal identification.

S13. Validate the analytical method just optimized at S12 by the evaluation of specificity, sensitivity, accuracy, recovery and detection capability with probabilities of false positive and false negative fixed at 0.05.

S14. Estimate the kinetic model of the migration of BPA from polycarbonate glasses into food simulant *D1* at two different incubation temperatures.

**f.** Regarding the study described in ***Chapter 5, Section 5.2 – Dealing with the ubiquity of phthalates in the laboratory when determining plasticizers by gas chromatography-mass spectrometry and PARAFAC***:

S15. Develop a GC-MS method to determine several plasticizers that fulfils the requirements in EU food migrant policies [65,67] and that is capable of cope with:

*i.* Complex finger-peak signals.

*ii.* Ubiquity of one of the analytes in the analytical environment, which results in a non-zero variable concentration of this compound in blank samples.

S16. Taking the matrix structure of the signal provided by GC-MS instrumentation into account, use of PARAFAC or PARAFAC2 decompositions on the experimental data obtained to fully ensure the unequivocal determination (both qualitative and quantitative) of every compound according to the analytical performance criteria in [71].

S17. Apply the designed analytical method to the study of the presence of the target compounds in a hexane extract from a dummy intended for infants.

## 1.5. Publications resulting from this doctoral thesis

- M.L. Oca, M.C. Ortiz, L.A. Sarabia, A.E. Gredilla, D. Delgado, *Prediction of Zamorano cheese quality by near-infrared spectroscopy assessing false non-compliance and false compliance at minimum permitted limits stated by designation of origin regulations*, Talanta 99 (2012) 558–565.

- L. Rubio, M.L. Oca, L. Sarabia, I. García, M.C. Ortiz, *Desirability functions as response in a D-optimal design for evaluating the extraction and purification steps of six tranquillizers and an anti-adrenergic by liquid chromatography-tandem mass spectrometry*, Journal of Chemometrics 30 (2016) 58–69.

- M.L. Oca, L. Rubio, M.C. Ortiz, L.A. Sarabia, I. García, *Robustness testing in the determination of seven drugs in animal muscle by liquid chromatography–tandem mass spectrometry*, Chemometrics and Intelligent Laboratory Systems 151 (2016) 172–180.

- M.L. Oca, M.C. Ortiz, A. Herrero, L.A. Sarabia, *Optimization of a GC/MS procedure that uses parallel factor analysis for the determination of bisphenols and their diglycidyl ethers after migration from polycarbonate tableware*, Talanta 106 (2013) 266–280.

- M.L. Oca, M.C. Ortiz, A. Herrero, L.A. Sarabia*, Optimum pH for the determination of bisphenols and their corresponding diglycidyl ethers by gas chromatography–mass spectrometry. Migration kinetics of bisphenol A from polycarbonate glasses*, Journal of Chromatography A 1360 (2014) 23–38.

- M.L. Oca, L. Rubio, L.A. Sarabia, M.C. Ortiz, *Dealing with the ubiquity of phthalates in the laboratory when determining plasticizers by gas chromatography/mass spectrometry and PARAFAC*, Journal of Chromatography A 1464 (2016) 124–140.

- M.C. Ortiz, S. Sanllorente, A. Herrero, C. Reguera, L. Rubio, M.L. Oca, L. Valverde-Som, M.M. Arce, M.S. Sánchez, L.A. Sarabia, *Three-way PARAFAC decomposition of chromatographic data for the unequivocal identification and quantification of compounds in a regulatory framework*, Chemometrics and Intelligent Laboratory Systems 200 (2020) 104003.

## 1.6. References

[1] D.C. Montgomery, *Introduction to Statistical Quality Control*, John Wiley & Sons, Hoboken, New Jersey (USA), 6th ed., 2009

[2] European Commission – Knowledge Centre for Food Fraud and Quality (KC-FFQ) (2021). Knowledge for Policy. *Food Quality*. Retrieved from https://knowledge4policy.ec.europa.eu/food-fraud-quality/topic/food-quality_en on 21 September 2021.

[3] U.S. Food and Drug Administration, *Guidance for Industry. Process Validation: General Principles and Practi*ces, Current Good Manufacturing Practices (2011). Available at https://www.fda.gov/regulatory-information/search-fda-guidance-documents/process-validation-general-principles-and-practices.

[4] International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use, *ICH Harmonised Tripartite Guideline Q8(R2): Pharmaceutical Development*, Geneva, Switzerland, 2009.

[5] W. S. Schlindwein, M. Gibson (eds.), *Pharmaceutical quality by design: a practical approach*. John Wiley & Sons, Hoboken, New Jersey (USA), 1st ed., 2018.

[6] E. Rozet, P. Lebrun, B. Debrus, B. Boulanger, P. Hubert, *Design Spaces for analytical methods*, Trends in Analytical Chemistry 42 (2013) 157-167.

[7] International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use, *ICH Harmonised Tripartite Guideline Q9: Quality Risk Management*, Geneva, Switzerland, 2005.

[8] International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use, *ICH Harmonised Tripartite Guideline Q10: Pharmaceutical Quality System*, Geneva, Switzerland, 2008.

[9] International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use, *ICH Harmonised Tripartite Guideline Q11: Development and Manufacture of Drug Substances (Chemical Entities and Biotechnological/Biological Entities)*, Geneva, Switzerland, 2012.

[10] International Council for Harmonisation (ICH) of Technical Requirements for Pharmaceuticals for Human Use, *ICH Harmonised Guideline Q12: Technical and Regulatory Considerations for Pharmaceutical Product Lifecycle Management*, Geneva, Switzerland, 2019.

[11] U.S. Food and Drug Administration, *Guidance for Industry. PAT – A Framework for Innovative Pharmaceutical Manufacturing and Quality Assurance*. Pharmaceutical CGMPs (2004). Available at https://www.fda.gov/regulatory-information/search-fda-guidance-documents/pat-framework-innovative-pharmaceutical-development-manufacturing-and-quality-assurance.

[12] K.A. Welch, C.A. Fung, S.R. Schmidt, *Risk/Science-Based Approach to Validation: A Win-win-win for Patients, Regulators, and Industry*, Center for Risk-based Strategies, Inc. 58 (2004) pp. 15-23.

[13] R. Murugan, *Design Space for Analytical Methods – QbD Concept*, Cutting Edge Spinco Biotech (2014) 7-11.

[14] J. Teng, C. Zhu, J. Lyu, L. Pan, M. Zhang, F. Zhang, H. Wu, *Analytical Lifecycle Management (ALM) and Analytical Quality by Design (AQbD) for analytical procedure development of related substances in tenofovir alafenamide fumarate tablets*, Journal of Pharmaceutical and Biomedical Analysis (2021). doi: https://doi.org/10.1016/j.jpba.2021.114417.

[15] International Council for Harmonization (ICH) of Technical Requirements for Pharmaceuticals for Human Use, *ICH Q14 Guideline: Analytical Procedure Development and Revision of Q2(R1): Analytical Validation*. Final Concept Paper (2018). Available at https://database.ich.org/stes/default/files/Q2R2-Q14_EWG_Concept_Paper.pdf

[16] M. Schweitzer, M. Pohl, M. Hanna-Brown, P. Nethercote, P. Borman, G. Hansen, K. Smith, J. Larew, *Implications and opportunities of applying QbD principles to analytical measurements*, Pharmaucetical Technology 34 (2010) 52-59.

[17] J. Ermer, *Quality by design: A lifecycle concept for pharmaceutical analysis*, European Pharmaceutical Review 16 (2011) 16-24.

[18] M.M. Arce, S. Sanllorente, S. Ruiz, M.S. Sánchez, L.A. Sarabia, M.C. Ortiz, *Method Operable Design Region obtained with a PLS2 model inversion in the determination of ten polycyclic aromatic hydrocarbons by liquid chromatography with fluorescence detection*, Journal of Chromatography A (2021-09). doi: https://doi.org/10.1016/j.chroma.2021.462577.

[19] M.M. Arce, S. Ruiz, S. Sanllorente, M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, *A new approach based on inversion of a Partial Least Squares model searching for a preset analytical target profile. Application to the determination of five bisphenols by liquid chromatography with diode*

*array detector*, Analytica Chimica Acta, 1149 (2021) 338217.doi: https://doi.org/10.1016/j.aca.2021.338217.

[20] M.C. Ortiz, L.A. Sarabia, A. Herrero, C. Reguera, S. Sanllorente, M.M. Arce, O. Valencia, S. Ruiz, M.S. Sánchez, *Partial Least Squares model inversion in the chromatographic determination of triazines in water*, Microchemical Journal, 164 (2021) 105971. doi: https://doi.org/10.1016/j.microc.2021.105971.

[21] A. Smilde, R. Bro, P. Geladi. *Multi-way Analysis with Applications in the Chemical Sciences*. John Wiley & Sons Ltd, West Sussex, 2004.

[22] M.K. Parr, A.H. Schmidt, *Life Cycle Management of Analytical Methods*, Journal of Pharmaceutical and Biomedical Analysis 147 (2017) 506-517.

[23] P.M. Kroonenberg, *Applied Multiway Data Analysis*, John Wiley & Sons, Inc., Hoboken, New Jersey (USA), 2008.

[24] E. Meléndez, L.A. Sarabia, M.C. Ortiz, *Parallel factor analysis for monitoring data from a grape harvest in qualified designation of origin Rioja including spatial and temporal variability*, Chemometrics and Intelligent Laboratory Systems 146 (2015) 347-353.

[25] P. Geladi, H. Grahn, *Multivariate image analysis*, John Wiley & Sons, Ltd, Chichester, 1996.

[26] M.C. Ortiz, L. Sarabia, *Quantitative determination in chromatographic analysis based on n-way calibration strategies*, Journal of Chromatography A 1158 (2007) 94-110.

[27] M.L. Oca M.C. Ortiz, A. Herrero, L.A. Sarabia, *Optimization of a GC/MS procedure that uses parallel factor analysis for the determination of bisphenols and their diglycidyl ethers after migration from polycarbonate tableware*, Talanta 106 (2013) 266-280.

[28] M.L. Oca, L.A. Sarabia, A. Herrero, M.C. Ortiz, *Optimum pH for the determination of bisphenols and their corresponding diglycidyl ethers by gas chromatography–mass spectrometry. Migration kinetics of bisphenol A from polycarbonate glasses*, Journal of Chromatography A 1360 (2014) 23-38.

[29] M.L. Oca, L. Rubio, L.A. Sarabia, M.C. Ortiz, *Dealing with the ubiquity of phthalates in the laboratory when determining plasticizers by gas chromatography/mass spectrometry and PARAFAC*, Journal of Chromatography A 1464 (2016) 124-140.

[30] R.A. Harshman, *Foundations of the PARAFAC procedure: mode and conditions for an explanatory multi-mode factor analysis*, UCLA Working Papers in Phonetics 16 (1970) 1-84.

[31] J.D. Carroll, J. Chang, *Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition*, Psychometrika 35 (1970) 283-319.

[32] R.B. Cattell, *Parallel proportional profiles and other principles for determining the choice of factors by rotation*, Psychometrika 9 (1944) 267-283.

[33] A.C. Olivieri, G.M. Escandar, *Practical Three-Way Calibration*, Elsevier, Amsterdam (The Netherlands), 2014.

[34] R. Bro, H.A.L. Kiers, *A new efficient method for determining the number of components in PARAFAC models*. Journal of Chemometrics 17 (2003) 274-286.

[35] J.M.F. ten Berge, *Partial uniqueness in CANDECOMP/PARAFAC*, Journal of Chemometrics 18 (2004) 12-16.

[36] R. Bro, C.A. Andersson, H.A.L. Kiers, *PARAFAC2 - Part II. Modeling chromatographic data with retention time shifts*, Journal of Chemometrics 13 (1999) 295-309.

[37] European Commission (2021). European Commission website. *Quality schemes explained*. Retrieved from https://ec.europa.eu/info/food-farming-fisheries/food-safety-and-quality/certification/quality-labels/quality-schemes-explained_en on 21 September 2021.

[38] Regulation (EU) 2017/625 of the European Parliament and of the Council of 15 March 2017 on official controls and other official activities performed to ensure the application of food and feed law, rules on animal health and welfare, plant health and plant protection products, amending Regulations (EC) No 999/2001, (EC) No 396/2005, (EC) No 1069/2009, (EC) No 1107/2009, (EU) No 1151/2012, (EU) No 652/2014, (EU) 2016/429 and (EU) 2016/2031 of the European Parliament and of the Council, Council Regulations (EC) No 1/2005 and (EC) No 1099/2009 and Council Directives 98/58/EC, 1999/74/EC, 2007/43/EC, 2008/119/EC and 2008/120/EC, and repealing Regulations (EC) No 854/2004 and (EC) No 882/2004 of the European Parliament and of the Council, Council Directives 89/608/EEC, 89/662/EEC, 90/425/EEC, 91/496/EEC, 96/23/EC, 96/93/EC and 97/78/EC and Council Decision 92/438/EEC, Official Journal of the European Union, L 95, 1-142.

[39] Regulation (EU) No 1151/2012 of the European Parliament and of the Council of 21 November 2012 on quality schemes for agricultural products and foodstuffs, Official Journal of the European Union, L 343, 1-29.

[40] Council Regulation (EC) No 510/2006 of 20 March 2006 on the protection of geographical indications and designations of origin for agricultural products and foodstuffs, Official Journal of the European Union, L 93, 12-25.

[41] Commission Regulation (EC) No 1107/96 of 12 June 1996 on the registration of geographical indications and designations of origin under the procedure laid down in Article 17 of Council Regulation (EEC) No 2081/92, Official Journal of the European Communities, L 148-I, 1-15.

[42] Commission Delegated Regulation (EU) No 664/2014 of 18 December 2013 supplementing Regulation (EU) No 1151/2012 of the European Parliament and of the Council with regard to the establishment of the Union symbols for protected designations of origin, protected geographical indications and traditional specialities guaranteed and with regard to certain rules on sourcing, certain procedural rules and certain additional transitional rules, Official Journal of the European Union, L 179, 17-22.

[43] Commission Implementing Regulation (EU) No 668/2014 of 13 June 2014 laying down rules for the application of Regulation (EU) No 1151/2012 of the European Parliament and of the Council on quality schemes for agricultural products and foodstuffs, Official Journal of the European Union, L 179, 36-61.

[44] Commission Communication – Guidelines on the labelling of foodstuffs using protected designations of origin (PDOs) or protected geographical indications (PGIs) as ingredients, Official Journal of the European Union, C 341:3-4.

[45] Orden de 6 de mayo de 1993 por la que se aprueba el Reglamento de la Denominación de Origen "*Queso Zamorano*" y su Consejo Regulador, BOE-A-1993-13159, 15311–15316.

[46] Junta de Castilla y León (2011). *Pliego de condiciones de la DOP "Queso Zamorano"* (1st Rev.)

[47] ISO 21543:2006 (IDF 201:2006) – *Milk products – Guidelines for the application of near infrared spectrometry*, International Organization for Standardization, Geneva, 1st ed., 2006.

[48] ISO 21543:2020 (IDF 201:2020) – *Milk and milk products – Guidelines for the application of near infrared spectrometry*, International Organization for Standardization, Geneva, 2nd ed., 2020.

[49] European Parliament (2021). Food Safety. *Fact Sheets on the European Union*. Retrieved from https://www.europarl.europa.eu/factsheets/en/sheet/51/food-safety on 20 September 2021.

[50] Regulation (EC) No 178/2002 of the European Parliament and of the Council of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety, Official Journal of the European Communities, L 31, 1-24.

[51] Directive 2001/18/EC of the European Parliament and of the Council of 12 March 2001 on the deliberate release into the environment of genetically modified organisms and repealing Council Directive 90/220/EEC, Official Journal of the European Communities, L 106, 1-39.

[52] Council Regulation (EEC) No 315/93 of 8 February 1993 laying down Community procedures for contaminants in food, Official Journal of the European Union, L 37, 1-5.

[53] Commission Regulation (EU) No 1881/2006 of 19 December 2006 setting maximum levels for certain contaminants in foodstuffs, Official Journal of the European Union, L 364, 5-24.

[54] Commission Regulation (EU) No 1259/2011 of 2 December 2011 amending Regulation (EC) No 1881/2006 as regards maximum levels for dioxins, dioxin-like PCBs and non-dioxin-like PCBs in foodstuffs, Official Journal of the European Union, L 320, 18-23.

[55] Regulation (EC) No 396/2005 of the European Parliament and of the Council of 23 February 2005 on maximum residue levels of pesticides in or on food and feed of plant and animal origin and amending Council Directive 91/414/EEC, Official Journal of the European Communities, L 70, 1-16.

[56] Commission Regulation (EC) No 401/2006 of 23 February 2006 laying down the methods of sampling and analysis for the official control of the levels of mycotoxins in foodstuffs, Official Journal of the European Union, L 70, p. 12.

[57] Commission Regulation (EU) No 644/2017 of 5 April 2017 laying down methods of sampling and analysis for the control of levels of dioxins, dioxin-like PCBs and non-dioxin-like PCBs in certain foodstuffs and repealing Regulation (EU) No 589/2014 Official Journal of the European Union, L 92, 9-34.

[58] Commission Regulation (EC) No 333/2007 of 28 March 2007 laying down the methods of sampling and analysis for the control of the levels of lead, cadmium, mercury, inorganic tin, 3-MCPD and benzo(a)pyrene in foodstuffs, Official Journal of the European Union, L 88, 29-38.

[59] Council Directive 96/23/EC of 29 April 1996 on measures to monitor certain substances and residues thereof in live animals and animal products, Official Journal of the European Communities, L 125, 10-32.

[60] Regulation (EC) No 470/2009 of the European Parliament and of the Council of 6 May 2009 laying down Community procedures for the establishment of residue limits of pharmacologically active substances in foodstuffs of animal origin, Official Journal of the European Union, L 152, 11-22.

[61] Commission Regulation (EU) No 37/2010 of 22 December 2009 on pharmacologically active substances and their classification regarding maximum residue limits in foodstuffs of animal origin, Official Journal of the European Union, L 15, 1-72.

[62] Commission Implementing Regulation (EU) 2021/808 of 22 March 2021 on the performance of analytical methods for residues of pharmacologically active substances used in food-producing animals and on the interpretation of results as well as on the methods to be used for sampling and repealing Decisions 2002/657/EC and 98/179/EC, Official Journal of the European Union, L 180, 84-109.

[63] Commission Decision 2002/657/EC of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Official Journal of the European Communities, L 221, 8-36.

[64] European Commission (2021). Food Safety. *Food Contact Materials*. Retrieved from https://ec.europa.eu/food/safety/chemical-safety/food-contact-materials_en on 25 September 2021.

[65] Regulation (EC) No 1935/2004 of the European Parliament and of the Council of 27 October 2004 on materials and articles intended to come into contact with food and repealing Directives 80/590/EEC and 89/109/EEC, Official Journal of the European Union, L 338, p. 4.

[66] European Commission (2021). Food Safety. *Legislation.* Retrieved from https://ec.europa.eu/food/safety/chemical-safety/food-contact-materials/legislation_en#plastic_materials on 23 September 2021.

[67] Commission Regulation (EU) No 10/2011of 14 January 2011 on plastic materials and articles intended to come into contact with food, Official Journal of the European Union, L 12, 1-89.

[68] Commission Implementing Regulation (EU) No 321/2011 of 1 April 2011 amending Regulation (EU) No 10/2011 as regards the restriction of use of Bisphenol A in plastic infant feeding bottles, Official Journal of the European Union, L 87, 1-2.

[69] Regulation No 609/2013 of 12 June 2013 of the European Parliament and of the Council on food intended for infants and young children, food for special medical purposes, and total diet replacement for weight control and repealing Council Directive 92/52/EEC, Commission Directives 96/8/EC, 1999/21/EC, 2006/125/EC and 2006/141/EC, Directive 2009/39/EC of the European Parliament and of the Council and Commission Regulations (EC) No 41/2009 and (EC) No 953/2009, Official Journal of the European Union, L 181, p. 35.

[70] Commission Regulation (EU) 2018/213 of 12 February 2018 on the use of bisphenol A in varnishes and coatings intended to come into contact with food and amending Regulation (EU) No 10/2011 as regards the use of that substance in plastic food contact materials, Official Journal of the European Union, L 41, 6-12.

[71] S. Bratinova, B. Raffael, C. Simoneau, *Guidelines for performance criteria and validation procedures of analytical methods used in controls of food contact materials - EUR 24105 EN (JRC53034)*, Publications Office of the European Union, Luxembourg, 1st ed., 2009.

[72] ECHA (European Chemicals Agency). *Annex XVII to REACH – Conditions of restriction. Restrictions on the manufacture, placing on the market and use of certain dangerous substances, mixtures and articles*. Entry 66, p. 1.

[73] ECHA (European Chemicals Agency). All news. *Bisphenol S has replaced bisphenol A in thermal paper* (ECHA/NR/20/22). Retrieved from https://echa.europa.eu/es/-/bisphenol-s-has-replaced-bisphenol-a-in-thermal-paper on 04 October 2021.

[74] Directive 2009/48/EC of the European Parliament and of the Council of 18 June 2009 on the safety of toys, Official Journal of the European Union, L 170, 1-37.

[75] Commission Directive (EU) 2017/898 of 24 May 2017 amending, for the purpose of adopting specific limit values for chemicals used in toys, Appendix C to Annex II to Directive 2009/48/EC of the European Parliament and of the Council on the safety of toys, as regards bisphenol A, Official Journal of the European Union, L 138, 128-130.

[76] Commission Regulation (EC) No 1895/2005 of 18 November 2005 on the restriction of use of certain epoxy derivatives in materials and articles intended to come into contact with food, Official Journal of the European Union, L 302, 28-32.

## CHAPTER 2

# ENSURING FOOD QUALITY IN REGULATED PRODUCTS THROUGH MULTIVARIATE METHODOLOGIES: PREDICTION OF *QUESO ZAMORANO*'S QUALITY BY NEAR-INFRARED SPECTROSCOPY ASSESSING FALSE NON-COMPLIANCE AND FALSE COMPLIANCE AT MINIMUM PERMITTED LIMITS STATED BY PROTECTED DESIGNATION OF ORIGIN REGULATIONS

## 2.1. Introducción[1]

Este capítulo gira en torno a cómo garantizar la Calidad Alimentaria mediante metodologías que permitan verificar estadísticamente el cumplimiento de ciertos atributos que, en suma, hacen que un determinado producto se diferencie del resto de los de su clase por sus excepcionales características, vinculadas tanto a su origen geográfico como a su saber hacer tradicional. En particular, se presentará a continuación un método de análisis rápido y no destructivo mediante espectroscopía NIR y regresión PLS para determinar simultáneamente el contenido de grasa, extracto seco y proteína total en *Queso Zamorano*, de acuerdo con el estándar internacional ISO 21543:2006 sobre productos lácteos y espectrometría NIR, en vigor en el momento en el que se realizó la experimentación y el posterior análisis de datos.

Este trabajo no recoge simplemente un caso más de calibración multivariante a partir de señales espectrales: el producto *Queso Zamorano* está oficialmente registrado como PDO-ES-0089, una de las Denominaciones de Origen Protegidas (DOP; en inglés, PDO (*Protected Designation of Origin*)) de la Unión Europea. Un producto alimenticio con tal mención debe satisfacer un conjunto de atributos detallados en el reglamento específico de su DOP, lo que hace obligatorio el control de su calidad. Algunos de estos atributos son cuantitativos y hacen referencia al balance de varios macronutrientes y de diversas propiedades físico-químicas en el producto final. En este sentido y para el caso particular del *Queso Zamorano*, los porcentajes en peso de proteína total, de extracto seco y de grasa sobre extracto seco en este producto al término de su maduración tienen fijados unos límites mínimos permitidos en el Reglamento de la Denominación de Origen "*Queso Zamorano*" y de su Consejo Regulador. Por lo tanto, los modelos de calibrado PLS resultantes no estarán únicamente enfocados a predecir las cantidades de los macronutrientes del *Queso Zamorano* sin necesidad de ejecutar tediosos análisis de referencia, sino, lo que es más importante, a evaluar la capacidad de detección de cada uno de esos macronutrientes con respecto a su límite mínimo permitido correspondiente.

Así, habiendo considerado como representativo un conjunto de 42 quesos clasificados como pertenecientes a la DOP "*Queso Zamorano*", se estimaron sus porcentajes en peso de grasa, extracto seco, proteína total y grasa sobre extracto seco mediante regresión PLS a partir de los datos obtenidos por espectroscopia de transmitancia en el infrarrojo cercano (*near-infrared transmittance* (NIT)) y de los valores de referencia de esas variables respuesta calculados mediante métodos validados descritos en estándares ISO. El número óptimo de

[1] Realicé la totalidad de la experimentación detallada en este capítulo durante una estancia predoctoral en el laboratorio de análisis físico-químico de la *Estación Tecnológica de la Leche* (Palencia, España), entidad acreditada de acuerdo con la normativa UNE-EN ISO/IEC 17025 para llevar a cabo análisis de referencia de productos lácteos de Castilla y León.

variables latentes en cada modelo fue seleccionado por validación cruzada, teniendo cuidado de incluir todas las réplicas de un mismo queso en el mismo grupo de cancelación para evitar el sobreajuste del modelo. Por otro lado, el preprocesado de los datos de absorbancia en la región espectral del infrarrojo cercano, necesario para este tipo de información, se llevó a cabo siguiendo 8 pretratamientos distintos, lo que supuso la estimación de 8 modelos PLS para cada respuesta. La comparación de las características de esos 8 modelos concluyó que la segunda derivada del resultado de la aplicación del método Savitzky-Golay a los espectros NIT originales con un ancho de ventana de 9 puntos y habiendo ajustado un polinomio de segundo orden seguida de una transformación SNV (*standard normal variate*) de dicha derivación era la mejor opción de preprocesado, puesto que, en general, daba lugar al modelo PLS más coherente, esto es, a aquel que explicaba el mayor porcentaje de varianza de la variable respuesta correspondiente con el menor número de variables latentes y sin necesidad de haber eliminado del conjunto de aprendizaje un número excesivo de datos anómalos. Concretamente, los modelos NIT+PLS finales presentan valores de la raíz cuadrada del error en validación cruzada que varían entre 0.41 y 0.76, entre 92.93% y 96.57% para el caso del coeficiente de determinación y entre 0.66% y 1.05% si se trata de la media en valor absoluto del error relativo en las muestras de calibración. Adicionalmente, se contrastó la veracidad de dichos modelos con un 99% de confianza, no habiéndose encontrado evidencia alguna de que los resultados obtenidos a partir de aquellos eran estadísticamente diferentes a los generados por el correspondiente método de referencia seguido hasta ese momento en el laboratorio de análisis físico-químico de la *Estación Tecnológica de la Leche*.

Por último, se verificó que la nueva metodología NIT+PLS desarrollada durante este trabajo era capaz de asegurar el control de la calidad de *Queso Zamorano*, definida, entre otras cosas, mediante límites mínimos demandados para el contenido en proteína total, extracto seco y grasa sobre extracto seco. Para ello, fue necesario adaptar los conceptos de *límite de decisión* ($CC\alpha$) y *capacidad de detección* ($CC\beta$) cuando se estiman modelos multivariantes al caso de la existencia de un límite mínimo permitido, resultando así $CD\alpha$ y $CD\beta$, respectivamente. Dada la buena calidad de los calibrados NIT+PLS estimados, el porcentaje en peso de cada macroconstituyente de *Queso Zamorano* bajo control que podía ser significativamente diferenciado de su respectivo límite mínimo permitido tenía un valor muy próximo a este último, siendo la probabilidad de falso no cumplimiento ($\alpha$) igual a 0.05 y la de falso no cumplimiento ($\beta$) igual o menor a 0.05. Esta fue la primera ocasión hasta el momento en la que esta aproximación se usaba en el control de la calidad de un producto alimenticio protegido por una DOP de la Unión Europea.

Como laboratorio autorizado por parte del Consejo Regulador de la DOP "*Queso Zamorano*" para realizar análisis físico-químicos de este producto, los modelos PLS

resultantes de esta investigación están siendo actualmente usados de manera rutinaria y periódicamente revisados en la *Estación Tecnológica de la Leche*.

## 2.2. Prediction of *Zamorano* cheese quality by near-infrared spectroscopy assessing false non-compliance and false compliance at minimum permitted limits stated by Designation of Origin regulations[2]

### 2.2.1. Abstract

Near-infrared transmittance spectroscopy has been used to predict the percentage in weight of the fat, dry matter, protein and fat-to-dry matter contents in *Zamorano* cheeses, protected with a Designation of Origin by the European Union. A total of 42 cheeses submitted to official control have been previously analysed by reference methods. Samples have been next scanned (850-1050 nm) and predictive equations have been developed using partial least-squares regression with a cross-validation step. Eight different spectra pretreatments have been firstly considered, being the most appropriate one that performing the second derivative (using a Savitzky–Golay method with a nine-point window and a second-order polynomial) followed by a standard normal variate transformation. In this case, the values obtained for the root mean square error in cross-validation, the coefficient of determination and the mean of the absolute value of relative errors have been, respectively, for fat (0.62; 95.37%; 1.05%), dry matter (0.76; 96.03%; 0.83%), protein (0.41; 96.57%; 0.81%) and the fat-to-dry matter ratio (0.61; 92.93%; 0.66%). At 99% confidence level, the trueness of the PLS models for fat, dry matter and protein has been verified.

As the official regulation for *Zamorano* cheese demands minimum permitted limits on the percentage in weight of total protein (25%), dry matter (55%) and the ratio of fat to dry matter (45%), both the decision limit and the detection capability at these minimum values have been evaluated. The adaptation of the decision limit and the detection capability to the case of a minimum permitted limit (*CDα* and *CDβ*, respectively) when a partial least-squares calibration is used has been applied for the first time to a food product protected by a Designation of Origin. The values of *CDα* with a probability of false non-compliance equal to 0.05 and of *CDβ* when, in addition, the probability of false compliance was equal to or less than 0.05 have been, respectively, 24.78% and 24.57% for protein, 54.14% and 53.28% for dry matter, and 44.39% and 43.78% for the fat-to-dry matter ratio.

### 2.2.2. Introduction

Infrared (IR) spectroscopy has currently become one of the most common techniques for a wide range of analyses in various industries due to the achievement of a fast non-

[2] This section is published as *M.L. Oca et al. / Talanta 99 (2012) 558-565* (9 citations up to August 30, 2021 according to *Scopus*).

destructive method of analysis and the rapid development of multivariate calibration techniques.

Regarding dairy industry, both near-infrared (NIR) and mid-infrared (MIR) spectroscopy have been tested to determine their potential in acquiring information on process monitoring, determination of quality, geographical origin and adulteration of dairy products in processes such as milk, milk powder, butter and cheese production. Despite the low sensitivity of the NIR region [1], it is possible to use these signals to determine the concentration of nearly all major constituents of dairy products such as water, protein, fat and carbohydrates using absorption spectroscopy with sufficient accuracy. In particular, NIR spectroscopy [2,3,4,5] has been more widely used than MIR spectroscopy for cheese composition determination [6,7,8], since the radiation light of MIR, in spite of its better specificity, has a very short penetration depth (usually a few micrometres) and cannot go through glass, plastics and other materials. On the contrary, most packaging stuff is transparent to NIR light [9].

Since dairy products contain a lot of important intrinsic fluorophores such as vitamin A and the aromatic aminoacids tryptophan, tyrosine and phenylalanine in proteins, the potential of front-face fluorescence spectroscopy to succeed in cheese evaluation has also been investigated [10,11,12], in some cases coupled with an IR spectroscopic technique [13,14]. Ref. [15] is a review of recent developments in this field. On the other hand, nuclear magnetic resonance (NMR) has proved to be a versatile spectroscopic technique in dairy research [12,16], since some processes such as pressure, heating or changes in pH, which alter the milk protein conformation and/or the aggregation state, can be studied by NMR [17].

However, for most food samples, all this chemical information is obscured by changes in the spectra caused by physical properties such as the particle size of powders. This means that each of these techniques and, in particular NIR spectroscopy, becomes a comparative strategy requiring calibration against a reference method for the constituent of interest [18]. Nowadays, most NIR spectroscopic applications are carried out by using Partial Least Squares (PLS) regression [19,20]. This methodology was devised to find a few linear combinations (latent variables) of the spectral intensities, $\mathbf{X}$, in order to explain the values of the reference method, $\mathbf{y}$, given that PLS is a biased regression method to achieve the highest prediction capacity. The $m$-th latent variable, $\mathbf{lv}_m$, is the result of maximizing the product $\text{corr}^2(\mathbf{y}, \mathbf{X\alpha})\cdot\text{Var}(\mathbf{X\alpha})$ with the restrictions $\|\boldsymbol{\alpha}\| = 1$ and $\mathbf{lv}_i^t\mathbf{S}\boldsymbol{\alpha} = 0\ (i = 1, ..., m-1)$ to ensure that $\mathbf{X\alpha}$ is uncorrelated with all the previous linear combinations $\mathbf{X}\mathbf{lv}_i$. Regarding dairy products, the International Standard ISO 21543:2006 "Milk products –

Guidelines for the application of near infrared spectrometry" governs the performance criteria of this type of analyses with multivariate calibration techniques [21].

Council Regulation (EC) No. 510/2006 of 20 March 2006 [22] lays down the definitions for Protected Designation of Origin (PDO) and Protected Geographical Indication, the two regulatory figures in a framework of protection applied on agricultural products and foodstuffs. For the purpose of this Regulation, "designation of origin" (DO) means the name of a region, a specific place or, in exceptional cases, a country, used to describe an agricultural product or a foodstuff: *i)* originated in that region, specific place or country, *ii)* whose quality or characteristics are essentially or exclusively due to a particular geographical environment with its inherent natural and human factors, and *iii)* whose production, processing and preparation take place in the defined geographical area.

*Zamorano* cheese is protected with a DO since 1993 and registered as PDO in Commission Regulation (EC) No. 1107/96 [23]. According to the Order of 6 May 1993 (BOE No. 120 of 20 May 1993) [24], that name refers to a compressed-paste fatty cheese made from milk of ewes of the Spanish *Churra* and *Castellana* breeds and with a ripening step of at least 100 days' duration.

This work was aimed to develop a simple and rapid analytical tool for monitoring simultaneously the fat, dry matter and protein contents in *Zamorano* cheese samples by near-infrared transmittance (NIT) spectroscopy. Calibration models based on a cross-validated PLS regression of the percentage in weight of every constituent on the recorded spectra were estimated. Several pretreatments of the raw signals were first evaluated by comparing the resulting PLS models in order to achieve the best correlations with the values of every constituent obtained by reference methods. Finally, trueness was assessed, and the minimum contents of protein, dry matter and the fat-to-dry matter ratio according to [24] were calculated in terms of detection capability of the corresponding procedure by evaluating the probabilities of false non-compliance and false compliance. Currently, these calibration models are being applied systematically at *Estación Tecnológica de la Leche* to accredit the NIT+PLS procedure for the routine quality control of *Zamorano* cheese.

## 2.2.3. Materials and methodology

### *2.2.3.1. Cheese samples*

Forty-two *Zamorano* cheeses that dated either 2010 or 2011 were randomly supplied by the Regulatory Council of this PDO from all the producers included in it. Cheese samples were vacuum-packaged and stored at frozen conditions (−20 ± 5 ºC) until analysis. Sampling

and subsequent analyses were performed in accordance with the general recommendations included in the International Standard ISO 21543:2006 [21].

#### *2.2.3.2. Reference analyses*

All reference analyses were carried out in *Estación Tecnologica de la Leche* (*Instituto Tecnológico Agrario de Castilla y León*, ITACyL) in Palencia (Spain), which is an accredited laboratory for the performance of such determinations according to the criteria established in the regulation UNE-EN ISO/IEC 17025 [25].

The determination of the fat, dry matter and protein contents was achieved following official methods [26,27,28]. As a step in their official accreditation, the uncertainty of these analytical methods had been previously estimated as standard deviation and resulted in: *i)* 0.36% when the fat percentage in weight ranged from 0.90% to 45.0%; *ii)* 0.64% when the dry matter percentage in weight varied from 20.0% to 80.0%; and *iii)* 0.41% when the protein percentage in weight ranged from 6.0% to 30.0%.

All the reference analyses were performed in duplicate. For every cheese analysed, the mean of these two replicates was used as the final value of the fat, dry matter or protein content.

#### *2.2.3.3. Near-infrared spectroscopy*

All samples were allowed to reach room temperature (25 ± 2 °C) before analysis. Then each sample was grated after being removed from its plastic package and once all outer portions had been cut off. Approximately 40 g of each homogeneous grated cheese were softly packed into a 100 mm-diameter Petri dish in order to achieve an air hole-free optical path length of about 10 mm. To minimise sampling error, all samples were analysed in triplicate, except for two of them, which were in duplicate.

These samples were measured using NIT mode on a FOSS FoodScan Lab spectrophotometer. Because of its effect on spectral response [29], the temperature of the spectroscopic measurements was controlled within the range 26−30 ºC. NIT spectra were collected from 850 to 1050 nm at 2 nm intervals in a log $1/T$ format, where $T$ is the sample transmittance. Each spectrum was an average of 16 sub-spectra recorded at sixteen different points by rotating the Petri dish automatically in the analyser. Therefore, 124 NIT spectra (shown in Figure 2.1.a) were finally obtained, which made up the calibration set.

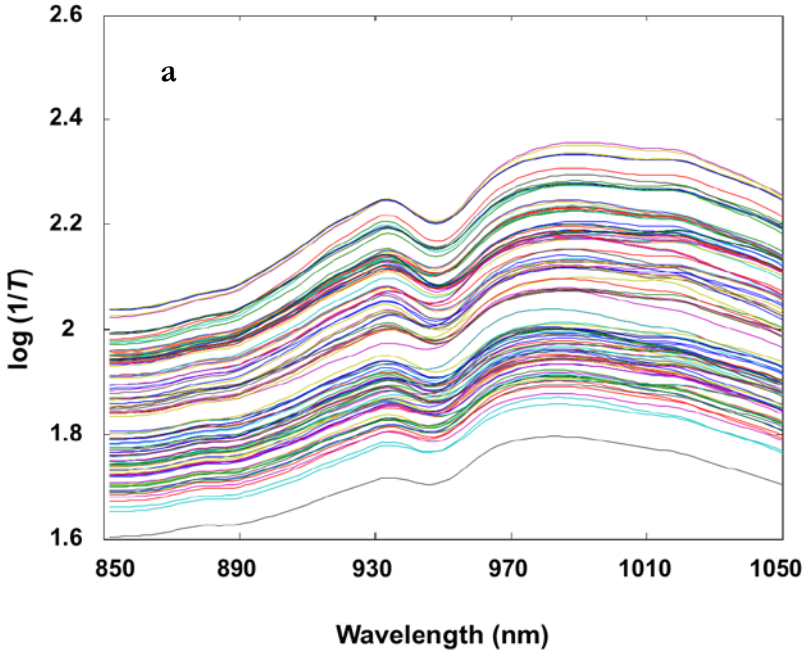


**Figure 2.1.a.** Raw NIT spectra of the 124 *Zamorano* cheese samples in the original calibration set. On the *x*-axis, the spectral region (from 850 to 1050 nm). On the *y*-axis, the logarithm of the inverse of the transmittance *T* of the sample.

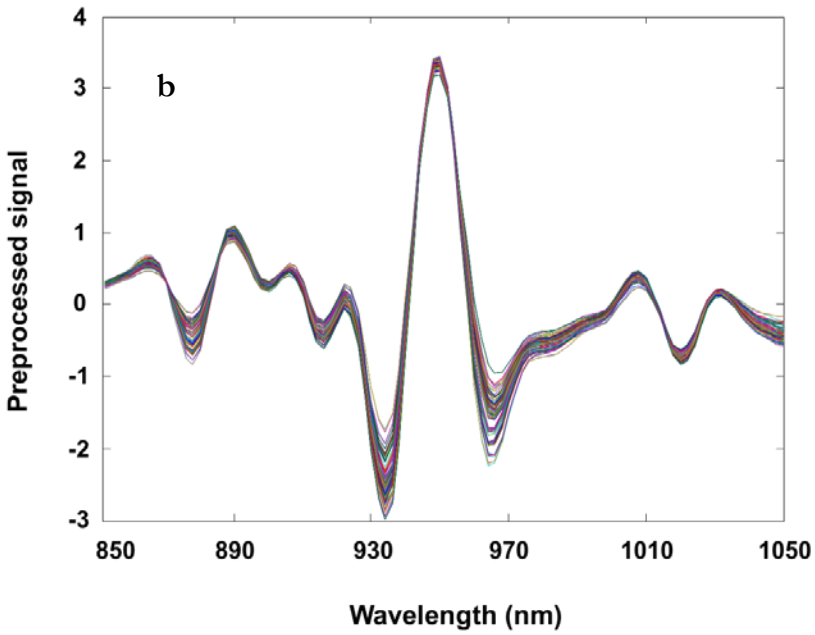


**Figure 2.1.b.** Preprocessed spectra by means of the Savitzky-Golay second derivative (nine-point window, second-order polynomial) of the 124 raw signals followed by a SNV transformation.

### *2.2.3.4. Chemometrics: multivariate analysis*

#### 2.2.3.4.1. Data preprocessing

Both a slightly curved baseline and a baseline offset can be seen in Figure 2.1.a, which made preprocessing necessary to transform the data in such a way that the multivariate signals would better adhere to Beer's law. Eight pretreatments were evaluated by comparing the calibration models estimated from the corresponding preprocessed data:

a) Standard normal variate (SNV) is an autoscaling of spectra (rows of the data matrix), so a reference spectrum is not needed and the pretreatment for each sample is independent of the others. However, as there is not a regression step, SNV does not remove noise [9]. This preprocessing is weighted towards considering the spectrum values that deviate from the mean more heavily than those near the mean.

b) Derivatives are a common method used to remove the unimportant baseline signal from spectra by taking the derivative with respect to the variable number. This method is adequate for NIR signals because variables are strongly related to each other and the adjacent variables contain similar correlated signals. The first derivative of every spectrum has been carried out using the Savitzky–Golay (SG) method [30] with a window of nine points and a second-order polynomial. This preprocessing procedure will be indicated as SG (9,2,1) in the following.

c) Second derivative of the NIT signals also using the SG technique (9-point window and second-order polynomial), named SG (9,2,2).

d) SNV followed by detrending. Detrend fits a polynomial (of second order in this work) to the entire spectra containing both baseline and signal and subtracts this polynomial. As such, it works optimally when the largest source of spectral signal in each sample is background interference. The detrend method was introduced along with the SNV transformation by Barnes et al [31].

e) SNV followed by SG (9,2,1).

f) SNV followed by SG (9,2,2).

g) SG (9,2,1) followed by SNV.

h) SG (9,2,2) followed by SNV.

2.2.3.4.2. Partial Least Squares regression

Data were arranged in a matrix **X** with dimensions (124×100), where 124 referred to the total number of samples analysed and 100 did to the set of wavelengths (predictor variables) recorded. For each pretreatment, four PLS models were estimated where the response **y** was the fat, dry matter, protein and the fat-to-dry matter percentage in every cheese sample, respectively. The values of the first three responses had been previously obtained by means of reference analytical methods (see Section 2.2.3.2), while those for the fat-to-dry matter ratio were next calculated by dividing the corresponding percentages obtained for these two constituents.

For the cross-validation step, three cancellation groups were chosen so that the three (or two) replicates that had been assigned the same reference value always lay in the same group as they referred to the same cheese; otherwise, there would have been a risk of overfitting.

As stated in [32,33,34], the procedure followed for the PLS regression in every instance was:

a) Preprocess **X** data with the procedures a) to h) commented in Section 2.2.3.4.1 and autoscale the response **y**.

b) Determine the optimum number of latent variables by plotting the root mean square error in cross-validation (*RMSECV*) *versus* the number of latent variables in the model. Generally, the best solution should be the one yielding the lowest *RMSECV* from the fewest latent variables, seeking to ensure this *RMSECV* value is not lower than that of the uncertainty for the corresponding reference method to avoid overfitting the model.

c) Remove those samples with standardized residual (in absolute value) greater than 2.5 (considered as **y**-outliers) and/or with both $Q$ and Hotelling's $T^2$ values higher than their corresponding thresholds at 95% confidence level (considered as outliers in the calibration subspace). The $Q$ residual index indicates the difference or residual between the value of the sample and its projection onto the subspace of the model, while the Hotelling's $T^2$ statistic is a measurement of the Mahalanobis distance from each sample to the centroid, measured in the projection plane (hyperplane) of the model considered.

d) Repeat steps b) and c) until there are no outliers of any kind.

After a PLS model had been obtained for each response, with the aim of checking the validity of the method, a least-squares (LS) linear regression between the PLS calculated

values of the response and the values obtained by the corresponding official method was performed in each case. The LS regression is unbiased and with the lowest variance providing that residuals have the same normal distribution in all samples. But this distributional property of the residuals is false if some datum lies outside the linear tendency, and the good inferential properties of the LS regression are cancelled. This could occur if there were outlier samples with erroneous values either calculated from the PLS model or wrongly determined from reference analyses. Therefore, an outlier detection step based on a least median of squares (LMS) regression [35,36] is previously necessary, which provides an objective and robust criterion to detect outliers. Every sample whose standardized residual and/or diagnostic resistance with regard to the LMS model is higher than 2.5 in absolute value will be thus removed from the data set. The final LS linear regression, actually called *reweighted LS regression*, of predicted values of the response *versus* true values will be carried out with the remaining objects.

Finally, two hypothesis tests enabled to verify the trueness of the method by checking if, at a significance level $\alpha$, there were no statistically significant differences between the values obtained, respectively, for the slope ($b_1$) and 1, and the intercept ($b_0$) and 0. In such a case, it could be concluded that the trueness of the PLS method built was confirmed at the confidence level 1 - $\alpha$.

2.2.3.4.3. Assessing the minimum permitted limit with a multivariate signal

When, as in this work, a minimum permitted limit $x_0$ has been established for a substance, the following one-tailed hypothesis test is posed about the presence of the analyte in the problem sample:

$H_0$: $x \geq x_0$ (the concentration of the analyte is greater than or equal to $x_0$). (2.1)

$H_a$: $x < x_0$ (the concentration of the analyte is lower than $x_0$).

The decision limit ($CD\alpha$) of the method is the concentration below which it can be decided with a statistical certainty of 1 – $\alpha$ that the minimum permitted limit has been truly exceeded. $CD\alpha$ is related with the probability of false non-compliance, $\alpha$. That is, $\alpha$ is the probability of concluding that the tested sample is not compliant when in fact it is (false non-compliance decision or type I error). Formally, $\alpha = \text{pr}\{\text{reject } H_0 / H_0 \text{ is true}\}$.

But it is also necessary to assess the probability, $\beta$, of false compliant decision (type II error). $\beta$ is the probability of affirming that the tested sample has a concentration of the analyte greater than or equal to $x_0$, i.e., to state that it is compliant, when it is not actually. Formally, $\beta = \text{pr}\{\text{accept } H_0 / H_0 \text{ is false}\}$.

Therefore, the capability of detection ($CD\beta$) of the method is the concentration above which it can be assessed that the probability of false compliance is $\beta$ and that of false non-compliance is $\alpha$. $CD\beta$, which is the critical value of the hypothesis test in Eq. (2.1), depends on $\alpha$, $\beta$, the number of replicate measurements in the test sample and both the sensitivity and the precision of the method [37].

Once the probability $\alpha$ has been established, the plot of $\beta$ *versus* $CD\beta$ represents the operative curve of the hypothesis test in Eq. (2.1). The latter shows the capacity of the method to discriminate a specific quantity with regard to the minimum permitted limit.

Analytical methods provide signals and both $CD\alpha$ and $CD\beta$ are concentrations, so a calibration curve is necessary to relate signals and concentrations. If $x_0$ is either null or the maximum limit permitted for a compound, the null hypothesis $H_0$ in Eq. (2.1) is formulated as "$x \leq x_0$ (the concentration of the analyte is equal to or lower than $x_0$)", while the alternative hypothesis $H_a$ is as "$x > x_0$ (the concentration of the analyte is greater than $x_0$)". If the analytical signal is univariate and the calibration curve is linear, then $CD\alpha$ and $CD\beta$ are named, respectively, $CC\alpha$, (decision limit) and $CC\beta$ (capability of detection). This definition of detection capability has been accepted by IUPAC [38], ISO 11843 [39] and some European regulations [40].

In the case of using multivariate and/or multiway signals, the approach as hypothesis test is the same, but the calibration model "signal equal to a linear function of concentration" is no longer valid. Instead, the concentration **y** is the response to be fitted as a function of a matrix, **X**, when signals are multivariate, or a tensor, $\underline{\mathbf{X}}$, if signals are multiway. The previous concepts of $CC\alpha$ and $CC\beta$ have been generalized for these signals with multivariate calibrations such as PLS [41] or multiway such as *n*-PLS or PARAFAC [42]. This method can also be applied to any kind of calibration function such as a neural network [43]. The use of $CD\alpha$ and $CD\beta$ for assuring a minimum value has been developed in [32], but the present work means its application for the first time to ensure a minimum value regulated in a food product. The procedure for that is based on the LS regression "*PLS-calculated concentration versus True concentration*". A tutorial about how to evaluate type I and type II errors in various kinds of chemical analyses can be consulted in [37].

#### 2.2.3.4.4. Software

The FoodScan software (FOSS, Hilleroed, Denmark) was used to acquire the spectra. Raw data were exported to MATLAB using WinISI III, version 1.60 (Infrasoft International, Port Matilda, PA, USA). PLS regressions were computed with the PLS_Toolbox [44] for use with MATLAB version 7.9.0.529 (The MathWorks, Inc.). LS regressions and hypothesis testing were done with STATGRAPHICS [45]. LMS regressions for the

detection of outliers were carried out with PROGRESS [35]. A home-made program, NWAYDET, was used to estimate $CD\alpha$ and $CD\beta$ for protein, dry matter and the fat-to-dry matter ratio.

### 2.2.4. Results and discussion

#### *2.2.4.1. Construction of PLS models for fat, dry matter and protein*

The whole set of 124 NIT spectra was selected as the predictor **X** to develop PLS models following the procedure explained in Section 2.2.3.4.2. According to the results provided by the reference methods, percentages in weight of the three constituents analysed in that way ranged: *i)* from 33.30% to 42.94% for fat; *ii)* from 63.95% to 77.20% for dry matter; *iii)* from 22.21% to 28.73% for protein.

A PLS model was built to estimate each of these three response variables from each one of the eight sets of pretreated NIT data. These 24 PLS models (not shown) were compared in order to select the best preprocessing method. Thus, the chosen pretreatment was SG (9,2,2) followed by SNV. On the whole, that transformation of the NIT spectra involved the lowest number of latent variables in the final model together with the highest percentage of explained variance of the response without having removed too many outlier data from the calibration set. Figure 2.1.b shows the 124 NIT spectra after the SG (9,2,2) + SNV pretreatment.

The results of the application of the PLS procedure to estimate the relation between each response and the NIT spectra are summarized in Table 2.1. For example, with regard to fat, for the initial calibration set (124 objects), the optimal *RMSECV* value (0.74) was obtained from the 7-latent variable model, with a root mean square error in calibration (*RMSEC*) of 0.60 and 93.17% of explained variance of fat. Samples with numbers 75 and 76 had standardized residuals greater than 2.5 in absolute value when that number of latent variables was considered, so they were removed from the calibration set. The process was repeated with the 122 remaining objects. The evolution in the building of the PLS model for fat is displayed in the first four rows of Table 2.1. Finally, the optimum number of latent variables was 7 for the fourth model (*RMSEC* = 0.50; *RMSECV* = 0.62; 95.37% of explained variance of fat). In this case, as no objects presented either standardized residuals higher than 2.5 in absolute value or $Q$ and Hotelling's $T^2$ values greater than their corresponding threshold values at 95% confidence level, this was considered the final PLS model for fat. With this model, the mean, median and standard deviation of the absolute value of the relative error in calibration, which ranged from 0.03% to 3.26%, were, respectively, 1.05%, 0.87% and 0.78%.

Regarding dry matter and protein contents, the same procedure was applied for the estimation of the corresponding PLS model that related each response with the pretreated NIT spectra. As can be seen in Table 2.1 (rows 5–9), as for dry matter, after the removal of the 11 outliers detected, the number of latent variables that achieved the optimum value for *RMSECV* was 5 (*RMSEC* = 0.69; *RMSECV* = 0.76; 96.03% of explained variance of dry matter). With this PLS model, the mean, median and standard deviation of the absolute value of the relative error in calibration, which ranged from 0.02% to 1.99%, were, respectively, 0.83%, 0.74% and 0.53%. On the other hand, for the protein content, once 3 outlier data had been rejected, the PLS model that fitted this response best was the one estimated from 10 latent variables (*RMSEC* = 0.27; *RMSECV* = 0.41; 96.57% of explained variance of protein), while the mean, median and standard deviation of the absolute value of the relative error in calibration, which ranged from 0.00% to 2.67%, were, respectively, 0.81%, 0.68% and 0.65%.

It must be noticed that, in every instance, the value of *RMSEC* was quite similar to that of *RMSECV*, which proves that all models were stable.

**Table 2.1**. Evolution of the development of the PLS models for the calibration of the fat, dry matter and protein contents *versus* the NIT spectra recorded.

| Response variable y | PLS model (final number of objects) | Number of LV[a] | Explained variance of X (%) | Explained variance of y (%) | Sample indexes considered outlier data | *RMSEC*[b] | *RMSECV*[c] |
|---|---|---|---|---|---|---|---|
| Fat | 1st (124) | 7 | 99.99 | 93.17 | 75 ($SR^d$ = 3.36)<br>76 ($SR$ = 3.14) | 0.60 | 0.74 |
| | 2nd (122) | 7 | 99.99 | 94.52 | 27 ($SR$ = −2.53)<br>26 ($SR$ = −2.67) | 0.54 | 0.65 |
| | 3rd (120) | 6 | 99.99 | 94.68 | 28 ($SR$ = −2.60) | 0.53 | 0.64 |
| | 4th (119) | 7 | 99.99 | 95.37 | — | 0.50 | 0.62 |
| Dry matter | 1st (124) | 4 | 99.97 | 90.33 | 21 ($SR$ = 3.51)<br>22 ($SR$ = 3.54)<br>86 ($SR$ = −2.56) | 1.08 | 1.18 |
| | 2nd (121) | 4 | 99.97 | 92.98 | 24 ($SR$ = −2.51)<br>40 ($SR$ = 2.57)<br>75 ($SR$ = 2.61)<br>76 ($SR$ = 2.70)<br>88 ($SR$ = −2.76) | 0.93 | 1.03 |
| | 3rd (116) | 4 | 99.97 | 94.83 | 38 ($SR$ = 2.74)<br>39 ($SR$ = 2.65) | 0.80 | 0.86 |
| | 4th (114) | 5 | 99.98 | 95.75 | 87 ($SR$ = −2.64) | 0.71 | 0.77 |
| | 5th (113) | 5 | 99.98 | 96.03 | — | 0.69 | 0.76 |

**Table 2.1. (cont.)**

| Response variable y | PLS model (final number of objects) | Number of LV[a] | Explained variance of X (%) | Explained variance of y (%) | Sample indexes considered outlier data | *RMSEC*[b] | *RMSECV*[c] |
|---|---|---|---|---|---|---|---|
| Protein | 1st (124) | 9 | 100 | 95.73 | 115 (*SR* = −2.89) | 0.30 | 0.44 |
| | 2nd (123) | 9 | 100 | 96.02 | 77 ($Q^e$ = 0.02; $T^{2\,e}$ = 23.92) | 0.29 | 0.43 |
| | 3rd (122) | 10 | 100 | 96.21 | 113 (*SR* = −2.55) | 0.28 | 0.42 |
| | 4th (121) | 10 | 100 | 96.57 | — | 0.27 | 0.41 |

(a) Latent variables.

(b) Root mean square error in calibration.

(c) Root mean square error in cross-validation.

(d) Standardized residual of data considered *y*-outliers.

(e) Indexes defined in Section 2.2.3.4.2.

#### *2.2.4.2. Trueness of the PLS models for fat, dry matter and protein*

After a PLS model had been estimated for each of these three responses, an LS linear regression was performed between the PLS-predicted values of every response and the values obtained by the corresponding reference method. Before the LS fitting, an LMS regression was carried out to detect possible outlier data caused by some mistake occurred either at the construction of the PLS model or during the analyses by the official methods. The results of the LMS and LS regressions and the *p*-values for the two hypothesis tests posed to check the trueness of every method are listed in Table 2.2. As all *p*-values were higher than 0.01, it could be asserted that the analytical behaviour of the PLS method was the same as that of the corresponding reference procedure at 99% confidence level. Plots of the percentage in weight of each constituent obtained from its PLS model as a linear function of the amount determined from the respective official method are illustrated in Figures 2.2.a (fat), 2.2.b (dry matter) and 2.2.c (protein).

**Table 2.2.** Results of the LMS and LS regressions performed with the values of each response calculated from the corresponding PLS model *versus* the true values.

| Response variable | LMS regression "*Predicted response versus True response*" | | | | | Reweighted LS regression "*Predicted response versus True response*" | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | Number of outliers | Outlier indexes | *SR*[a] | *DR*[b] | *n* | $b_0$ | $b_1$ | $R^2$ (%) | $s_{yx}$ | *p*-value for the test $b_0 = 0$ | *p*-value for the test $b_1 = 1$ |
| Fat | 119 objects | 4 | 10 | 2.52 | 1.75 | 115 objects | 1.59 | 0.96 | 96.16 | 0.45 | 0.02 | 0.02 |
| | | | 87 | 2.60 | 2.05 | | | | | | | |
| | | | 88 | 2.57 | 2.03 | | | | | | | |
| | | | 89 | 2.76 | 2.16 | | | | | | | |
| Dry matter | 113 objects | 0 | | | | 113 objects | 2.82 | 0.96 | 96.03 | 0.68 | 0.03 | 0.03 |
| Protein | 121 objects | 9 | 4 | 2.70 | 1.70 | 112 objects | 0.56 | 0.98 | 97.82 | 0.22 | 0.12 | 0.13 |
| | | | 24 | −2.78 | 2.01 | | | | | | | |
| | | | 50 | −3.18 | 2.15 | | | | | | | |
| | | | 51 | −2.78 | 1.91 | | | | | | | |
| | | | 52 | −2.59 | 1.80 | | | | | | | |
| | | | 74 | −3.32 | 2.21 | | | | | | | |
| | | | 82 | −2.67 | 1.78 | | | | | | | |
| | | | 99 | 2.63 | 1.71 | | | | | | | |
| | | | 112 | −2.90 | 2.02 | | | | | | | |

(a) Standardized residual of the object.

(b) Diagnostic resistance of the object.

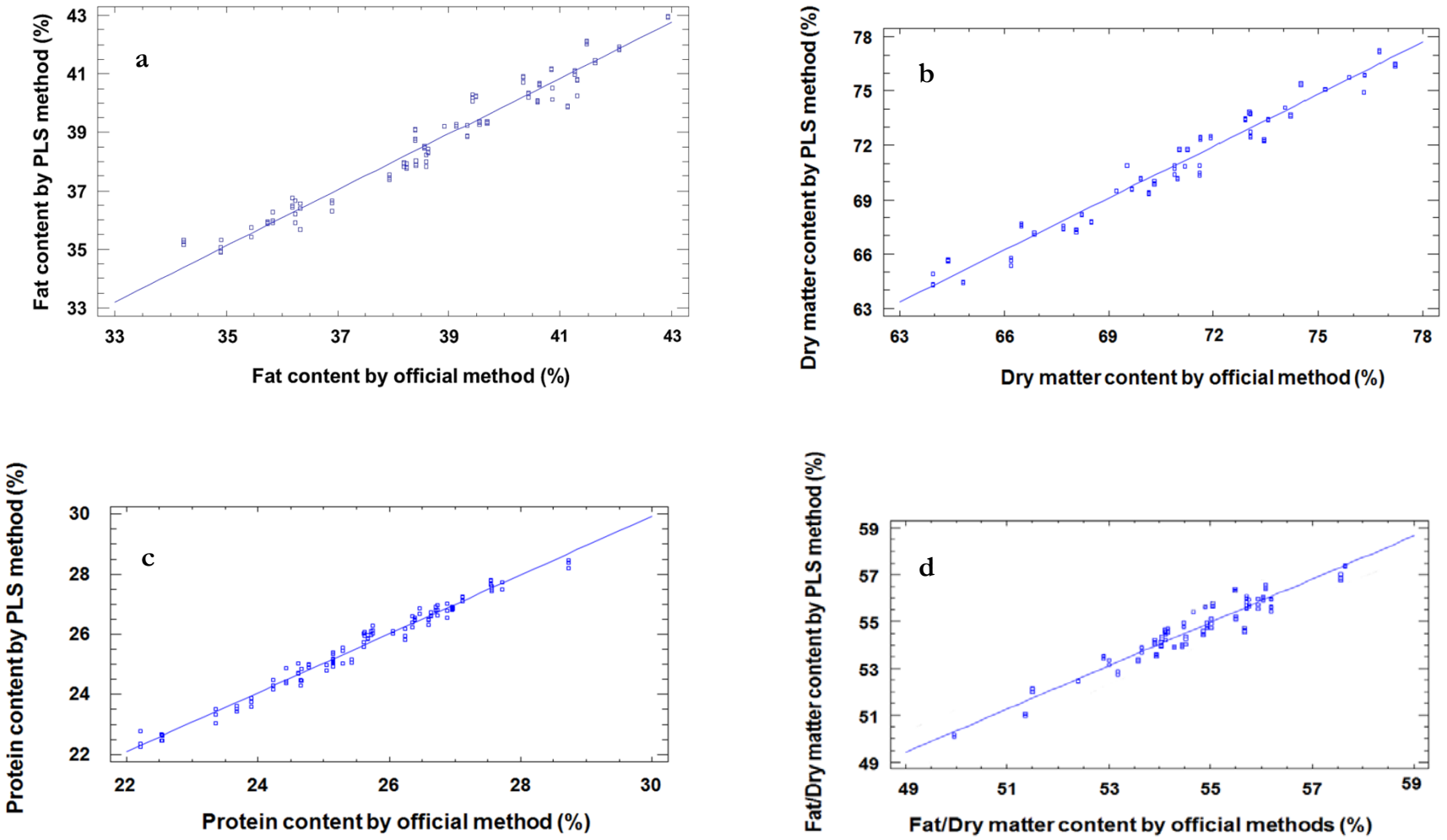


**Figure 2.2.** Regression line of the PLS-predicted values of each response *versus* the values obtained from the corresponding reference method: **a.** Fat content; **b.** Dry matter content; **c.** Protein content; **d.** Fat/dry matter content.

#### *2.2.4.3. Detection capability for protein, dry matter and fat-to-dry matter ratio*

In the chapter IV of [24], which lays down the regulations and specifications of *Zamorano* cheese, the following physicochemical properties are established for this food product:

*i)* The protein content must not be lower than 25% in weight.

*ii)* The dry matter content must not be lower than 55% in weight.

*iii)* The fat-to-dry matter ratio, expressed as percentage in weight, must not be lower than 45%.

In this context, the detection capability of every method at its respective minimum permitted level with $\alpha = \beta = 0.05$ was estimated from the corresponding reweighted LS regression *"Predicted response versus True response"* performed previously (see Section 2.2.4.2). So in each case and for a specific number of replicates, the operative curve of the hypothesis test in Eq. (2.1) was determined when $\alpha = 0.05$. This graph showed the probability of false compliance, $\beta$, as a function of the percentage in weight of the corresponding regulated constituent (protein, dry matter and fat-to-dry matter ratio) at its minimum permitted level. This nominal value was, in each case, that of $x_0$ in Eq. (2.1).

For the total protein percentage in *Zamorano* cheese, several operative curves for the nominal minimum protein content of 25% appear in Figure 2.3. Particularly, for $\alpha = \beta = 0.05$, the decision limit, $CD\alpha$, was 24.78%, and the detection capability, $CD\beta$, equalled 24.57% when 3 replicates were considered (see curve **(c)**). So the NIT+PLS-based method for total protein was able to distinguish 24.57% from 25% with probabilities of false non-compliance and false compliance equal to 0.05.

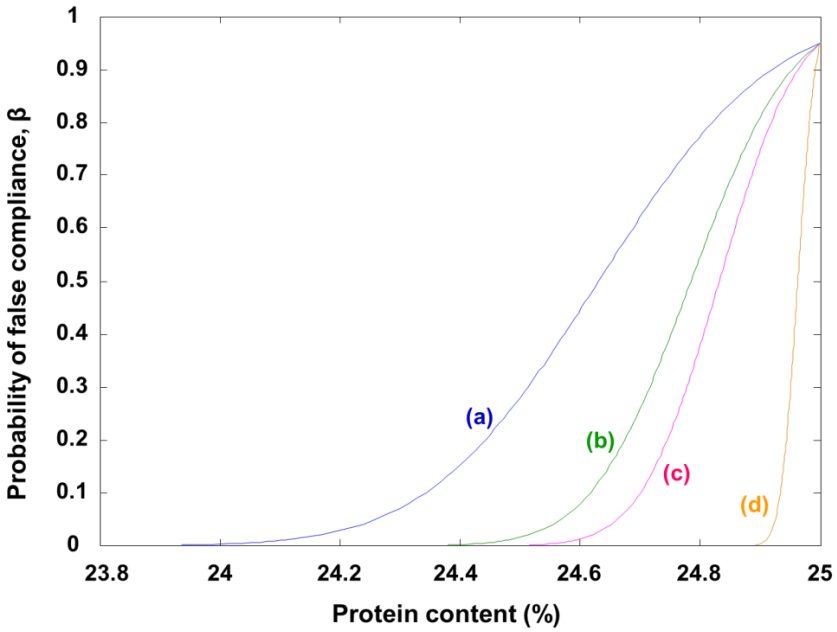


**Figure 2.3.** Probability of false compliance, $\beta$, *versus* protein percentage in weight (probability of false non-compliance, $\alpha$, fixed at 0.05). **(a)** One replicate; **(b)** two replicates; **(c)** three replicates; **(d)** infinite replicates. Nominal minimum value: 25%.

As for the dry matter percentage, Figure 2.4 shows four power curves for the minimum limit of 55% laid down in [24]. For 3 replicates and $\alpha = \beta = 0.05$, $CD\alpha$ and $CD\beta$ would be 54.14% and 53.28%, respectively (see curve **(c)**). However, since the reference values for the dry matter content in the 42 cheeses analysed ranged from 63.95% to 77.20%, all of them higher than the minimum permitted level for this constituent, both $CD\alpha$ and $CD\beta$ were extrapolated values, so they were overestimated. That is, the NIT+PLS-based method for the dry matter content was able to distinguish 53.28% from 55% with a probability of false non-compliance equal to 0.05 and a probability of false compliance lower than 0.05.

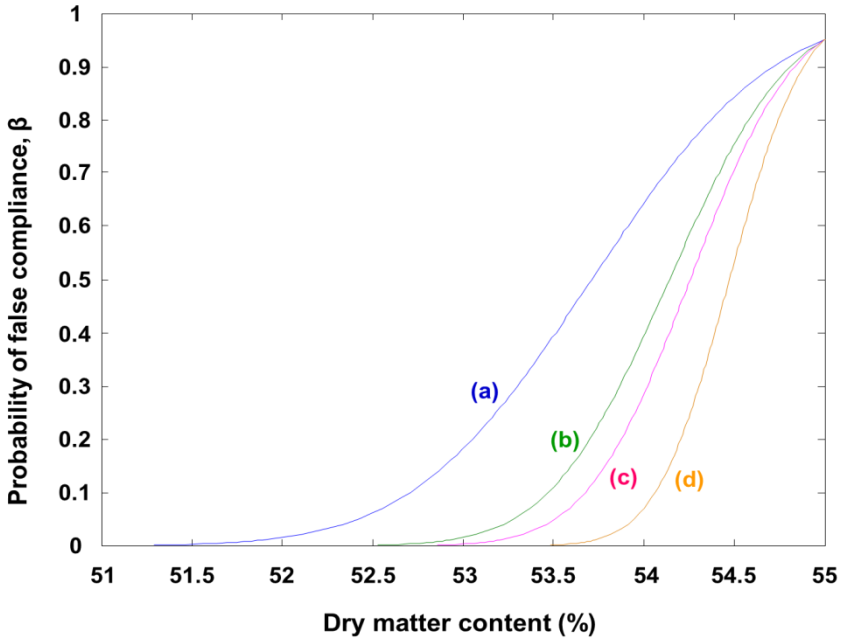


**Figure 2.4.** Probability of false compliance, $\beta$, *versus* dry matter percentage in weight (probability of false non-compliance, $\alpha$, fixed at 0.05). **(a)** One replicate; **(b)** two replicates; **(c)** three replicates; **(d)** infinite replicates. Nominal minimum value: 55%.

In the case of the fat/dry matter percentage, the PLS model that related this variable to the SG (9,2,2)+SNV-pretreated spectra was first built instead of calculating the PLS-predicted value of this ratio by dividing the corresponding values of fat and dry matter estimated from their PLS regressions. The reference percentages in weight of this ratio ranged from 49.96% to 59.87%. For this response, the evolution of the procedure designed to achieve the PLS regression is collected in Table 2.3, where it is shown that, after 11 outlier objects had been removed from the calibration set, the model that fitted the fat/dry matter ratio best was that with 8 latent variables (*RMSEC* = 0.42; *RMSECV* = 0.61; 92.93% of explained variance of the response). The mean, median and standard deviation of the absolute value of the relative error in calibration, which ranged from 0.03% to 1.98%, were, respectively, 0.66%, 0.59% and 0.43%.

**Table 2.3**. Evolution of the PLS calibration to determine the fat-to-dry matter ratio, expressed as percentage in weight, from the pretreated NIT spectra.

| PLS model (final number of objects) | Number of LV[a] | Explained variance of X (%) | Explained variance of y (%) | Sample indexes considered to be outliers | *RMSEC*[b] | *RMSECV*[c] |
|---|---|---|---|---|---|---|
| 1st (124) | 7 | 99.99 | 72.83 | 21 ($SR^d$ = −5.16)<br>22 ($SR$ = −4.98) | 0.91 | 0.99 |
| 2nd (122) | 7 | 99.99 | 83.66 | 40 ($SR$ = −2.58)<br>86 ($SR$ = 2.52)<br>87 ($SR$ = 2.87)<br>88 ($SR$ = 2.91) | 0.65 | 0.71 |
| 3rd (118) | 8 | 99.99 | 88.41 | 38 ($SR$ = −3.21)<br>39 ($SR$ = −3.33) | 0.55 | 0.62 |
| 4th (116) | 8 | 99.99 | 91.26 | 75 ($SR$ = 2.70) | 0.46 | 0.52 |
| 5th (115) | 8 | 99.99 | 91.91 | 76 ($SR$ = 2.61) | 0.44 | 0.51 |
| 6th (114) | 8 | 99.99 | 92.46 | 77 ($Q^e$ = 0.02;<br>$T^{2\,e}$ = 29.81) | 0.43 | 0.48 |
| 7th (113) | 8 | 100.00 | 92.54 | — | 0.43 | 0.49 |

(a) Latent variables.

(b) Root mean square error in calibration.

(c) Root mean square error in cross-validation.

(d) Standardized residual of data considered **y**-outliers.

(e) Indexes defined in Section 2.2.3.4.2.

Next, an LMS regression *"PLS-predicted values versus True values"* was performed, and no outlier data were found. The accuracy line next estimated had slope 0.94 and intercept 4.07 with a standard error of 0.4 and a coefficient of determination equal to 92.51%. This regression line is plotted in Figure 2.2.d. Both *p*-values for the two hypothesis tests to the trueness of the method were 4 $10^{-3}$, which meant that the PLS regression for the fat-to-dry matter ratio was biased at 99% confidence level. This bias, which was statistically significant, should be assessed in practice because it is equal to an error in the determination between 0.78% and 1.14% when using the PLS regression for the fat-to-dry matter ratio. However, if desired, the bias could be corrected using Eq. (2.2):

$$\text{Fat/Dry matter} = \frac{\text{Fat/Dry matter}_{\text{PLS predicted}} - b_0}{b_1} = \frac{\text{Fat/Dry matter}_{\text{PLS predicted}} - 4.07}{0.93} \qquad (2.2)$$

Lastly, the detection capability of the method for the fat-to-dry matter percentage at the minimum permitted level of 45% was estimated with $\alpha = \beta = 0.05$. The operative curves determined for several numbers of replicates and $\alpha = 0.05$ are depicted in Figure 2.5. If 3 replicates were specifically considered and probabilities of false non-compliance, $\alpha$, and false compliance, $\beta$, were fixed at 0.05, $CD\alpha$ would be 44.39%, while $CD\beta$ would equal 43.78% (see curve **(c)**).

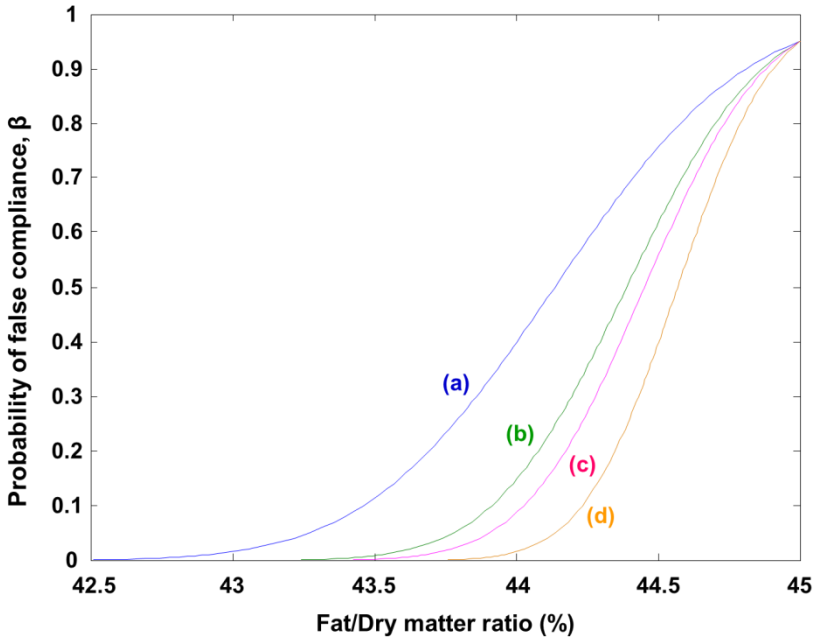


**Figure 2.5.** Probability of false compliance, $\beta$, *versus* fat/dry matter percentage in weight (probability of false non-compliance, $\alpha$, fixed at 0.05). **(a)** One replicate; **(b)** two replicates; **(c)** three replicates; **(d)** infinite replicates. Nominal minimum value: 45%.

The reference values of the fat-to-dry matter ratio of the 113 samples comprising the final calibration set varied between 50.0% and 57.5%, which were higher than the minimum permitted quantity (45%). As a consequence, as happened in the case of the dry matter content, the conclusion was that the NIT+PLS-based method for the fat-to-dry matter ratio was able to distinguish 43.78% from 45% with a probability of false non-compliance equal to 0.05 and a probability of false compliance lower than 0.05.

### 2.2.5. Acknowledgements

The authors thank the financial support provided by Ministerio de Ciencia e Innovación (CTQ2011-26022) and Junta de Castilla y León (BU108A11-2) M.L. Oca is particularly grateful to Universidad de Burgos for her FPI grant.

## 2.2.6. References

[1] J.L. Rodríguez-Otero, M. Hermida, J. Centeno, *Analysis of Dairy Products by Near-Infrared Spectroscopy: A Review*, Journal of Agricultural and Food Chemistry 45 (1997) 2815-2819.

[2] J.L. Rodríguez-Otero, M. Hermida, A. Cepeda, *Determination of Fat, Protein, and Total Solids in Cheese by Near-Infrared Reflectance Spectroscopy*, Journal of AOAC INTERNATIONAL. 78 (1995) 802-806.

[3] M.M. Pierce, R.L. Wehling, *Comparison of sample handling and data treatment methods for determining moisture and fat in Cheddar cheese by near-infrared spectroscopy*, Journal of Agricultural and Food Chemistry 42 (1994) 2830-2835.

[4] C. Wittrup, L. Nørgaard, *Rapid Near Infrared Spectroscopic Screening of Chemical Parameters in Semi-hard Cheese Using Chemometrics*, Journal of Dairy Science 81 (1998) 1803-1809.

[5] D. McKenna, J. AOAC Int. 84 (2001) 623–628.

[6] T. Woodcock, C.C. Fagan, C.P. O'Donnell, *Application of Near and Mid-Infrared Spectroscopy to Determine Cheese Quality and Authenticity*, Food and Bioprocess Technology 1 (2008) 117-129.

[7] D.H. McQueen, R. Wilson, A. Kinnunen, E.P. Jensen, *Comparison of two infrared spectroscopic methods for cheese analysis*, Talanta 42 (1995) 2007-2015.

[8] M. Chen, J. Irudayaraj, D.J. McMahon, *Examination of full fat and reduced fat Cheddar cheese during ripening by Fourier transform infrared spectroscopy*, Journal of Dairy Science 81 (1998) 2791-2797.

[9] Da-Wen Sun, *Infrared Spectroscopy for Food Quality Analysis and Control*, Elsevier, 2009.

[10] S. Herbert, N. Mouhous Riou, M.F. Devaux, A. Riaublanc, B. Bouchet, J.D. Gallant, E. Dufour, *Monitoring the identity and the structure of soft cheeses by fluorescence spectroscopy*, Le Lait 80 (2000) 621-634.

[11] R. Karoui, E. Dufour, R. Schoonheydt, J. De Baerdemaeker, *Characterisation of soft cheese by front face fluorescence spectroscopy coupled with chemometric tools: Effect of the manufacturing process and sampling zone*, Food Chemistry 100 (2007) 632-642.

[12] C.M. Andersen, M.B. Fröst, N. Viereck, *Spectroscopic characterization of low- and non-fat cream cheeses*, International Dairy Journal 20 (2010) 32-39.

[13] R. Karoui, E. Dufour, L. Pillonel, D. Picque, T. Cattenoz, J-O. Bosset, *Determining the geographic origin of Emmental cheeses produced during winter and summer using a technique based on the concatenation of MIR and fluorescence spectroscopic data*, European Food Research and Technology 219 (2004) 184-189.

[14] R. Karoui, J-O. Bosset, G. Mazerolles, A. Kulmyrzaev, E. Dufour, *Monitoring the geographic origin of both experimental French Jura hard cheeses and Swiss Gruyère and L'Etivaz PDO cheeses using mid-infrared and fluorescence spectroscopies: a preliminary investigation*, International Dairy Journal 15 (2005) 275-286.

[15] E. Dufour, *Recent advances in the analysis of dairy product quality using methods based on the interactions of light with matter*, International Journal of Dairy Technoly 64 (2011) 153-165.

[16] S. De Angelis Curtis, R. Curini, M. Delfini, E. Brosio, F. D'Ascenzo, B. Bocca, *Amino acid profile in the ripening of Grana Padano cheese: a NMR study*, Food Chemistry 71 (2000) 495-502.

[17] R. Karoui, J. De Baerdemaeker, *A review of the analytical methods coupled with chemometric tools for the determination of the quality and identity of dairy products*, Food Chemistry 102 (2007) 621-640.

[18] B.G. Osborne, in: Robert A. Meyers (Ed.), *Encyclopaedia of Analytical Chemistry: applications, theory and instrumentation*, John Wiley & Sons Ltd, Chichester, 2000.

[19] H. Martens, T. Naes, *Multivariate calibration*, John Wiley & Sons, New York, 1989.

[20] H. M. Heise, R. Winzen, in: H.W. Siesler, Y. Ozaki, S. Kawata, H.M. Heise (Eds.), *Near Infrared Spectroscopy: Principles, Instruments, Applications*, Wiley-VCH Verlag GmbH, Weinheim, 2002.

[21] ISO 21543:2006 (IDF 201:2006) – *Milk products – Guidelines for the application of near infrared spectrometry*, first ed., International Organization for Standardization, Geneva, 2006.

[22] Council Regulation (EC) No. 510/2006 of 20 March 2006, Brussels, on the protection of geographical indications and designations of origin for agricultural products and foodstuffs, Official Journal of the European Communities, L 93, 12-25.

[23] Commission Regulation (EC) No. 1107/96 of 12 June 1996 on the registration of geographical indications and designations of origin under the procedure laid down in

Article 17 of Council Regulation (EEC) No 2081/92, Official Journal of the European Communities, L 148-I, 1-15.

[24] Orden de 6 de mayo de 1993 por la que se aprueba el Reglamento de la Denominación de Origen "Queso Zamorano" y su Consejo Regulador, BOE No. 120 de 20 de mayo de 1993, 15311–15316.

[25] UNE–EN–ISO/IEC 17025:2005 – *General requirements for the competence of testing and calibration laboratories*, International Organization for Standardization, Geneva, 2005.

[26] ISO 1735:2004 (IDF 5:2004) – *Cheese and processed cheese products – Determination of fat content – Gravimetric method (Reference method)*, International Organization for Standardization, Geneva, 2004.

[27] ISO 5534:2004 (IDF 4:2004) – *Cheese and processed cheese products – Determination of the total solids content – Gravimetric method (Reference method)*, International Organization for Standardization, Geneva, 2004.

[28] ISO 8968–1:2001 (IDF 20–1:2001) – *Milk – Determination of nitrogen content – Part 1: Kjeldahl method*, International Organization for Standardization, Geneva, 2001.

[29] R.L. Wehling, M.M. Pierce, *Determination of Moisture in Cheddar Cheese by Near Infrared Reflectance Spectroscopy*, Journal of Association of Official Analytical Chemists 71 (1988) 571-574.

[30] A. Savitzky, M.J.E. Golay, *Smoothing and Differentiation of Data by Simplified Least Squares Procedures*, Analytical Chemistry 36 (1964) 1627-1639.

[31] R.J. Barnes, M.S. Dhanoa, S.J. Lister, *Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra*, Applied Spectroscopy 43 (1989) 772-777.

[32] M.C. Ortiz, L. Sarabia, A. Jurado-López, M.D. Luque de Castro, *Minimum value assured by a method to determine gold in alloys by using laser-induced breakdown spectroscopy and partial least-squares calibration model*, Analytica Chimica Acta 515 (2004) 151-157.

[33] M.C. Ortiz, L. Sarabia, R. García-Rey, M.D. Luque de Castro, *Sensitivity and specificity of PLS-class modelling for five sensory characteristics of dry-cured ham using visible and near infrared spectroscopy*, Analytica Chimica Acta 558 (2006) 125-131.

[34] R. Díez, M.C. Ortiz, L. Sarabia, I. Birlouez-Aragón, *Potential of front face fluorescence associated to PLS regression to predict nutritional parameters in heat treated infant formula models*, Analytica Chimica Acta 606 (2008) 151-158.

[35] P.J. Rousseeuw, A.M. Leroy, *Robust regression and outliers detection*, John Wiley and Sons, New Jersey, 2001.

[36] M.C. Ortiz, L. Sarabia, A. Herrero, *Robust regression techniques. A useful alternative for the detection of outlier data in chemical analysis*, Talanta 20 (2006) 499-512.

[37] M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, *Tutorial on evaluation of type I and type II errors in chemical analyses: From the analytical detection to authentication of products and process control*, Analytica Chimica Acta 674 (2010) 123-142.

[38] J. Incédy, T. Lengyed, A.M. Ure, A. Gelencsér, A. Hulanicki, International Union of Pure and Applied Chemistry (IUPAC), Compendium of Analytical Nomenclature, Blackwell, Oxford, 1998.

[39] ISO 11843-2 *Capability of Detection. Methodology in Linear Calibration Case*, International Organization for Standardization, Geneva, 2000.

[40] Commission Decision (EC) No 2002/657/EC of 12 August 2002, implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Official Journal of the European Communities, L 221, 8-36.

[41] M.C. Ortiz, L.A. Sarabia, A. Herrero, M.S. Sánchez, M.B. Sanz, M.E. Rueda, D. Giménez, M.E. Meléndez, *Capability of detection of an analytical method evaluating false positive and false negative (ISO 11843) with partial least squares*, Chemometrics and Intelligent Laboratory Systems 69 (2003) 21-33.

[42] M.C. Ortiz, L.A. Sarabia, I. García, D. Giménez, E. Meléndez, *Capability of detection and three-way data*, Analytica Chimica Acta 559 (2006) 124-136.

[43] R. Morales, L.A. Sarabia, M.C. Ortiz, *D-optimal designs and N-way techniques to determine sulfathiazole in milk by molecular fluorescence spectroscopy*, Analytica Chimica Acta 707 (2011) 38-46.

[44] B.M. Wise, N.B. Gallagher, R. Bro, J.M. Shaver, W. Windig, R.S. Koch, *PLS Toolbox 6.0.1.*, Eigenvector Research Inc., Manson, WA (2010).

[45] *STATGRAPHICS Centurion XVI, version 16.1.05,* StatPoint Technologies, Inc, Warrenton VA (2010).

# CHAPTER 3

## ENSURING FOOD SAFETY IN REGULATED PRODUCTS THROUGH DESIGN OF EXPERIMENT APPROACHES: EFFECT ANALYSIS AND ROBUSTNESS ASSESSMENT OF SEVERAL SAMPLE PRETREATMENT FACTORS ON THE DETERMINATION OF SEVEN SEDATIVES IN ANIMAL MUSCLE BY LIQUID CHROMATOGRAPHY-TANDEM MASS SPECTROMETRY

## 3.1. Introducción[1]

Este capítulo aborda el tema de la Seguridad Alimentaria desde una perspectiva analítica al afrontar dos de las etapas que forman parte del desarrollo y de la validación de un procedimiento analítico, en concreto, de determinación de residuos de sustancias farmacológicamente activas suministradas a animales destinados a la producción de alimentos: en primer lugar, se ocupa de la optimización de la etapa de pretratamiento muestral para, a continuación, evaluar su robustez y dar así evidencias de la calidad analítica del procedimiento propuesto. La metodología detallada en este capítulo ha sido diseñada y contrastada de acuerdo con la legislación europea vigente en el ámbito de las sustancias farmacológicamente activas (Directiva del Consejo 96/23/CE, Reglamento (UE) No 470/2009 y Reglamento (UE) No 37/2010), que estipula la necesidad de realizar una evaluación científica antes de que un producto alimenticio sea autorizado, estableciendo para ello límites máximos de residuo (MRL, *maximum residue limit*) para algunos de esos fármacos e incluso prohibiendo el uso de otros en caso necesario. Específicamente, el procedimiento analítico desarrollado y presentado a lo largo de este capítulo tiene como objetivo último la correcta determinación de 6 tranquilizantes (azaperol, azaperona, clorpromazina, haloperidol, propionilpromazina y xilazina) y un antiadrenérgico (carazolol) en músculo animal mediante LC-MS/MS, siguiendo durante dicho desarrollo los requerimientos formulados en la Decisión de la Comisión 2002/657/CE.

En general, la validación de un procedimiento analítico está fuertemente relacionada con su desarrollo. La mayor parte de las características de funcionamiento que han de comprobarse o verificarse durante su validación (selectividad, sensibilidad analítica, límite de decisión, límite de detección, veracidad, precisión y robustez, entre otras) son evaluadas normalmente a lo largo de su desarrollo, al menos, de manera aproximada. Así, antes de la validación formal de la versión final de un procedimiento, su Espacio de Diseño Analítico debería estar claramente definido, de forma que permita garantizar la calidad y su control: en este punto es donde las estrategias de Calidad por el Diseño (QbD, *Quality by Design*) basadas en el Diseño de Experimentos (DoE, *Design of Experiments*) y en el análisis multivariante juegan un rol decisivo.

Con estos conceptos en mente, el apartado 3.2 de este capítulo presenta un empleo novedoso de las funciones de deseabilidad en combinación con diseños D-óptimos en la optimización de análisis multirresiduo mediante técnicas cromatográficas. Esta estrategia

[1] La experimentación descrita en este capítulo fue realizada personalmente junto con L. Rubio en el *Laboratorio de Salud Pública de Burgos (Servicio Territorial de Sanidad y Bienestar Social de la Junta de Castilla y León, España)* en el marco de una estancia predoctoral. Las conclusiones extraídas de esta experimentación fueron incluidas en los informes necesarios para justificar la validación formal del procedimiento analítico resultante, que está siendo rutinariamente empleado en la actualidad en el laboratorio anteriormente mencionado.

permite que la respuesta experimental bajo investigación refleje diversos aspectos del problema analítico en cuestión. A este respecto y como parte del desarrollo de un procedimiento analítico para la determinación de los sedantes mencionados anteriormente, se ejecutó un diseño D-óptimo *ad hoc* destinado a evaluar el efecto de cuatro factores relacionados con la etapa de pretratamiento de la muestra. Los resultados obtenidos fueron interpretados a partir de tres variables respuesta distintas: *i)* área absoluta de pico de cada compuesto (ya sea analito o estándar interno), *ii)* área relativa de pico de cada analito e *iii)* función de deseabilidad conjunta definida en base a criterios de aceptación de la relación entre los valores de las áreas de pico de cada analito y de su estándar interno. Como se verá más adelante, llevar a cabo esta triple interpretación de los resultados del diseño D-óptimo ejecutado permitirá poner de manifiesto que, dependiendo de la respuesta experimental escogida, se puede llegar a conclusiones muy diferentes en lo que al nivel óptimo de un factor se refiere, lo cual es un aspecto que ha de tenerse en mente de cara a establecer correctamente cualquier Espacio de Diseño Analítico, de acuerdo con lo recogido en el apartado 1.1 del capítulo 1 de esta tesis.

Una vez optimizado el procedimiento analítico, la metodología del diseño de experimentos se empleó nuevamente en el estudio de la robustez de la etapa de pretratamiento de la muestra, como se describe en el apartado 3.3 de este capítulo. De acuerdo con la legislación europea, se ejecutó un diseño factorial fraccional de Youden para evaluar la influencia de siete variables relacionadas con aquella etapa sobre la concentración finalmente cuantificada de los siete fármacos veterinarios. Para la interpretación de los resultados obtenidos a partir del plan experimental resultante y evaluar en consecuencia la significación de esos siete factores, se utilizaron varias estrategias estadísticas: *i)* test de hipótesis paramétrico usando una varianza externa previamente estimada, *ii)* método de Lenth e *iii)* análisis bayesiano. Las estrategias de Lenth y de Bayes permiten determinar el efecto de una variable sobre la respuesta en ausencia de una estimación de la variabilidad experimental: esta característica es de particular importancia cuando realizar réplicas de uno o más puntos de un diseño de experimentos no es posible por coste económico o de tiempo, teniendo que hacer frente a esquemas saturados. Independientemente de la aproximación utilizada para la evaluación de la influencia de los siete factores sobre la determinación de cada analito, se comprobó que la conclusión finalmente alcanzada era la misma, en concreto, que el proceso de pretratamiento muestral bajo investigación es robusto frente a los pequeños cambios considerados.

## 3.2. Desirability functions as response in a D-optimal design for evaluating the extraction and purification steps of six tranquillizers and an anti-adrenergic by liquid chromatography-tandem mass spectrometry[2]

### 3.2.1. Summary

Internal standards can be added at different stages of an analytical procedure. When they are added at the beginning of a multiresidue method and their behaviour is not exactly the same as that of the analytes, the intended correction for small variations within the analytical process could not be achieved. Because of this, in the present work, the use of D-optimal designs together with desirability functions is proposed to state the experimental response under study. The overall desirability function used relates two analytical criteria: to assess a similar chemical behaviour of each analyte in relation to its internal standard and to avoid a significant reduction of the absolute peak area of the internal standards. This strategy has been applied to the analysis of the effect of four factors related to the extraction and purification steps of six tranquillizers and a *β*-blocker from pig muscle analysed by liquid chromatography-tandem mass spectrometry (LC-MS/MS). The effect of those factors has been evaluated by means of an *ad hoc* D-optimal design consisting of only 11 experiments. The resulting levels of the four factors that enable to achieve the greatest overall desirability have also been compared to those obtained when either the standardized or absolute peak area has been considered as response. Differences in both the significant factors and their optimum levels have been observed. It is noticeable that the experimental effort necessary to study the effect of the factors has been reduced by more than 50% thanks to the D-optimal design.

### 3.2.2. Introduction

Tranquillizers and sedatives have been frequently used in animal production since 1970 [1,2] because some farm animals, especially pigs, are not able to adapt to stress situations and die prematurely during their transport. These drugs are usually injected a few hours before slaughtering, so residues might be present in foodstuffs of animal origin [2,3] and might constitute a hazard to human health. For this reason, specific rules on this matter have been laid down in European legislation, such as Commission Regulation (EU) No 37/2010 [4], where maximum residue limits (MRLs) for some pharmacologically active substances are established depending on the animal species and on the target tissue.

[2] This section is published as *L. Rubio, M.L. Oca et al. / J. Chemometrics 2016; 30: 58–69* (4 citations up to August 30, 2021 according to *Scopus*).

The substances considered in this work were five tranquillizers (azaperone (AZA), propionylpromazine (PROP), chlorpromazine (CHLOR), haloperidol (HAL) and xylazine (XYL)), one of the metabolites of azaperone (azaperol (AZOL), which is synthesized from azaperone by reduction) and a blocker agent of the *β*-adrenergic receptor (carazolol (CAR), used to control tachycardia). All their chemical structures together with their CAS Registry numbers and molecular weights appear in Figure 3.2.1. The MRL in porcine muscle tissue for azaperone, as the sum of azaperone and azaperol, is 100 μg $kg^{-1}$, while that for carazolol is 5 μg $kg^{-1}$. These MRLs have been set following the procedures foreseen in Regulation (EC) No 470/2009 [5].

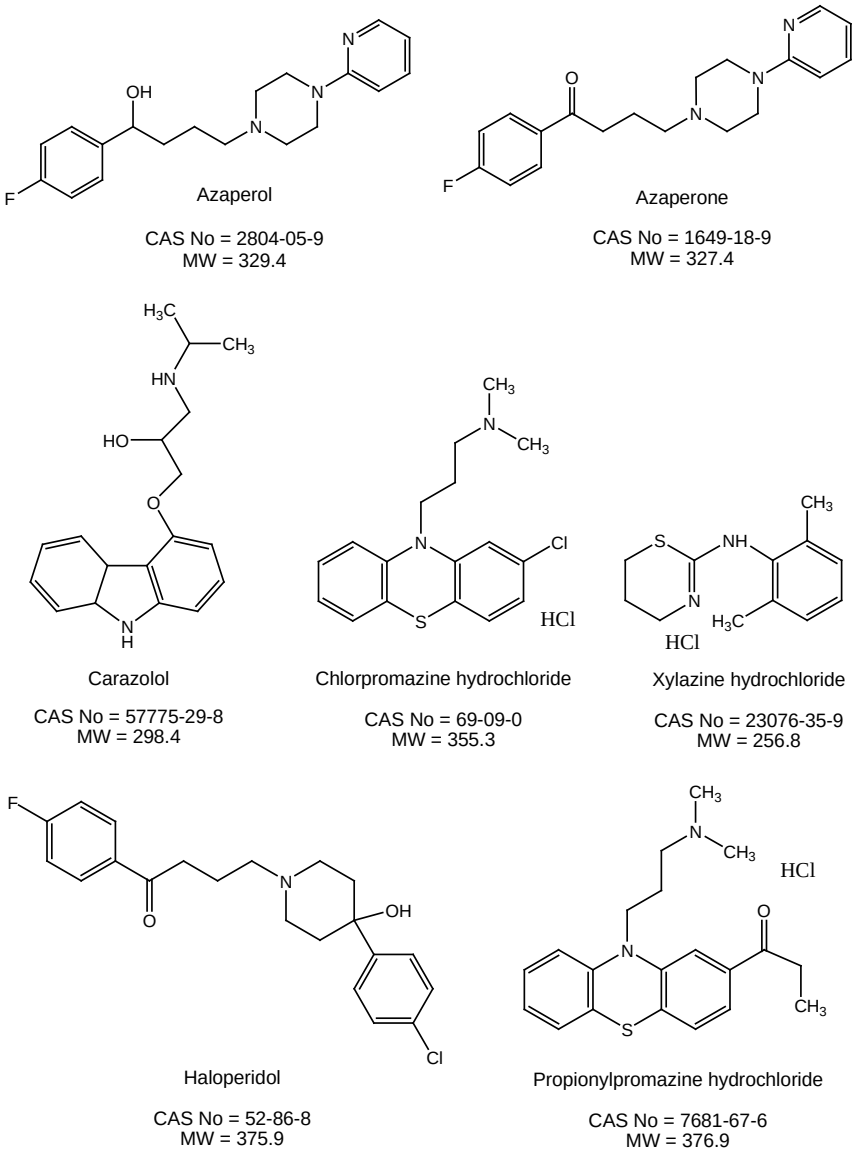


**Figure 3.2.1.** Chemical structure, CAS Registry Number and molecular weight (MW, in g/mol) of the beta-blocker carazolol and the six tranquillizers (three of them in their hydrochloride forms) analysed. Concentrations of these analytes are referred to their free forms in this work.

A great number of techniques are available for the particular determination of residues of tranquillizers and carazolol in animal tissues: radioimmunoassay [6], radioreceptor assay [7], enzyme-linked immunosorbent assay (ELISA) [8], thin-layer chromatography (TLC) [9,10], high-performance liquid chromatography (HPLC) with ultraviolet detection [10,11,12], with fluorescence detection [11,13,14] or with electrochemical detection [15].

However, according to the rules provided by Commission Decision 2002/657/EC [16], confirmatory methods for organic residues or contaminants shall provide information on the chemical structure of the analyte. So the use of mass spectrometric detection is compulsory in order to ensure the reliability of the results achieved. In fact, several works have used gas chromatography with mass spectrometric detection (GC/MS) [17,18,19] and liquid chromatography with mass spectrometric detection (LC-MS) or with tandem mass spectrometric detection (LC-MS/MS) [1,12,18,20,21,22] for the analysis of tranquillizers.

For the confirmation of banned substances (listed in Group A of Annex I of Directive 96/23/EC [23]), such as chlorpromazine, a minimum of 4 identification points shall be required, whereas for the confirmation of substances with a MRL (listed in Group B of Annex I of Directive 96/23/EC), a minimum of 3 identification points is needed [16]. All the drugs analysed in this work, except chlorpromazine, are classified in Group B2d of Directive 96/23/EC. In LC-MS/MS, the precursor ion and each transition product mean, respectively, 1 and 1.5 identification points, provided that the relative intensity of each detected ion fulfils the maximum permitted tolerances also established in [16].

In addition, Commission Decision 2002/657/EC sets another requirement for the unequivocal identification of the analytes: "The ratio of the chromatographic retention time of the analyte to that of the internal standard, i.e. the relative retention time of the analyte, shall correspond to that of the calibration solution at a tolerance of ± 0.5% for GC and ± 2.5% for LC". So, internal standards (ISs) are needed in LC-MS/MS to fulfil these confirmatory criteria.

The election of the stage of the analytical procedure in which the ISs are introduced is important. An IS can be added to the sample just before injection to correct for instrumental fluctuations [24]. However, sometimes it is necessary to add the IS to the sample at the beginning of the procedure prior to any chemical changes and as early as possible, acting as a surrogate standard. In this way, the IS goes through the extraction (and/or purification steps) and analysis together with the analyte of interest. Therefore, surrogate standards are used to correct for random and systematic errors of the entire analysis (including sample preparation, chromatography and detection). Any variations that occur to the analyte in all these processes are expected to be compensated by the IS [25,26], so, for ease of control, multiple ISs are added at various points in the analytical

process to account for all possible sources of variability in some cases. However, in a multiresidue method, where several analytes are to be standardized with the same IS, the dissimilarity between analytes and their ISs can result in quite different behaviours and unbalanced losses may occur.

A review of the responses considered in experimental designs in chromatography was conducted. "*Chromatography*", "*experimental design*" and "*area*" were the keywords entered in the search fields at ScienceDirect database. It could be concluded that the relative peak area was considered in some cases [27,28]; in others, the selected response variable was the absolute area [29,30,31,32,33,34,35], the mean of the absolute areas of several peaks [36] or their sum [34,37].

If the purpose of the analysis is to identify the analyte by its own mass spectrum, the absolute peak area could be thought as the best response at the experimental design because a peak just as large as possible would be pursued. But quantification in chromatography is generally performed by the prior standardization of peak areas. In fact, when the aim of the analytical method was to quantify, the standardized peak area was used in some of those works [27,29,31,32,36,37], despite that variable had not been chosen as response in the experimental design. If an optimum quantification is wanted, the most logical choice will be to use the standardized peak area as response. However, if the IS has been added at the beginning of the analytical procedure and the pretreatment steps affect the analyte and the IS in a different way, the $A_{analyte}/A_{IS}$ ratio becomes the quotient between two conflicting variables ($A_{analyte}$ and $A_{IS}$) that have changed at quite different extents. In this case, considering that ratio without bearing in mind the true analytical situation underneath would lead to erroneous decisions on which factors have effect on the response and which the best analytical conditions are. In this sense, the desirability function by Derringer [38] is a tool flexible enough to join the most relevant aspects of this problem together yielding an only response variable. Desirability functions are usually used in optimizations based on continuous factors (response surface methodology) [39]. However, their application on discrete factors is not usual.

Another well-known fact is that there are, in general, many factors involved in an analytical procedure. D-optimal designs [40] are quite useful for estimating their effects and interactions with enough precision although the size of the experimental design has been reduced to feasible levels.

In this work, the use of D-optimal designs together with desirability functions is proposed to state the experimental response under study for the analysis of the effect of the extraction and purification steps of six tranquillizers and a $\beta$-blocker from pig muscle analysed by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS)

where the ISs have been added at the beginning of the procedure. Some of the selected factors in this analysis are continuous but they are considered as discrete in the experimental design. The results are then compared to those obtained when either standardized or absolute peak areas are considered as response.

### 3.2.3. Methodology

#### *3.2.3.1. D-optimal experimental design*

In the experimental design methodology, three items must be considered before creating the experimentation plan: the experimental domain under study, the factors to be investigated and the response or responses to be analysed and optimized. Once all this has been set, the type of experimental design that will be carried out must be selected. At this point, it is important to notice that D-optimal designs are useful to adapt the experimental plan to the problem studied, reducing the number of observations substantially without losing efficiency and precision in the results. In the following, the fundamental concepts of this approach are described. A complete description can be found elsewhere in [40].

An experimental domain consisting of all the factors under study and their respective levels is defined. Then a linear model is supposed for the relation between the experimental response **y** and the values of the factors

$$\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{3.2.1}$$

where $\mathbf{X} = (x_{ij})$ is the model matrix with dimensions $n \times p$ ($n$, number of experiments; $p$, number of coefficients of the model; they both take the experimental design and the chosen model into account); **y** is the vector of the experimental responses; $\boldsymbol{\beta}$ is the vector of the coefficients and $\boldsymbol{\varepsilon}$ is the vector of the experimental errors.

The joint confidence region for the estimated coefficients, **b**, is a hyperellipsoid. It is said that **X** is D-optimal when the volume of the hyperellipsoid is minimum. A matrix model **X** is said to be G-optimal when the maximum value of the response variance is minimum.

After having defined the experimental domain, $N_C$ candidate observations are fixed in it, so that there is sufficient information to estimate the p coefficients of the proposed model ($N_C$ = 24 and $p$ = 7 in our case, see Section 3.2.5.1 below). Then, for each number of experiments, $N$ ($p \leq N \leq N_C$), a matrix $\mathbf{X}_N$, which is D-optimal, is determined. That is, from the $N_C$ candidates, the $N$ experiments that provide the most precise joint estimation of all the coefficients of the model are extracted. This D-optimal solution could not be unique. Next, the model matrix **X** is selected from among all the $\mathbf{X}_N$ matrices available in

order to achieve an acceptable value of the G-criterion from the lowest possible number of experiments.

#### *3.2.3.2. Desirability function*

The desirability function of Derringer and Suich [38] allows to find the factor levels to reach the best possible value for all the evaluated responses in order to fulfil some prior specifications. This is achieved by converting the multiple responses into a single one followed by its optimization.

In a first step, each response $i$ is transformed over the experimental domain into an individual desirability function, $d_i$, which ranges between $d_i$ = 0 % (undesirable response) to $d_i$ = 100% (optimal response).

In a second step, the overall desirability function, $D$, is calculated as the weighted geometric mean of the $s$ individual desirability functions:

$$D = \sqrt[w]{d_1^{w_1} \times d_2^{w_2} \times ... \times d_s^{w_s}} \quad , \quad w_i \geq 0, \;\; i = 1,...,s \;\text{ and }\; w = w_1 + w_2 + ... + w_s \qquad (3.2.2)$$

The analyst introduces the relative importance given to each response through the weightings, $w_i$

It should be kept in mind that the goal for an optimization procedure is to find a good set of conditions that will meet all the goals, but not to get a $D$ value equal to 100%.

### 3.2.4. Experimental

#### *3.2.4.1. Reagents and standard solutions*

HPLC-grade methanol, analytical-grade ethanol, HPLC-grade acetonitrile, formic acid (98-100% purity), 32% ammonia solution, analytical-grade potassium dihydrogen phosphate and acetic acid were all purchased from Merck (Darmstadt, Germany). Analytical-grade orthophosphoric acid and potassium hydroxide (both 85% purity) were supplied by Panreac (Barcelona, Spain).

Azaperol (97% minimum purity) was supplied by Dr. Ehrenstorfer GmbH (Augsburg, Germany). Xylazine hydrochloride, haloperidol, chlorpromazine hydrochloride, propionylpromazine hydrochloride, azaperone and azaperone-$d_4$ (used as the IS for both azaperone and azaperol) were obtained from Sigma-Aldrich (Steinheim, Germany), all of them with a 98% minimum purity. Chlorpromazine-$d_3$ (98% minimum purity, used as the IS for both chlorpromazine and propionylpromazine), and haloperidol-$d_4$ (98% minimum

purity, used as the IS for haloperidol, carazolol and xylazine), were purchased from LGC Standards (Teddington, UK). Carazolol (Suacron®, 98% minimum purity) was a generous gift from Divasa-Farmavic, S.A. (Barcelona, Spain).

Stock solutions of azaperone, azaperol, carazolol, propionylpromazine, xylazine, chlorpromazine and haloperidol at 0.1 g $L^{-1}$ and individual stock solutions of azaperone-$d_4$ (100 mg $L^{-1}$), chlorpromazine-$d_3$ (1 mg $L^{-1}$) and haloperidol-$d_4$ (1 mg $L^{-1}$) were prepared in ethanol. All these standard solutions were stored at 4 ºC in amber bottles for a maximum of 1 year. Concentrations of these analytes must be referred to their free forms, not to the hydrochloride ones.

Two working solutions were prepared in ethanol: one containing the seven non-deuterated drugs at the following concentrations: 2500 μg $L^{-1}$ for azaperone, 2000 μg $L^{-1}$ for azaperol, 125 μg $L^{-1}$ for carazolol, 100 μg $L^{-1}$ for chlorpromazine, 50 μg $L^{-1}$ for haloperidol, 150 μg $L^{-1}$ for propionylpromazine and 100 μg $L^{-1}$ for xylazine; and the other with the three ISs at concentrations of 200 μg $L^{-1}$ for chlorpromazine-$d_3$, 40 μg $L^{-1}$ for haloperidol-$d_4$ and 2500 μg $L^{-1}$ for azaperone-$d_4$. Both solutions were stable for 2 months and stored at 4 ºC in amber bottles.

Two solutions of phosphate buffer (0.1 M) with a pH value of 6 and 7.4, respectively, were also prepared.

#### *3.2.4.2. Instrumental*

Millex-AP prefilters were obtained from Millipore Corporation (Bedford MA, USA). A Jouan C3i centrifuge (Thermo Electron Corporation, San Jose, CA, USA) was used. The vacuum manifold used for the solid-phase extraction (SPE) step was purchased from Waters Corporation (Milford, MA, USA), and SPE cartridges were 300 mg/6 mL Varian Bond Elut LRC Certify® (Agilent Technologies, Waldbronn, Germany).

#### *3.2.4.3. Sample preparation and purification procedure*

In each experiment a 5-g aliquot of homogenised pig muscle was weighed in a 50 mL polypropylene tube; then, every sample was fortified with the three ISs by adding a 50 μL portion of the working internal standard solution. 100 μL of the working standard solution of tranquillizers and carazolol was added to all tubes, except for the blank sample, which was prepared in parallel with the performance of the D-optimal design. Next, 10 mL of 0.1 M phosphate buffer solution was pipetted to every sample and the mixture was stirred in a vortex mixer for 30 s. The pH value of the buffer solution (7.4 or 6) was that shown in Table 3.2.1 for the corresponding experiment. After 10 min in an ultrasonic bath, 3 mL of

1% orthophosphoric acid was added, and all tubes were vortex mixed again for 30 s and then centrifuged for 10 min at 6800 rpm at room temperature. Before the SPE step, the supernatant of every homogenate was clarified using a Millex-AP prefilter coupled to a 20 mL syringe. Each filtrate was collected in a 15 mL amber tube.

The purification phase by SPE was performed at the corresponding level for the flow rate (10 or 2 mL/min) following the design in Table 3.2.1. Firstly, SPE cartridges were placed on a multistation vacuum manifold, conditioned with 3 mL of methanol and then equilibrated with 3 mL of Milli-Q water. After every sample filtrate had been loaded, cartridges were rinsed with 3 mL of 1 M formic acid and next with 3 mL of a washing solution (methanol/1 M formic acid (55/45 v/v) or methanol/1 M acetic acid (55/45 v/v), see Table 3.2.1). Whenever the washing mixture containing acetic acid was used, the SPE cartridge was first rinsed with 1 M acetic acid instead of with 1 M formic acid. All SPE cartridges were dried by passing an air stream under vacuum for 10 min. Elution of the analytes was performed with a volume of the eluting solution (ammonium hydroxide/ethanol 2/98 v/v) equal to the value shown for each experiment in Table 3.2.1 (3, 4 or 5 mL). Each eluate was collected in a 15-mL amber tube, evaporated to dryness under nitrogen at 38 ºC, reconstituted with 250 µL of 30% acetonitrile and placed in an ultrasonic bath for 5 min. Every final extract was clarified through a 0.22-µm nylon filter coupled to a 1-mL syringe, and the filtrate was transferred to an insert contained into a vial for LC-MS/MS analysis.

The preparation of the blank sample followed the same procedure detailed above, except that the pH value of the phosphate buffer was 6, the rinsing solution in SPE was methanol/1 M formic acid (55/45 v/v) and 4 mL of the elution mixture was required in the SPE final step.

**Table 3.2.1.** Experimental matrix and values of the absolute peak area for each of the seven analytes and the three internal standards. Factor 1: SPE flow rate (10 or 2 mL/min, levels A or B, respectively). Factor 2: pH of the phosphate buffer (7.4 or 6, level A or B, respectively). Factor 3: SPE washing solution (Methanol/Formic acid, level A, or Methanol/Acetic acid, level B). Factor 4: Volume of eluting solution (3, 4 or 5 mL, levels A, B or C, respectively).

| Exp. | Level | | | | CAR | XYL | AZOL | AZA | HAL | PROP | CHLOR | AZA-$d_4$ | HAL-$d_4$ | CHLOR-$d_3$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Factor 1 | Factor 2 | Factor 3 | Factor 4 | | | | | | | | | | |
| 1 | B | A | A | A | 14541 | 47444 | 1059606 | 1344758 | 6678 | 4701 | 783 | 408423 | 1812 | 854 |
| 2 | A | B | A | A | 19679 | 48091 | 1318343 | 1673400 | 10118 | 7930 | 1786 | 523213 | 2790 | 1932 |
| 3 | A | A | B | A | 10892 | 48928 | 941931 | 1031595 | 4840 | 0 | 0 | 324853 | 1519 | 0 |
| 4 | B | B | B | A | 22680 | 47127 | 1125323 | 1220836 | 7750 | 3483 | 593 | 367845 | 2009 | 563 |
| 5 | A | A | A | B | 16918 | 55058 | 1236715 | 1462197 | 7727 | 5206 | 903 | 448741 | 2033 | 894 |
| 6 | B | A | A | B | 18508 | 49621 | 1151126 | 1396737 | 8566 | 5340 | 1029 | 425087 | 2524 | 940 |
| 7 | A | B | B | B | 23379 | 55711 | 1233259 | 1456196 | 8537 | 5980 | 1148 | 453934 | 2274 | 1081 |
| 8 | B | B | B | B | 23737 | 52246 | 1172705 | 1290610 | 6454 | 3244 | 567 | 395821 | 1701 | 528 |
| 9 | A | B | A | C | 25153 | 51786 | 1493910 | 1812107 | 13127 | 9627 | 2014 | 556404 | 3582 | 2025 |
| 10 | B | B | A | C | 24438 | 45228 | 1429542 | 1700984 | 12018 | 8935 | 1952 | 507898 | 3310 | 1998 |
| 11 | B | A | B | C | 16820 | 50553 | 1019783 | 1178097 | 5189 | 3379 | 587 | 367405 | 1505 | 550 |

*3.2.4.4. Liquid chromatography coupled to tandem mass spectrometry analysis*

Chromatographic separation was carried out on an Agilent 1200RR series liquid chromatograph (Agilent Technologies, Waldbronn, Germany) equipped with a binary pump, a degasser, an autosampler and a column heater, where a Gemini NX C18 column (50 mm × 2 mm I.D.; particle size, 3 μm; Phenomenex, Macclesfield, UK) was placed at 30 ºC. Two mobile phases were used in a gradient program, where the mobile phase A consisted of acetonitrile containing 0.5% ammonium hydroxide and mobile phase B was water containing 0.5% ammonium hydroxide. The initial composition was 30% A and 70% B, which was increased linearly to 85% A in the first 7 min, held at these conditions for 2 min and returned to the initial composition from 9 to 11 min. The column was then equilibrated in these conditions for 4 min before the next injection (the total run time was 15 min). The flow rate was set at 0.3 mL $min^{-1}$ and the injection volume was 10 μL.

Tandem mass spectrometry was performed on an Agilent 6410B triple quadrupole mass analyzer with a multimode ionization source which was programmed to operate in mixed conditions of atmospheric pressure chemical ionization (APCI) and positive ion electrospray ($ESI^+$). The source temperature was held at 250 ºC and the capillary voltage was set at 4300 V. Nitrogen was used as nebuliser gas and collision gas. Analytes were detected using the multiple reaction monitoring (MRM) mode with a dwell time of 70 ms per ion. Both Q1 and Q3 quadrupoles were maintained at unit resolution. For the confirmation of the presence of each drug, its molecular ion was selected as the precursor ion and two transitions were monitored, as it can be seen in Table 3.2.2. Quantification was carried out with the most intense transition. Regarding the three ISs, only one transition from its corresponding molecular ion was considered in each case.

**Table 3.2.2.** Transitions monitored for the confirmation of the seven analytes and the three internal standards.

| Compound | Most intense transition | Less intense transition |
|---|---|---|
| Xylazine | 221.2 > 90.0 | 221.2 > 164.1 |
| Azaperol | 330.3 > 121.1 | 330.3 > 149.2 |
| Carazolol | 299.2 > 116.1 | 299.2 > 222.1 |
| Azaperone | 328.1 > 165.0 | 328.1 > 121.1 |
| Propionylpromazine | 341.2 > 86.1 | 341.2 > 58.1 |
| Chlorpromazine | 319.1 > 86.2 | 319.1 > 58.1 |
| Haloperidol | 376.2 > 165.1 | 376.2 > 123.2 |
| Azaperone-$d_4$ | 332.0 > 169.0 | - |
| Chlorpromazine-$d_3$ | 322.2 > 61.0 | - |
| Haloperidol-$d_4$ | 380.0 > 169.0 | - |

## 3.2.5. Results and discussion

### *3.2.5.1. Selection of factors, experimental domain and response variables for the analysis of the effect of the sample preparation step by means of a D-optimal design*

The four factors are related to the extraction and the purification steps of the analytes from the samples:

- The SPE flow rate (Factor 1) affects the capability of the analytes to get retained in the cartridge. If the flow rate is too great, the analytes may not be efficiently retained. The levels considered were A (10 mL/min) and B (2 mL/min).
- Tranquillizers are drugs with basic properties; that is why basic extracting solvents are expected to provide better recoveries. In this work, Factor 2 (pH of the extracting solvent) was selected at two levels: 7.4 (slightly basic, level A) and 6.0 (slightly acid, level B).
- The washing step of the SPE (Factor 3) is also critical in method development. Washing mixtures are selected in such a way that matrix interferents are eluted from the SPE cartridge while analytes remain retained in it. Factor 3 was studied at two levels: washing solution containing formic acid (level A) and acetic acid (level B), both at the same concentration.
- The final step of the SPE procedure is the elution of the analytes from the cartridge. The volume of the elution solvent (Factor 4) is also an important variable. It is advisable to elute all the analytes in a volume as low as possible. However, if the elution volume is too small, analytes may remain retained in the cartridge, whereas if it is too large, more interferents could be eluted at this step. Besides, in this latter case, the time of the evaporation step of the eluted extracts will increase, and as a consequence, analytes could be thermally degraded. That is the reason why Factor 4 was studied at three levels in case its effect on the response was not linear: 3 mL (level A), 4 mL (level B) and 5 mL (level C).
- The possible interaction between Factors 1 and 2 was also taken into account.

The full factorial design necessary to handle three factors at two levels and another one at three levels would have $2^3 \times 3^1 = 24$ experiments. The mathematical model that relates the levels of the four factors and the interaction to the response variable considered is expressed in Eq. (3.2.3).

$$y = \beta_0 + \beta_{1B\backslash A} X_{1B\backslash A} + \beta_{2B\backslash A} X_{2B\backslash A} + \beta_{3B\backslash A} X_{3B\backslash A} + \beta_{4C\backslash A} X_{4C\backslash A} + \beta_{4C\backslash B} X_{4C\backslash B} + \beta_{1B\backslash A2B\backslash A} X_{1B\backslash A} X_{2B\backslash A} + \varepsilon \quad (3.2.3)$$

The binary variables $X_{i\,M\backslash N}$, $i$ = 1,…, 4; $M$ = B, C (reference level) and $N$ = A, B, have the value of 1 when the $i$-th factor is at level $N$ and of 0 in the opposite case. For example, when Factor 1 is at level A, $\beta_{1B\backslash A}$ is added but not when it is at level B, so the coefficient $\beta_{1B\backslash A}$ is the variation of the response when Factor 1 changes from level B to level A. In the particular case of the 3-level factor (Factor 4), $\beta_{4C\backslash A}$ means the variation of the response when that factor moves from level C to level A, and $\beta_{4C\backslash B}$ that when it does to level B. Figure 3.2.2 describes graphically the meaning of these coefficients.

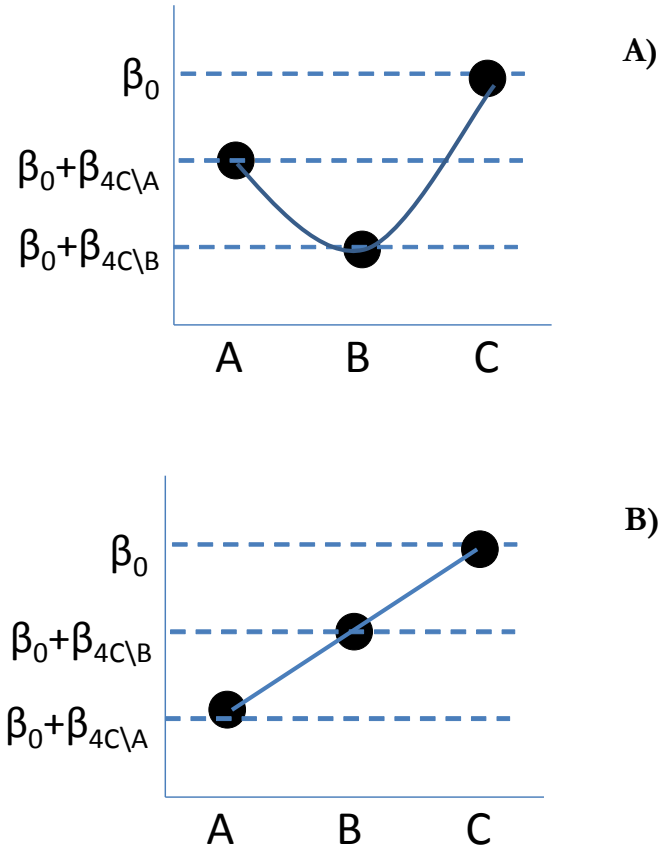


**Figure 3.2.2.** Schematic representation of the model of Eq. (3.2.3). When Factor 4 is at level C, the variables $X_{4C\backslash A}$ and $X_{4C\backslash B}$ have the value of 0. Then, the response $y$ has the value $\beta_0$ (supposing that the other factors are at the reference level). However, if Factor 4 is at level A, $\beta_{4C\backslash A}$ is added because the corresponding variables $X_{4C\backslash A}$ and $X_{4C\backslash B}$ are equal to 1 and 0, respectively. Similarly, if that factor is at level B, $\beta_{4C\backslash B}$ is then added as $X_{4C\backslash A}$ and $X_{4C\backslash B}$ have the value of 0 and 1, respectively. **A)** Non-linear effect and **B)** Linear effect of the levels of Factor 4.

24 is a number of experiments too high for the whole analysis to be completed within one day. The use of a D-optimal design will enable to reduce the experimental effort to that strictly necessary to estimate all the effects and interactions of interest with enough

precision. The model in Eq. (3.2.3) has got 7 coefficients, so at least 7 out from the 24 experiments in the experimental domain of the full factorial design would be required to estimate them.

Considering the full factorial design, $N_c$ = 24, the value for the variance inflation factor (VIF) for both $\beta_{4C\backslash A}$ and $\beta_{4\,C\backslash B}$ was 1.33, whereas it was 1 for the rest of the coefficients. In addition, the maximum value for the variance function (*d(u)*) was 0.29, so that achieving good precision in the estimation of both the coefficients and the response was feasible.

After the building of the D-optimal experimental matrices containing between 7 and 24 observations [41], an 11-experiment D-optimal design was elected because its maximum VIF was equal to 1.26. The values of the variance function *d(u)* for this design varied between 0.51 and 0.72, which were acceptable because they were all less than 1. Reducing the number of experiments from 24 to 11 meant an important saving in time and costs. The 11 experiments that made up the D-optimal design are collected in coded variables in Table 3.2.1 (columns 2 – 5).

Starting from experimental data, the election of the response variable considered for the study must reflect the aim of the experimentation faithfully. This election will determine the interpretation of the results. This fact will be shown by the analysis of the same experimental data using three different response functions:

- In Case 1 (see Section 3.2.5.1.1), the most relevant aspects of the problem under study will be considered together: to assess a similar chemical behaviour of each analyte in relation to its IS and to obtain the maximum peak area of the internal standards (keeping them above a level); so a global desirability function will be defined.
- In Case 2 (see Section 3.2.5.1.2), the aim is to get the experimental conditions of the four factors so that the absolute peak area of each analyte ($A_{analyte}$) will be maximized. As the aim is to develop a multiresidue method, if different experimental conditions are obtained for the maximization of the peak area of every analyte, either those for chlorpromazine (banned compound) or those saving costs or time will be chosen.
- In Case 3 (see Section 3.2.5.1.3), the peak area of each analyte standardized by the peak area of the corresponding internal standard (standardized peak area) will be studied.

These three studies have been individually performed for all seven analytes.

3.2.5.1.1. Case 1

In the analysis of the effect of the extraction and purification steps, if the analytes have different chemical behaviours, the internal standards ought to be chosen so that they will behave in a similar way to the analytes of interest. The ideal situation would be to have a deuterated IS for every analyte. But, in practice, it is not possible to achieve that in a multiresidue method, where several analytes are to be standardized with the same IS, in general, because of their different analytical sensibility, the commercial availability or their expensive prices. The key problem is to evaluate the behaviour of the IS, that is, to study when the standardization is adequate. This will occur when the IS behaves in the same way as the deuterated derivative of the analyte. In other words, when the previous phase of sample preparation affects both the analyte and the IS in a similar way. Thus the working hypothesis is that the previous pretreatment procedure affects both the analyte and its IS equally, so the $A_{analyte}/A_{IS}$ ratio should be constant throughout the sample preparation stage. In this ideal situation, the experimental conditions that would obtain the best response would be the same when both the absolute and the standardized peak area were considered.

In this work, the three ISs used were added at the beginning of the analytical procedure, so they underwent all the analysis steps. In addition, these ISs were deuterated derivatives of three of the analytes of interest, but they were also used to correct the variations of the rest. As the structures of the remaining analytes differed from those of the ISs, a dissimilar behaviour and subsequent unbalanced losses were likely to appear in this case. Therefore, the effect of the change in the experimental factors will be mainly noticed on those analytes that have not been corrected by its own deuterated derivative. If the change in the level of a factor leads to a major variation in the $A_{analyte}/A_{IS}$ ratio, that level must be chosen so as to minimize this change. In order to set the variability margin acceptable to conclude that a factor was significant, the upper and lower action limits for the $A_{analyte}/A_{IS}$ ratio were considered, which are part of the laboratory standard operating procedure (Public Health Laboratory) and that are collected in Table 3.2.3. These were the values of the standardized area of every analyte obtained from a set of standards containing 100 µL of the working standard solution of tranquillizers and carazolol.

**Table 3.2.3.** Limits for the values of standardized peak area when 100 µL of the working standard solution are considered.

| | **CAR** | **XYL** | **AZOL** | **AZA** | **HAL** | **PROP** | **CHLOR** |
|---|---|---|---|---|---|---|---|
| Lower | 2.30 | 2.70 | 1.30 | 2.20 | 3.06 | 2.58 | 0.82 |
| Upper | 10.88 | 14.94 | 3.54 | 3.24 | 5.20 | 11.02 | 1.52 |

If one of the values of the $A_{analyte}/A_{IS}$ ratio lay out of the corresponding limits, the procedure to which that sample had been subjected should be considered to be inappropriate, since the chemical behaviour of the analyte and that of the IS would be different from each other. Therefore, the object was not to maximize either the absolute or the standardized peak areas, but to keep the $A_{analyte}/A_{IS}$ ratios within the respective margins in Table 3.2.3, so a desirability function with the shape shown in Figure 3.2.3.a was defined.

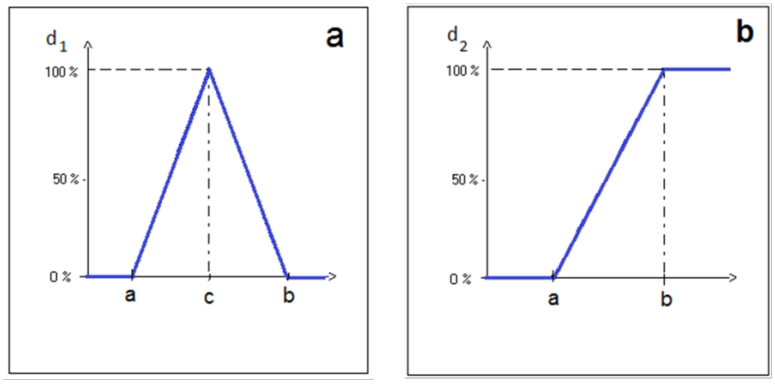


**Figure 3.2.3.** **a.** Shape of the individual desirability function $d_1$. **a** and **b** symbolize, respectively, the lower and upper limits of the validity interval for the standardized peak area of each analyte (see Table 3.2.3); **c** is the mean value of these two limits.

**b.** Shape of the individual desirability function $d_2$, where **a** and **b** represent, respectively, 50% and 80% of the maximum absolute peak area obtained for each internal standard after the performance of the experimental design.

For every analyte, the bilateral function $d_1$ was equal to 0% (undesirable response) for those values of standardized peak area either less or higher than the lower or upper action limit established, respectively. $d_1$ had a value of 100% (optimum response) when the mean value of the corresponding action limits was achieved. A linear function between the desirability values of 0% and 100% was built from the data included in Table 3.2.4. All the values of $d_1$ determined in this way are gathered in Table 3.2.5.

In addition, it must be borne in mind that the level of one or more factors could lead to such a relevant decrease in the IS peak area that this value turned out to be inadequate for the correction of the analytes. The way of defining this matter was by means of a unilateral desirability function $d_2$ (whose shape is depicted in Figure 3.2.3.b) that equalled: 0% when the absolute peak area of the IS was less than 50% of the greatest signal registered for this IS according to the experimental plan; 100% when the absolute peak area of the IS was higher than 80% of that greatest signal; and linear between 0% and 100% when the absolute peak area of the IS lay between the two preceding conditions. The values obtained for $d_2$ can be seen in Table 3.2.5 and were calculated from the IS peak areas collected in Table 3.2.1.

**Table 3.2.4.** Experimental matrix and values of the relative peak area for each of the seven analytes. The levels of every factor are codified in the same way as in Table 3.2.1.

| Exp. | Level | | | | CAR | XYL | AZOL | AZA | HAL | PROP | CHLOR |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Factor 1 | Factor 2 | Factor 3 | Factor 4 | | | | | | | |
| 1 | B | A | A | A | 8.03 | 26.19 | 2.59 | 3.29 | 3.69[a] | 5.51 | 0.92 |
| 2 | A | B | A | A | 7.05 | 17.24 | 2.52 | 3.20 | 3.63 | 4.11 | 0.92 |
| 3 | A | A | B | A | 7.17 | 32.21 | 2.90 | 3.18 | 3.19 | – [a] | – [a] |
| 4 | B | B | B | A | 11.29 | 23.46 | 3.06 | 3.32 | 3.86 | 6.18 | 1.05 |
| 5 | A | A | A | B | 8.32 | 27.08 | 2.75 | 3.26[a] | 3.80[a] | 5.82 | 1.00 |
| 6 | B | A | A | B | 7.33[a] | 19.66 | 2.71 | 3.29 | 3.39 | 5.68 | 1.09[a] |
| 7 | A | B | B | B | 10.28 | 24.50 | 2.72 | 3.21 | 3.75 | 5.53[a] | 1.06 |
| 8 | B | B | B | B | 13.95 | 30.71 | 2.96 | 3.26 | 3.79 | 6.15 | 1.07 |
| 9 | A | B | A | C | 7.02 | 14.46 | 2.68 | 3.26 | 3.66 | 4.75 | 0.99 |
| 10 | B | B | A | C | 7.38 | 13.66 | 2.81 | 3.35 | 3.63 | 4.47[a] | 0.98 |
| 11 | B | A | B | C | 11.18 | 33.60 | 2.78 | 3.21[a] | 3.45 | 6.14 | 1.07 |

[a] Outlier datum.

**Table 3.2.5.** Values of the individual desirability functions $d_1$ (%), $d_2$ (%) and of the overall desirability $D$ (%) for each of the seven analytes. The number of the experiment and the levels of every factor are the same as in Table 3.2.1.

| Exp | CAR | | | XYL | | | AZOL | | | AZA | | | HAL | | | PROP | | | CHLOR | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $d_1$ | $d_2$ | $D$ | $d_1$ | $d_2$ | $D$ | $d_1$ | $d_2$ | $D$ | $d_1$ | $d_2$ | $D$ | $d_1$ | $d_2$ | $D$ | $d_1$ | $d_2$ | $D$ | $d_1$ | $d_2$ | $D$ |
| 1 | 66.5 | 1.9 | 11.3 | 0 | 1.9 | 0 | 84.5 | 78.0 | 81.2 | 0 | 78.0 | 0 | 58.5 | 1.9 | 10.6 | 69.3 | 0 | 0 | 27.7 | 0 | 0 |
| 2 | 89.2 | 92.9 | 91.0 | 0 | 92.9 | 0 | 91.1 | 100 | 95.4 | 8.1 | 100 | 28.4 | 53.0 | 92.9 | 70.2 | 36.1 | 100 | 60.1 | 30.0 | 100 | 54.8 |
| 3 | 86.5 | 0 | 0 | 0 | 0 | 0 | 57.1 | 28.0 | 40.0 | 12.3 | 28.0 | 18.5 | 11.8 | 0 | 0 | – | 0 | 0[a] | – | 0 | 0[a] |
| 4 | 0 | 20.3 | 0 | 0 | 20.3 | 0 | 43.0 | 53.7 | 48.0 | 0 | 53.7 | 0 | 74.6 | 20.3 | 38.9 | 85.4 | 0 | 0 | 66.6 | 0 | 0 |
| 5 | 59.7 | 22.5 | 36.7 | 0 | 22.5 | 0 | 70.0 | 100 | 83.7 | 0 | 100 | 0 | 69.3 | 22.5 | 39.5 | 76.9 | 0 | 0 | 54.3 | 0 | 0 |
| 6 | 82.7 | 68.2 | 75.1[b] | 0 | 68.2 | 0 | 74.3 | 88.0 | 80.9 | 0 | 88.0 | 0 | 31.2 | 68.2 | 46.1 | 73.5 | 0 | 0 | 78.6 | 0 | 0 |
| 7 | 13.9 | 44.9 | 25.0 | 0 | 44.9 | 0 | 73.5 | 100 | 85.7 | 6.2 | 100 | 24.8 | 65.0 | 44.9 | 54.0 | 69.9 | 11.32 | 28.1 | 68.9 | 11.3 | 27.9 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 51.5 | 70.5 | 60.3 | 0 | 70.5 | 0 | 68.5 | 0 | 0[b] | 84.6 | 0 | 0 | 72.9 | 0 | 0 |
| 9 | 89.9 | 100 | 94.8 | 7.9 | 100 | 28.1 | 76.3 | 100 | 87.4 | 0 | 100 | 0[b] | 56.5 | 100 | 75.1 | 51.5 | 100 | 71.8 | 50.0 | 100 | 70.7 |
| 10 | 81.5 | 100 | 90.3 | 20.9 | 100 | 45.7 | 64.7 | 100 | 80.5 | 0 | 100 | 0 | 53.4 | 100 | 73.1 | 44.8 | 100 | 67.0 | 44.9 | 100 | 67.1 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0[b] | 68.3 | 53.4 | 60.4 | 6.4 | 53.4 | 18.5 | 36.3 | 0 | 0 | 84.4 | 0 | 0 | 70.6 | 0 | 0 |

[a] Assigned a value of the overall desirability $D$ equal to zero even though $d_1$ could not be calculated.

[b] Influential datum.

Finally, to take both individual desirability functions into account, the geometric mean of these was calculated to obtain the overall desirability function *D* (see values in Table 3.2.5), so that one criterion could not be amended by the other.

In the particular cases of xylazine and azaperone, $d_1$ was different from 0% only in two and four out of the eleven experiments performed, respectively, i.e., changes in the four factors introduced quite severe alterations to the values of the $A_{analyte}/A_{IS}$ ratio with regard to the validated ones. $d_2$, which regularized those cases in which $A_{IS}$ became too small, was equal to 0% in experiments 8 and 11 for all the analytes except for azaperol and azaperone. The extreme case was reflected in Experiment 3, since neither propionylpromazine nor chlorpromazine nor their IS were recovered (the absolute peak areas $A_{PROP}$, $A_{CHLOR}$ and $A_{CHLOR\text{-}d3}$ equalled 0) because of the experimental conditions involved in that experiment (pH of the phosphate buffer adjusted to 7.4 and methanol/acetic acid as the SPE washing solution). Therefore, it was impossible to calculate the standardized peak areas for these two analytes (see Table 3.2.1), and the overall desirability was thus considered to be 0%.

The overall desirability function *D*, which reproduced the aim of the study, was the response used for estimating the coefficients in Eq. (3.2.3), from which the effect of the experimental factors was to be evaluated.

Since the model in Eq. (3.2.3) had 7 coefficients, $\beta_{iM}$, and 11 experimental results were available for every analyte, there were enough degrees of freedom to assess the residual variance and to decide the statistical significance of both the model and every coefficient. The null hypothesis ($H_0$) of the test for the significance of the model is that the proposed model is not significant, while the alternative hypothesis ($H_a$) poses that it is. In the case of the test for the significance of every coefficient, $H_0$ postulates that the coefficient equals 0 in contrast to $H_a$, where the coefficient is different from 0. The models built for the responses related to azaperol, propionylpromazine and chlorpromazine were significant at a 5% significance level. However, after the models for the rest of the analytes had been estimated, one influential datum was detected in every case according to the values of the studentized residuals and Cook's distance [42]. The influential data and the maximum VIF of the resulting designs were:

- For carazolol, Experiment 6 (maximum VIF = 1.60).
- For xylazine, Experiment 11 (maximum VIF = 2.21).
- For azaperone, Experiment 9 (maximum VIF = 1.20).
- For haloperidol, Experiment 8 (maximum VIF = 1.49).

Furthermore, for three out of these four analytes (xylazine, azaperone and haloperidol), the desirability *D* associated with its respective influential experiment was 0%, as can be seen in Table 3.2.5. Once these data had been removed, the new models turned out to be significant at a 5% significance level (*p*-level between 0.014 and 0.044). The models fitted to the data of the 7 experimental responses had values of $R^2$ that ranged from 0.92 to 0.97.

Although the model in Eq. (3.2.3) is that obtained by least squares regression, the interpretation of the effect of the experimental factors will be easier if the model considered, after a change of variables, is the one in Eq. (3.2.4):

$$\begin{aligned} y = {} & \beta'_0 + \beta'_{1A} X_{1A} + \beta'_{1B} X_{1B} + \beta'_{2A} X_{2A} + \beta'_{2B} X_{2B} + \beta'_{3A} X_{3A} + \beta'_{3B} X_{3B} + \beta'_{4A} X_{4A} + \\ & \beta'_{4B} X_{4B} + \beta'_{4C} X_{4C} + \beta'_{1A2A} X_{1A} X_{2A} + \beta'_{1A2B} X_{1A} X_{2B} + \beta'_{1B2A} X_{1B} X_{2A} + \\ & \beta'_{1B2B} X_{1B} X_{2B} + \varepsilon \end{aligned} \qquad (3.2.4)$$

where $X_{iM}$ , $i$ = 1,…4 , $M$ = A, B, C, is 1 if the factor $i$ is at level $M$ and 0 if not. $\beta'_{iM}$ is therefore a quantity that is added to the response when the factor $i$ is at level $M$. A description of a screening design and the conversion of Eq. (3.2.4) to Eq. (3.2.3) can be seen in pages 344 to 346 and 469 to 472 in [40], respectively. Briefly explained, by taking into account the interdependence of the presence-absence variables in Eq. (3.2.4), that is,

$$X_{1A} + X_{1B} = X_{2A} + X_{2B} = X_{3A} + X_{3B} = X_{4A} + X_{4B} + X_{4C} = 1 \qquad (3.2.5)$$

the following relationships are obtained:

$$\left\{ \begin{array}{c} \beta'_{1B} = -\left(1/2\right) \beta_{1B\backslash A} \\ \beta'_{1A} = \beta_{1B\backslash A} + \beta'_{1B} = \left(1/2\right) \beta_{1B\backslash A} \\ \beta'_{2B} = -\left(1/2\right) \beta_{2B\backslash A} \\ \beta'_{2A} = \beta_{2B\backslash A} + \beta'_{2B} = \left(1/2\right) \beta_{2B\backslash A} \\ \beta'_{3B} = -\left(1/2\right) \beta_{3B\backslash A} \\ \beta'_{3A} = \beta_{3B\backslash A} + \beta'_{3B} = \left(1/2\right) \beta_{3B\backslash A} \\ \beta'_{4B} = -\left(1/3\right) (\beta_{4C\backslash A} + \beta_{4C\backslash B}) = -\left(1/3\right) \beta_{4C\backslash A} - \left(1/3\right) \beta_{4C\backslash B} \\ \beta'_{4A} = \beta_{4C\backslash A} - \beta'_{4C} = \left(2/3\right) \beta_{4C\backslash A} - \left(1/3\right) \beta_{4C\backslash B} \\ \beta'_{4B} = \beta_{4C\backslash B} - \beta'_{4C} = -\left(1/3\right) \beta_{4C\backslash A} + \left(2/3\right) \beta_{4C\backslash B} \end{array} \right. \qquad (3.2.6)$$

In general, if the *i*-th factor has $l_i$ levels ($i$=1,…,$k$), *M* is the reference level (level B or C above) and *N* are the others (levels A or A and B above), then:

$$\left\{ \begin{array}{c} \beta'_{iM} = -\frac{1}{l_i}\sum_{N=1}^{l_i-1}\beta_{iM\backslash N} \\ \beta'_{iN} = \beta_{iM\backslash N} - \beta'_{iM} \\ \beta'_0 = \beta_0 - \sum_{i=1}^{k}\beta'_{iM} \end{array} \right\} \quad (3.2.7)$$

The values of the coefficients in Eq. (3.2.4) for each of the four factors and for each of the seven analytes are illustrated in Figure 3.2.4. Dash-dotted lines in this figure represent the borders of the critical region of the test for the significance of every coefficient at 5%. To get a response as great as possible, every factor should be at the level in which its effect is positive (or in which its effect is the most positive for a 3-level factor), being positive those coefficients on the right of the vertical line shown for each analyte.

From the analysis of this figure, the following conclusions were drawn with the aim of achieving the maximum desirability:

- For Factor 1, it was clear that the SPE flow rate was not significant for either xylazine or azaperol or haloperidol, while it was for both carazolol and azaperone. On the other hand, it was nearly significant for both propionylpromazine and chlorpromazine. This factor had the same behaviour for these last four analytes, and it should be at level A (10 mL/min).

- Factor 2 was not significant for either xylazine or azaperol or azaperone, whereas it was for the rest of the analytes. Its effect was positive when it was at level B (pH of the phosphate buffer equal to 6). The results obtained in this work show that the acid buffer performed better than the basic one. The reason might be that the basic buffer also extracted more basic matrix compounds (interferents) that were not removed at the purification step and inhibited the ionization of the tranquillizers during the analysis by LC-MS/MS.

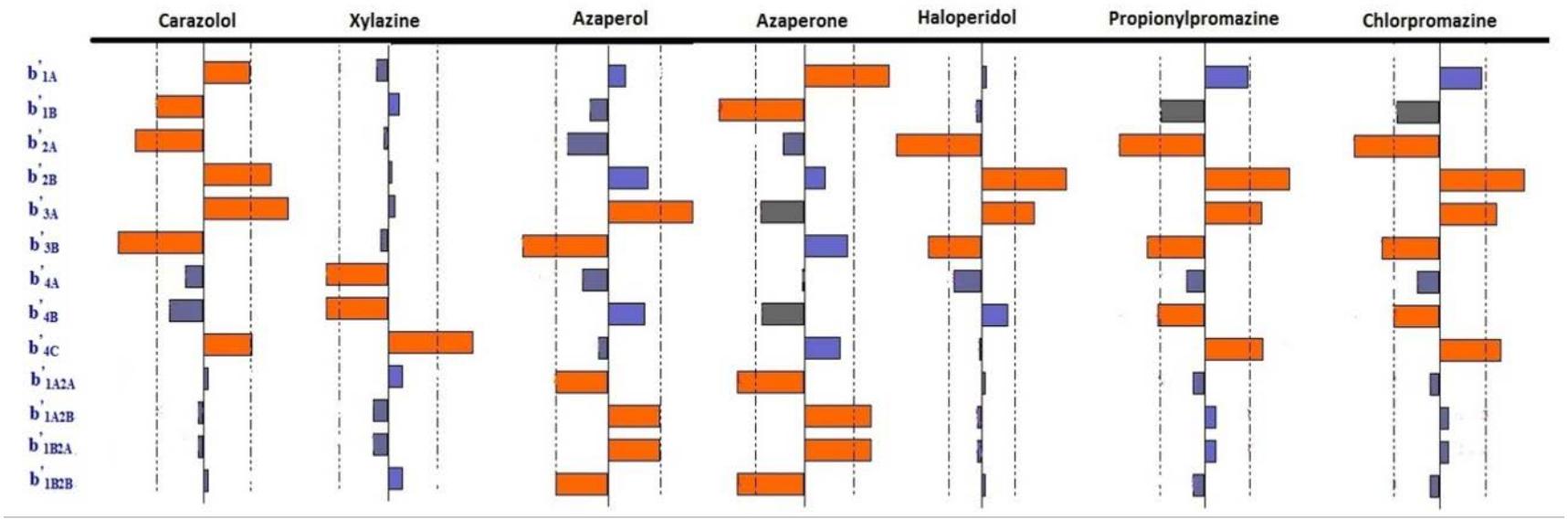


**Figure 3.2.4.** Effect of the factors on the overall desirability *D* for each analyte. Dash-dotted lines indicate the thresholds beyond which factors are significant at a 95% confidence level**.**

– Except for both xylazine and azaperone, Factor 3 was significant for the remaining drugs. In this case, this factor should best be at level A, since methanol/formic acid seemed more appropriate as SPE washing solution than methanol/acetic acid. Formic acid is stronger than acetic acid. During the washing step, analytes are wanted to get retained in the cartridge, but some losses always occur. This undesirable elution of basic compounds should be smaller with stronger acids. That might be the reason why formic acid performed better than acetic acid. This result was in accordance with the basic properties of the tranquillizers.

– Whenever Factor 4 was significant (carazolol, xylazine, propionylpromazine and chlorpromazine), its behaviour showed the same curvature, so that it would be advisable to use an elution volume equal to 5 mL (level C).

– Interaction 1-2 was only significant for azaperol and azaperone, for which just two out of the four combinations should be considered as they both had a positive effect on the overall desirability.

The results of the analysis of Figure 3.2.4 are summarized in Table 3.2.6 (Case 1 (*D*)). The significant factors are marked in bold upper case and the non-significant ones in plain lower case.

**Table 3.2.6.** Results of the analysis: level that represents the value at which the corresponding factor should be. The levels of the significant factors are written in bold upper case. The levels of the non-significant factors are in lower case. The response considered in each case appears between brackets in the head of the table (*D* is the overall desirability function and *A* indicates the absolute peak area).

| Analyte | Case 1 (*D*) | | | | Case 2 ($A_{analyte}$) | | | | Case 3 ($A_{analyte}/A_{IS}$) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Factor 1** | **Factor 2** | **Factor 3** | **Factor 4** | **Factor 1** | **Factor 2** | **Factor 3** | **Factor 4** | **Factor 1** | **Factor 2** | **Factor 3** | **Factor 4** |
| Carazolol | **A** | **B** | **A** | **C** | b | **B** | a | b | b | b | b | b |
| Xylazine | b | b | a | **C** | **A** | a | b | **B** | a | **A** | **B** | b |
| Azaperol | a | b | **A** | b | **A** | **B** | **A** | **C** | b | b | **B** | c |
| Azaperone | **A** | b | b | c | **A** | **B** | **A** | **C** | **B** | b | b | c |
| Haloperidol | a | **B** | **A** | b | a | **B** | **A** | c | b | **B** | b | a |
| Propionylpromazine | a | **B** | **A** | **C** | a | **B** | **A** | **C** | b | a | a | c |
| Chlorpromazine | a | **B** | **A** | **C** | a | **B** | **A** | **C** | a | a | **B** | c |
| Azaperone-$d_4$ | --- | --- | --- | --- | **A** | **B** | **A** | **C** | --- | --- | --- | --- |
| Haloperidol-$d_4$ | --- | --- | --- | --- | a | **B** | **A** | c | --- | --- | --- | --- |
| Chlorpromazine-$d_3$ | --- | --- | --- | --- | a | **B** | **A** | **C** | --- | --- | --- | --- |
| Conditions in which the extraction/ pretreatment should be carried out | **A** | **B** | **A** | **C** | **A** | **B** | **A** | **C** | **B** | **A** or **B** | **B** | c |

As an example, Figure 3.2.5 shows the total ion chromatogram (TIC) and MRM chromatograms of every analyte and internal standard for the most intense transition obtained in a spiked sample pretreated under these optimum conditions. This sample was fortified with a concentration equal to the decision limit for the banned substances [23] and at a concentration equal to the limit of quantification for the permitted substances.

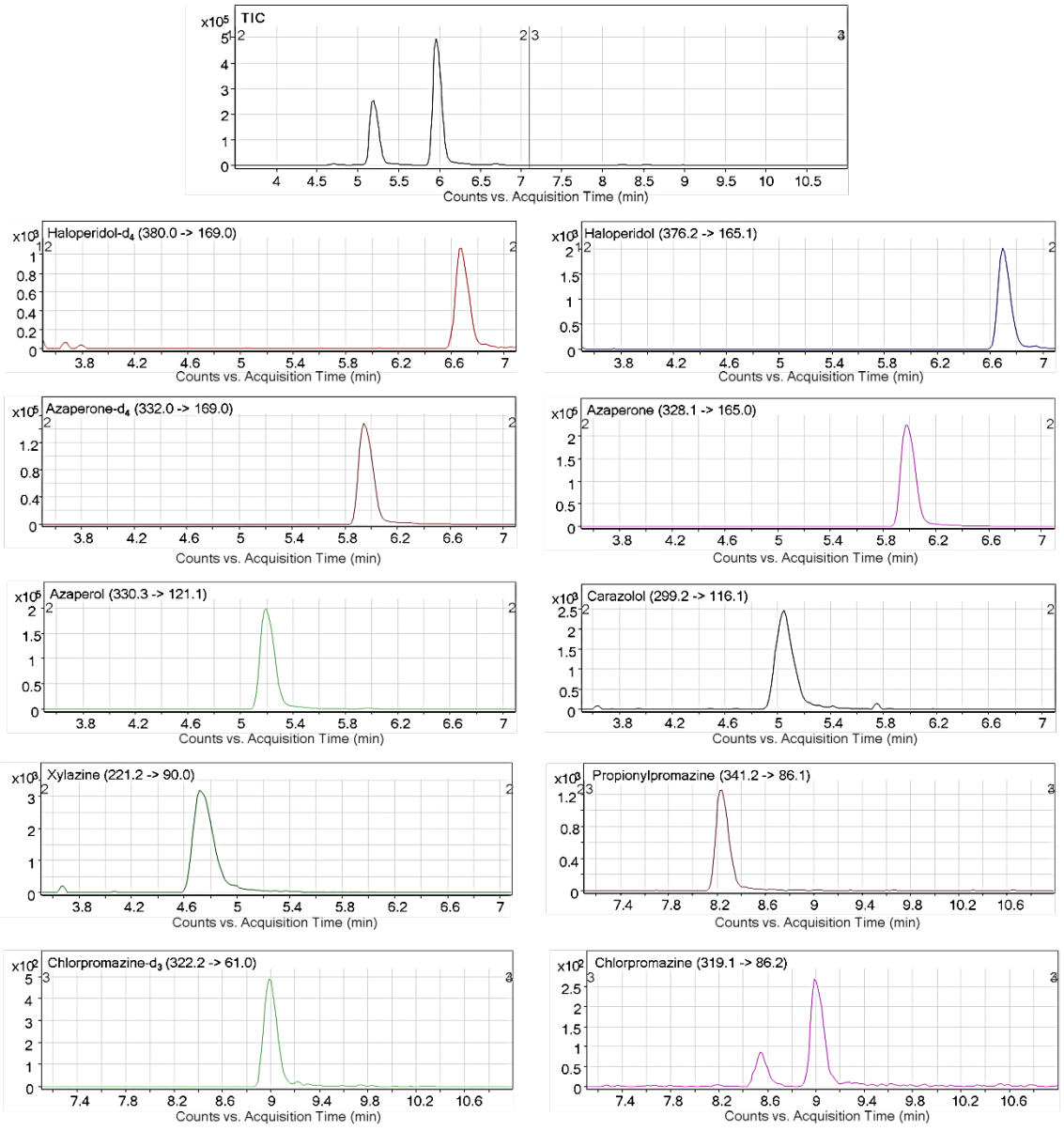


**Figure 3.2.5.** Total ion chromatogram (TIC) and MRM chromatograms of the seven analytes and of the three internal standards for the most intense transition in a spiked sample pretreated under the optimum conditions found in Case 1. The fortification levels were 0.50 μg $kg^{-1}$ for haloperidol, 25 μg $kg^{-1}$ for azaperone, 20 μg $kg^{-1}$ for azaperol, 1.30 μg $kg^{-1}$ for carazolol, 0.86 μg $kg^{-1}$ for xylazine, 1.40 μg $kg^{-1}$ for propionylpromazine and 0.90 μg $kg^{-1}$ for chlorpromazine.

3.2.5.1.2. Case 2

In this case, the absolute peak area of all seven analytes and the three internal standards was maximized. The experimental results are listed in Table 3.2.1. As it can be seen, the conditions in Experiment 3 were too extreme for propionylpromazine, chlorpromazine and chlorpromazine-$d_3$ and these compounds were not detected. All the models estimated from the ten experimental responses were significant at a 90% confidence level without any need for influential data to be removed. Depending on the analyte, the values of the coefficient of determination ($R^2$) ranged from 0.88 to 1.0.

The estimates of the coefficients $\beta'_{iM}$ (Eq. (3.2.4)) for each of the four factors and for each of the ten compounds are shown in Figure 3.2.6.

- Factor 1 (SPE flow rate) was only significant for xylazine, azaperol, azaperone and azaperone-$d_4$. The maximum peak area would be obtained at level A (10 mL/min).
- Factor 2 (pH of the extracting solvent) was significant for all the analytes and internal standards but for xylazine. In consequence, greater peak areas would be achieved at level B (pH 6).
- Factor 3 (SPE washing solution) was significant for all compounds except for carazolol and xylazine. The greatest peak areas would be obtained at level A (methanol/formic acid as washing solution).
- Factor 4 (volume of eluting solution) was significant for all the analytes but for carazolol, haloperidol and haloperidol-$d_4$. The highest peak areas for xylazine were obtained at level B (4 mL), which shows that the response of this factor was not linear: when the elution was performed with 4 mL instead of 3 mL, more xylazine was recovered, but when 5 mL instead of 4 mL was used, less xylazine was eluted. For the rest of the compounds, the highest peak areas were obtained at level C (5 mL).
- Interaction 1-2 was significant for both azaperone and azaperone-$d_4$. For both analytes, the highest peak areas were obtained when the SPE flow rate (Factor 1) was 10 mL/min and the pH (Factor 2) was 6.

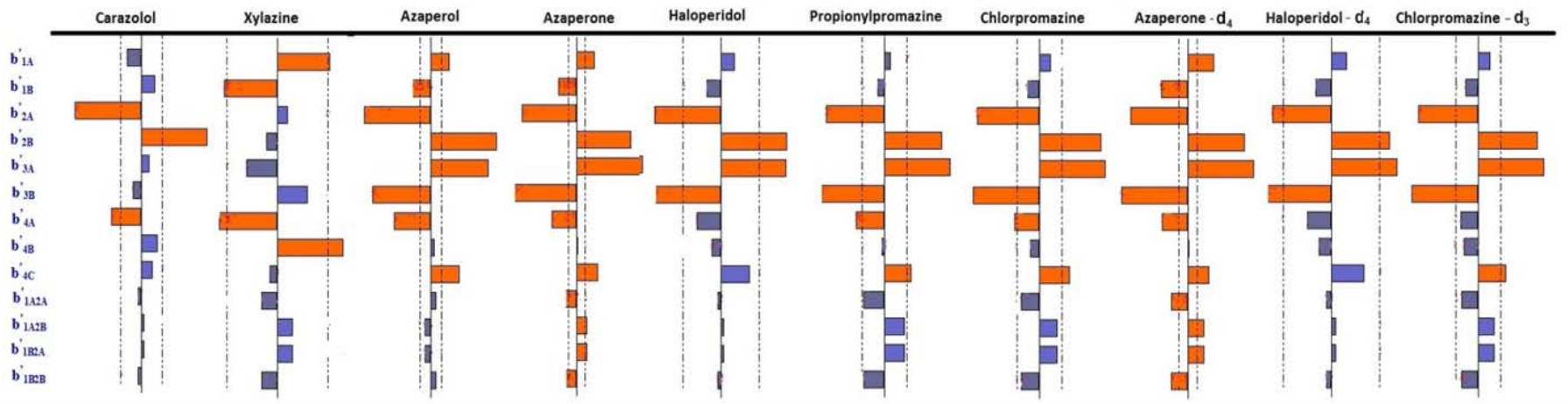


**Figure 3.2.6.** Effect of the factors on the absolute peak area for each analyte. Dash-dotted lines indicate the thresholds beyond which factors are significant at a 95% confidence level.

Table 3.2.6 (Case 2 ($A_{analyte}$)) summarizes the level of each factor that maximizes the peak area of every analyte. As can be seen in this table, these experimental conditions were the same as the ones obtained when the overall desirability function *D* was used as response (Case 1). However, with an eye on a following step of quantification, the response considered should be the overall desirability function *D* proposed in this work, because the issues concerning the $A_{analyte}/A_{IS}$ ratio are taken into account with that strategy.

3.2.5.1.3. Case 3

As it has been explained in Section 3.2.2, it is possible to study the effect of the factors on the standardized peak area, because internal standardization is an adequate tool for quantifying analytes in chromatography. When absolute signals are maximized, as in Case 2, the signal of the IS is also maximized. As a result, the standardized area could be lower for some of the analytes due to the different chemical behaviour of the analyte and its IS. Standardized areas of the seven compounds are shown in Table 3.2.4 and they fitted to the model in Eq. (3.2.3).

Models estimated for both xylazine and azaperol were significant at a 90% confidence level. The rest of the models turned out to be significant after the removal of some data. The outliers detected are shown in Table 3.2.4 with a lower case superscript letter. The seven final models had values of maximum VIF between 1.59 and 7.03 and of $R^2$ between 0.86 and 1.0.

If the aim of the analysis was to achieve a value for the $A_{analyte}/A_{IS}$ ratio as high as possible, the conclusions of the evaluation of the effects together with the corresponding condition (level) at which every factor should be for every analyte are collected in Table 3.2.6 (Case 3 ($A_{analyte}/A_{IS}$)). The coefficients of the fitted models are shown in Figure 3.2.7.

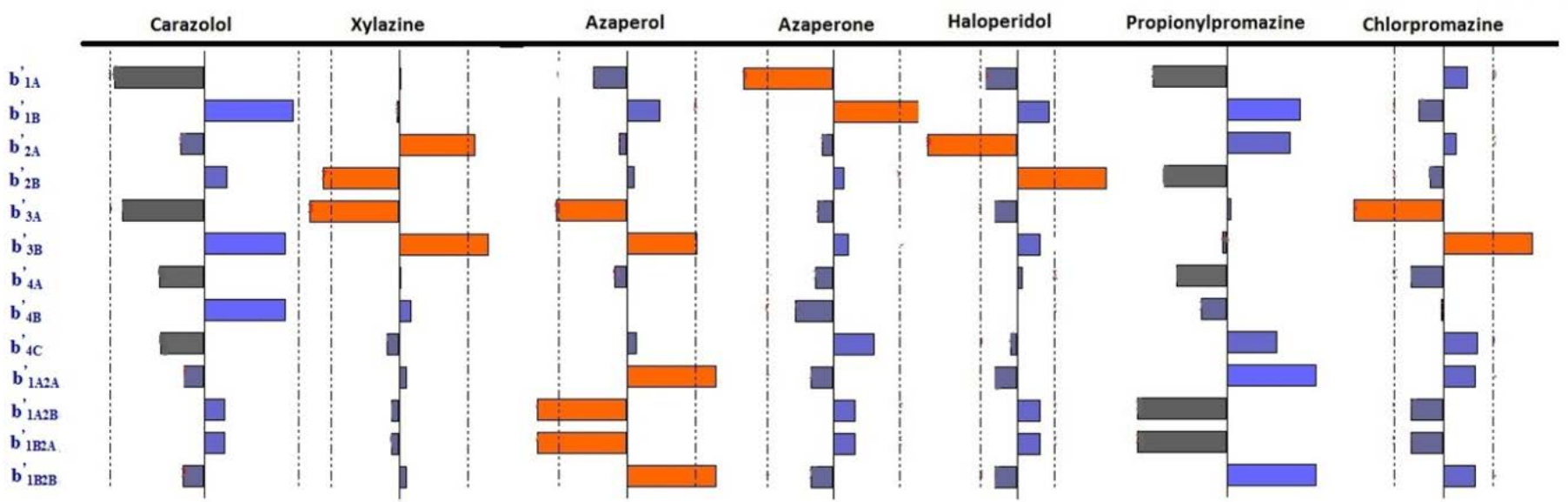


**Figure 3.2.7.** Effect of the factors on the standardized peak area for each analyte. Dash-dotted lines indicate the thresholds beyond which factors are significant at a 95% confidence level.

- Factor 1 (SPE flow rate) was only significant for azaperone and it should be at level B (2 mL/min); this result was in conflict with that obtained in Case 1 and Case 2.
- The pH of the phosphate buffer (Factor 2) was significant for xylazine and haloperidol, but for the former it should be at level A (pH = 7.4), while for the latter at level B (pH = 6). The level required for xylazine was different from that obtained in Case 1, but in Case 3 this factor was significant, while the level for haloperidol was the same as the one obtained in the two previous cases.
- Factor 3 (SPE washing solution) was significant for xylazine, azaperol and chlorpromazine and it should be at level B (methanol/acetic acid). This contradicts the result obtained in Case 1 and Case 2.
- Finally, the volume of eluting solution (Factor 4) was not significant for any tranquillizers, but for four of them 5 mL (level C) was required as in the two previous cases.
- It is deduced from the fitted model for each analyte that the interaction between Factors 1 and 2 was only significant for azaperol and that both factors should be at level B.

Overall, the experimental conditions that maximized the standardized peak area were quite different (see Table 3.2.6, Case 3) from those that maximized the overall desirability function *D* (Case 1) and the absolute peak area (Case 2). This is due to the fact that a maximum $A_{analyte}/A_{IS}$ ratio can be achieved by increasing the numerator and/or decreasing the denominator. This is an undesirable situation, because the $A_{analyte}/A_{IS}$ ratio will not be trustworthy if the value of $A_{IS}$ is not sufficient. In Case 3, the mathematical models described the response ($A_{analyte}/A_{IS}$) poorly because it is a ratio between two experimental quantities, although the statistical quality of the results was still sufficient.

It must be noticed that both the D-optimal design and the experimental data used were the same in the three cases analysed. However, the conclusions drawn regarding the level at which every factor should be were different depending on the response variable considered. This is the result of the behaviour of the analytes and the internal standards in the previous pretreatment steps.

### 3.2.6. Conclusions

The *ad hoc* procedure built to define the response variable (desirability functions) together with the experimental domain and the connection between factor levels and response (D-

optimal design) enables to adapt the methodology of experimental designs exactly to the analytical problem under study; that is, it makes performance of *ad hoc* designs possible.

In the analysis of the effect of the extraction and purification steps of azaperone, propionylpromazine, chlorpromazine, haloperidol, xylazine, azaperol and carazolol from pig muscle analysed by LC-MS/MS, the experimental effort necessary to study the effect of the SPE flow rate, the pH of phosphate buffer, the composition of the washing solution and the volume required of the eluting mixture, as well as the interaction between the first two factors, has been cut by more than 50% thanks to a D-optimal design (the number of experiments has been reduced from 24 to 11). Therefore, the methodology used in combination with the D-optimal design enables to save time and reagents considerably.

The desirability function used as experimental response quantifies the two analytical criteria of interest: i) assess a similar chemical behaviour of each analyte in relation to its internal standard and ii) avoid a significant reduction of the peak area of the internal standards. The optimum conditions for the determination of the analytes are achieved when the SPE flow rate is 10 mL/min, the pH of the phosphate buffer is 6, the rinsing solution is methanol/formic acid and the elution volume is 5 mL.

It is shown that the significance of the factors and their optimal levels depend on the experimental response considered. Although the absolute area may seem the best choice, the standardized peak area should also be taken into account to correct for variability in the experimental conditions. Moreover, considering the standardized peak area along with the quality of the signals of both the IS and the analyte is a more appropriate and flexible approach that reflects the analytical problem better.

### 3.2.7. Acknowledgments

The authors thank the financiation provided by MINECO (CTQ2014-53157-R). L. Rubio and M.L. Oca are grateful to University of Burgos for their FPI grants.

### 3.2.8. References

[1] P. Delahaut, C. Levaux, P. Eloy, M. Dubois, *Validation of a method for detecting and quantifying tranquillisers and a β-blocker in pig tissues by liquid chromatography–tandem mass spectrometry*, Analytica Chimica Acta 483 (2003) 335-340.

[2] N.A. Botsoglou, D.J. Fletouris, *Drug Residues in Foods: Pharmacology, food safety and analysis*. Marcel Dekker Inc.: New York, USA, 2001.

[3] S.B. Turnipseed, *Drug residues in meat: emerging issues*, in Y.H. Hui, W.K. Nip, R.W. Rogers, O.A. Young (Eds.), *Meat science and applications*, Marcel Dekker, Inc., New York, 2001.

[4] Commission Regulation (EU) No 37/2010 of 22 December 2009 on pharmacologically active substances and their classification regarding maximum residue limits in foodstuffs of animal origin, Official Journal of the European Union, L 15, 1-72.

[5] Regulation (EC) No 470/2009 of the European Parliament and of the Council of 6 May 2009 laying down Community procedures for the establishment of residue limits of pharmacologically active substances in foodstuffs of animal origin, Official Journal of the European Union, L 152,11-22.

[6] E. Rattenberger, P. Matzke, J. Neudegger, *Development of a radioimmunoassay for the detection of residues of the beta adrenoreceptor blocking agent 'Carazolol' in blood and urine of pigs*, Archiv fuer Lebensmittelhygiene 36 (1985) 85-87.

[7] S.A. Meenagh, J.D.G. McEvoy, C.T. Elliott, Determination of carazolol residues in porcine tissue by radioreceptor assay, Analytica Chimica Acta 462 (2002) 149-156.

[8] J. Cooper, P. Delahaut, T.L. Fodey, C.T. Elliot, *Development of a rapid screening test for veterinary sedatives and the beta-blocker carazolol in porcine kidney by ELISA*, Analyst 129 (2004) 169-174.

[9] N. Haagsma, E.R. Bathelt, J.W. Engelsma, *Thin-layer chromatographic screening method for the tranquillizers azaperone, propiopromazine and carazolol in pig tissues,* Journal of Chromatography 436 (1988) 73-79.

[10] L.A. Van Ginkel, P.L.W.J. Schwillens, M. Olling, *Liquid chromatographic method with on-line UV spectrum identification and off-line thin-layer chromatographic confirmation for the detection of tranquillizers and carazolol in pig kidneys*, Analytica Chimica Acta 225 (1989) 137-146.

[11] H.J. Keukens, M.M.L. Aerts, *Determination of residues of carazolol and a number of tranquillizers in swine kidney by high-performance liquid chromatography with ultraviolet and fluorescence detection*, Journal of Chromatography A 464 (1991) 149-161.

[12] Y. Aoki, H. Hakamata, Y. Igarashi, K. Uchida, H. Kobayashi, N. Hirayama, A. Kotani, F. Kusu, *Simultaneous determination of azaperone and azaperol in animal tissues by HPLC with confirmation by electrospray ionization mass spectrometry*, Journal of Chromatography B 877 (2009) 166-172.

[13] V. Cerkvenik-Flajs, *Determination of residues of azaperone in the kidneys by liquid chromatography with fluorescence detection*, Analytica Chimica Acta 586 (2007) 374-382.

[14] F. Kadir, J. Zuidema, *High-performance liquid chromatographic method for the determination of carazolol in serum or plasma of pigs*, Journal of Chromatography 527 (1990) 461-466.

[15] M.D. Rose, G. Shearer, *Determination of tranquilisers and carazolol residues in animal tissue using high-performance liquid chromatography with electrochemical detection*, Journal of Chromatography A 624 (1992) 471-477.

[16] Commission Decision 2002/657/EC of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Official Journal of the European Communities, L 221, 8-36.

[17] M.L. Olmos-Carmona, M. Hernández-Carrasquilla, *Gas chromatographic–mass spectrometric analysis of veterinary tranquillizers in urine: evaluation of method performance*, Journal of Chromatography B 734 (1999) 113-120.

[18] Y. Govaert, P. Batjoens, K. Tsilikas, J.M. Degroodt, S. Srebrnik, *Multi-residue analysis of tranquillizers in meat: confirmatory assays using mass spectrometry*, Analyst 123 (1998) 2507-2512.

[19] M.H. Spyridaki, E. Lyris, I. Georgoulakis, D. Kouretas, M. Konstantinidou, C.G. Georgakopoulos, *Determination of xylazine and its metabolites by GC-MS in equine urine for doping analysis*, Journal of Pharmaceutical and Biomedical Analysis 35 (2004) 107-116.

[20] J. Zhang, B. Shao, J. Yin, Y. Wu, H. Duan, *Simultaneous detection of residues of β-adrenergic receptor blockers and sedatives in animal tissues by high-performance liquid chromatography/tandem mass spectrometry*, Journal of Chromatography B 877 (2009) 1915-1922.

[21] K. Mitrowska, A. Posyniak, J. Zmudzki, *Rapid method for the determination of tranquilizers and a beta-blocker in porcine and bovine kidney by liquid chromatography with tandem mass spectrometry*, Analytica Chimica Acta 637 (2009) 185-192.

[22] C. Bock, C.S. Stachel, *Development and validation of a confirmatory method for the determination of tranquilisers and a β-blocker in porcine and bovine kidney by LC-MS/MS*, Food Additives & Contaminants: Part A 30 (2013) 1000-1011.

[23] Council Directive 96/23/EC of 29 April 1996 on measures to monitor certain substances and residues thereof in live animals and animal products, Official Journal of the European Communities, L 125, 10-32.

[24] Q. Alan Xu, T.L. Madden (Eds.), *LC-MS in drug bioanalysis*, Springer Science & Business Media, New York, USA, 2012.

[25] A. Tan, K. Awaiye, *Use of internal standards in LC-MS bioanalysis*, in: W. Li, J. Zhang, F.L.S. Tse (Eds.), *Handbook of LC-MS bioanalysis: Best practices, experimental protocols, and regulations*, John Wiley & Sons Inc., Hoboken, USA, 2013, pp. 217-227.

[26] K.J. Bronsema, R. Bischoff, N.C. van de Merbel, *Internal standards in the quantitative determination of protein biopharmaceuticals using liquid chromatography coupled to mass spectrometry*, Journal of Chromatography B 893-894 (2012) 1-14.

[27] M. Jalali-Heravi, H. Parastar, H. Ebrahimi-Najafabadi, *Characterization of volatile components of Iranian saffron using factorial-based response surface modeling of ultrasonic extraction combined with gas chromatography–mass spectrometry analysis*, Journal of Chromatography A 1216 (2009) 6088–6097.

[28] H. Sereshti, R. Heidari, S. Sadami, *Determination of volatile components of saffron by optimized ultrasound-assisted extraction in tandem with dispersive liquid–liquid microextraction followed by gas chromatography–mass spectrometry*, Food Chemistry 143 (2014) 499-505.

[29] G. Pieraccini, S. Furlanetto, S. Orlandini, G. Bartolucci, I. Giannini, S. Pinzauti, G. Moneti, *Identification and determination of mainstream and sidestream smoke components in different brands and types of cigarettes by means of solid-phase microextraction–gas chromatography–mass spectrometry*, Journal of Chromatography A 1180 (2008) 138-150.

[30] C. Cortada, L. Vidal, A. Canals, *Determination of nitroaromatic explosives in water samples by direct ultrasound-assisted dispersive liquid–liquid microextraction followed by gas chromatography–mass spectrometry*, Talanta 85 (2011) 2546-2552.

[31] M.I. Beser, J. Beltrán, V. Yusà, *Design of experiment approach for the optimization of polybrominated diphenyl ethers determination in fine airborne particulate matter by microwave-assisted extraction and gas chromatography coupled to tandem mass spectrometry*, Journal of Chromatography A 1323 (2014) 1-10.

[32] E. Gionfriddo, A. Naccarato, G. Sindona, A. Tagarelli, *Determination of hydrazine in drinking water: Development and multivariate optimization of a rapid and simple solid phase microextraction-gas chromatography-triple quadrupole mass spectrometry protocol*, Analytica Chimica Acta 835 (2014) 37-45.

[33] D. Feng, L. Liu, L. Zhao, Q. Zhou, T. Tan, *Determination of volatile nitrosamines in latex products by HS-SPME-GC-MS*, Chromatographia 74 (2011) 817-825.

[34] C. Cortada, L. Vidal, S. Tejada, A. Romo, A. Canals, *Determination of organochlorine pesticides in complex matrices by single-drop microextraction coupled to gas chromatography–mass spectrometry*, Analytica Chimica Acta 638 (2009) 29-35.

[35] A. Pérez Antón, A.M. Casas Ferreira, C. García Pinto, B. Moreno Cordero, J.L. Pérez Pavón, *Headspace generation coupled to gas chromatography–mass spectrometry for the automated determination and quantification of endogenous compounds in urine. Aldehydes as possible markers of oxidative stress*, Journal of Chromatography A 1367 (2014) 9-15.

[36] G. Lespes, V. Desauziers, C. Montigny, M. Potin-Gautier, *Optimization of solid-phase microextraction for the speciation of butyl- and phenyltins using experimental designs*, Journal of Chromatography A 826 (1998) 67-76.

[37] A. Melo, A. Aguiar, C. Mansilha, O. Pinho, I. Ferreira, *Optimisation of a solid-phase microextraction/HPLC/diode array method for multiple pesticide screening in lettuce*, Food Chemistry 130 (2012) 1090-1097.

[38] G. Derringer, R. Suich, *Simultaneous optimization of several response variables*, Journal of Quality Technology 12 (1980) 214-219.

[39] L.A. Sarabia, M.C. Ortiz, *Response surface methodology*, in S.D. Brown, R. Tauler, B. Walczak (Eds.), *Comprehensive Chemometrics: Chemical and biochemical data analysis*, Elsevier, Amsterdam, Holland, 2009, pp. 378-382.

[40] G.A. Lewis, D. Mathieu, R. Phan-Tan-Luu, *Pharmaceutical experimental design*, Marcel Decker Inc., New York, USA, 1999.

[41] D. Mathieu, J. Nony, R. Phan-Tan-Luu, *NemrodW (Versión 2007_03)*, L.P.R.A.I., Marseille, France, 2007.

[42] N.R. Draper, H. Smith, *Applied regression analysis*, John Wiley and Sons, New York, 3rd ed., 1998, pp. 210-214.

## 3.3. Robustness testing in the determination of seven drugs in animal muscle by liquid chromatography–tandem mass spectrometry[3]

### 3.3.1. Abstract

In this work, the robustness of the sample preparation procedure for the determination of six tranquillizers (xylazine, azaperone, propionylpromazine, chlorpromazine, haloperidol, and azaperol) and a beta-blocker (carazolol) in animal muscle by LC-MS/MS was assessed through the experimental design methodology. A $2_{III}^{7-4}$ fractional factorial design was performed to evaluate the influence of seven variables on the final concentration of the seven drugs in the samples, in accordance with what is laid down in Commission Decision No 2002/657/EC. The variation considered for each of those seven factors is likely to happen when preparing the samples, being the values chosen as level −1 the nominal operating conditions. The results of the experimentation were evaluated from different statistical strategies, such as hypothesis testing using an external variance previously estimated, Lenth's method, and Bayesian analysis. Both Lenth's and Bayes' approaches enabled to determine the effect of every variable even though no degrees of freedom were left to estimate the residual error. The same conclusion about the robustness of the extraction step was reached from the three methodologies, namely, none of the seven factors examined influenced on the method performance significantly, so the sample preparation procedure was considered to be robust.

### 3.3.2. Introduction

Assessing the potential sources of variability in one or several responses of an analytical procedure must be a key part of method development. This involves making deliberate and small changes in nominal experimental conditions and investigating their subsequent effect on performance to identify the variables with the most significant effect and ensure that they are closely controlled when using the method [1].

From this perspective, two terms referring to the evaluation of the method performance still coexist within the scientific vocabulary: robustness and ruggedness. They have often been used as synonyms [2,3,4,5], but a distinction between both has also been drawn in accordance with their information about different features of an analytical method: its practicability and stability related to experimental physicochemical variables that are internal to the method (robustness) and its interlaboratory transferability when the variables under study are external to the method (ruggedness) [6,7,8]. In this sense, the

[3] This section is published as *M.L. Oca et al. / Chemometrics and Intelligent Laboratory Systems 151 (2016) 172–180* (4 citations up to August 30, 2021 according to *Scopus*).

International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) defines the robustness of an analytical procedure as "a measure of its capacity to remain unaffected by small, but deliberate variations in method parameters and provides an indication of its reliability during normal usage" [9]. On the other hand, the United States Pharmacopeia and The National Formulary has adopted the ICH definition of robustness and defines the ruggedness of an analytical method as "the degree of reproducibility of test results obtained by the analysis of the same samples under a variety of conditions such as different laboratories, analysts, instruments, lots of reagents, elapsed assay times, assay temperatures, or days " [10]. However, there is still some confusion in scientific journals, guidelines and monograph literature regarding the use of these words when applied to analytical methods [11].

Information about ruggedness/robustness should be indicated in the laboratory procedure [1]. Anyway, the strategy to be followed in a robustness and/or ruggedness test is the same. It involves performing a screening study usually by means of experimental designs after the identification of the potentially influential factors and the definition of their variation ranges and of the responses to be determined. At this point, conducting either a Plackett–Burman design [12] or a fractional factorial design, as in the Youden test [13], is the most frequently used procedure for robustness/ruggedness evaluation. The choice of the design to be performed depends on the purpose of the test and on the number of factors to be examined [4]. Due to the minimum time and analytical effort required, Commission Decision (EC) No 2002/657 [14] encourages the application of the Youden approach to the compulsory verification of that performance characteristic both in screening and confirmatory methods for the monitoring of certain substances and residues thereof in live animals and animal products. Many examples in this field can be found in the literature [6,15,16,17,18,19]. The Eurachem Guide to Method Validation and Related Topics [1] recommends, whenever possible, the evaluation of the ruggedness/robustness of a method by using the Youden test. IUPAC [2] also recognizes the strategy described by Youden as adequate to study the ruggedness of an analytical method.

This work shows the evaluation of the performance of the sample preparation step prior to the simultaneous determination of seven drugs in animal muscle by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). More precisely, the substances analysed were five tranquillizers (xylazine (XYL), azaperone (AZA), propionylpromazine (PROP), chlorpromazine (CHLOR) and haloperidol (HAL)), one of the metabolites of azaperone (azaperol (AZOL), which is derived from the former by reduction), and a blocker agent of the $\beta$-adrenergic receptor (carazolol (CAR)). Their chemical structures are depicted in Figure 3.3.1.

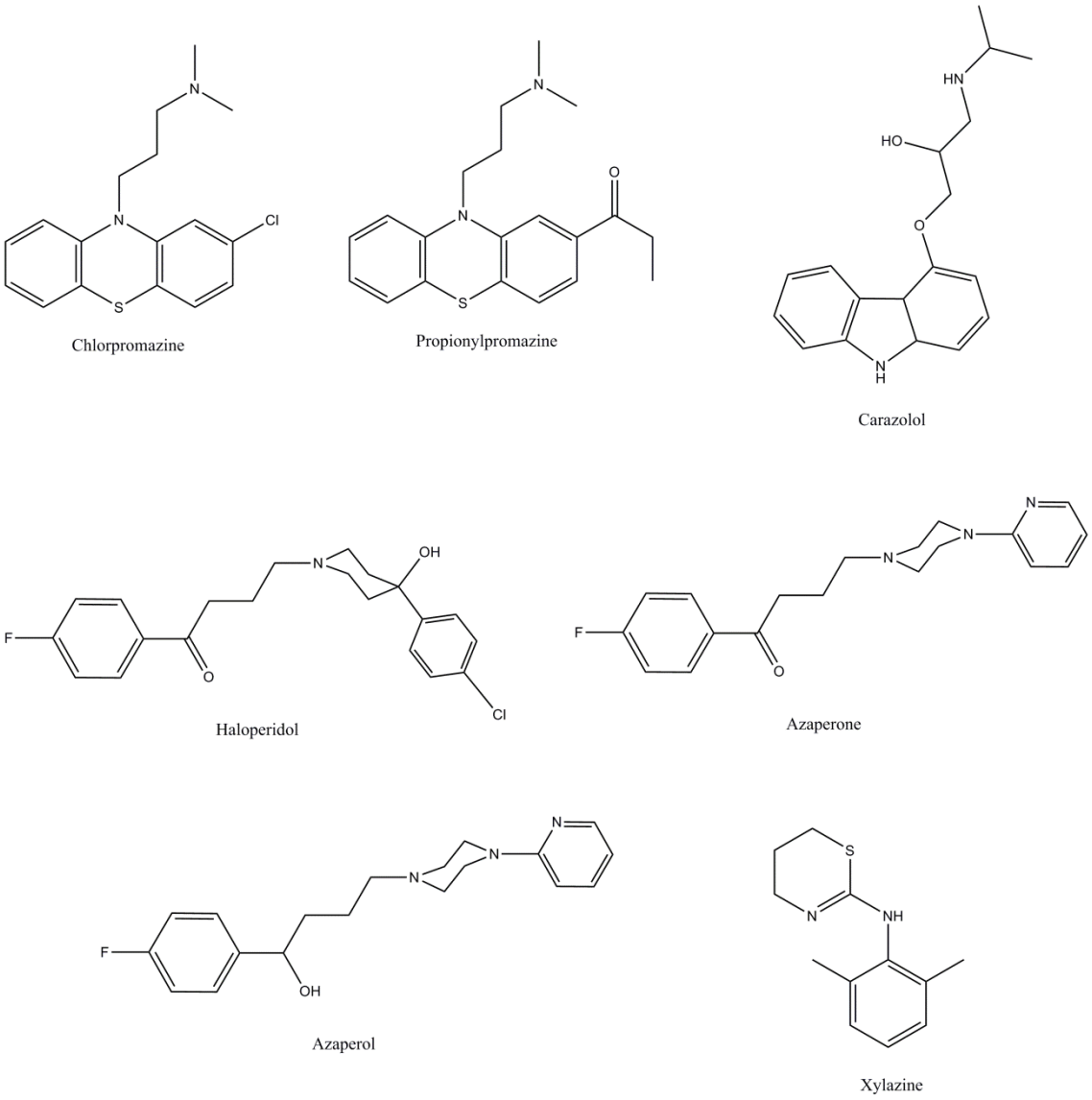


**Figure 3.3.1.** Chemical structures of the seven drugs analysed.

As the sample preparation stage includes sampling, pretreatment and solid-phase extraction (SPE) steps, it will be quite likely to be responsible of the highest errors in the determination. So, the effect of seven factors related to the sample preparation procedure on the final concentration of every drug in the sample was examined through an eight-experiment Youden design. These factors were deliberately changed between nominal and extreme conditions that represented the variability that may well occur when performing routine analyses. The results arising from the experimental plan were interpreted from several statistical methodologies in order to assess the robustness of the extraction step.

As the proposed design was saturated, an independent estimation of the experimental error as standard deviation at a previous stage of the method development was used to evaluate the significance of the factors. In addition, Lenth's and Bayes' approaches [20] have also

been applied for drawing conclusions on the robustness/ruggedness of the sample preparation stage.

Veterinary medicinal products are necessary to ensure animal health and welfare, but their administration to food-producing animals may leave residues in them. This is the case of sedatives, which are often used in animal production, especially in pigs. These are more aggressive than other farm animals and particularly sensitive to stress during handling and, more specifically, during their transport to the slaughterhouse. Not palliating all this stress will result in high premature mortality and meat of poorer quality called Pale Soft Exudative [21,22], so pigs are usually injected with sedatives a few hours before slaughtering to calm them down. Due to their potential effect on the activity of the human nervous system, residues of these drugs in foodstuffs of animal origin constitute a hazard to human health, which makes their analytical control necessary. Consequently, European legislation on this matter has been developed. As part of it, Commission Regulation (EU) No 37/2010 [23] establishes maximum residue limits (MRLs) for some allowed pharmacologically active substances depending on the animal species and on the target tissue according to the procedures in Regulation (EC) No 470/2009 [24]. Prohibited pharmacologically active substances for which a MRL cannot be established are also listed in [23]. In the specific case of the sedatives analysed in this study, the MRL for azaperone, as the sum of azaperone and azaperol, is 100 μg $kg^{-1}$, and that for carazolol is 5 μg $kg^{-1}$, both values in porcine muscle tissue, while the use of chlorpromazine in food-producing animals is not allowed.

The rules laid down in Commission Decision No 2002/657/EC [14] require confirmatory methods for organic residues such as those from veterinary drugs or contaminants to provide information on the chemical structure of the analyte. In fact, the European Union considers mass spectrometry the most suitable analytical tool for a correct identification of these compounds [25]. When full-mass spectra are not recorded, a system of identification points with different requirements for prohibited substances and for those with MRLs (listed in Group A or B, respectively, of Annex I of Directive 96/23/EC) shall be used to interpret the data.

### 3.3.3. Theory

Full factorial designs are useful for experimenting with relatively few factors. As the number of factors in a $2^k$ factorial design increases, the number of runs required for its consecution rapidly outgrows the resources of most experimenters. But it is possible to run only a fraction of the full factorial set that provides nearly as much information as in the corresponding full factorial design if it can be reasonably assumed that certain high-order

interactions are negligible. These fractional factorial designs are among the most widely used types of designs for product and process design and improvement, and industrial/business experimentation [26], being also greatly applied in screening studies such as those for evaluating the robustness/ruggedness of an analytical method.

A $2^k$ fractional factorial design containing $2^{k-p}$ runs is called a $1/2^p$ fraction of the $2^k$ design or, more simply, a $2^{k-p}$ fractional factorial design, where $k$ is the number of factors considered and $p$ the number of independent generators selected for the construction of the design, being $p < k$. More information about the construction of fractional factorial designs can be found in [26,27].

The resolution of a fractional factorial design shows how the estimated effects are confounded. In particular, resolution III fractional factorial designs $\left(2_{III}^{k-p}\right)$ are those in which no main effects are aliased with any other main effect, but they are aliased with two-factor interactions and some two-factor interactions may be aliased with each other [26]. So, as main effects can be estimated only in the absence of interaction effects, these will be considered to be negligible when estimating those.

Two-level resolution III designs can be constructed to evaluate $k = (N - 1)$ factors with $N$ runs, being $N$ a power of 2. An example of resolution III fractional factorial designs is the robustness test introduced by Youden and Steiner for analytical chemistry [13]; this kind of experimental designs was well-established in other fields of science, though, such as in agriculture since the 1920s [28].

The Youden robustness test makes use of a two-level $2_{III}^{7-4}$ fractional factorial design to study the influence of up to seven factors in eight experiments. This is the strategy recommended by different international regulatory bodies to check the robustness/ruggedness of an analytical method [1,2,13,14,29]. For the assignment of the two levels of every factor, although a symmetrical interval around every nominal operating condition (encoded as 0) between two extreme levels (encoded as −1 and +1, respectively) is often set, it is also stated in [1,2,13,14,29] that the different influential variables should be changed from their nominal values (encoded as −1) to those extreme (encoded as +1) meaning a slight change with regard to the nominal operating conditions. The existence of curvature in any of the monitored factors is not expected because the variation from level −1 to level +1 is quite small.

When seven factors are evaluated from eight runs performed following the experimental plan of the $2_{III}^{7-4}$ fractional factorial design proposed by Youden and Steiner, as no interactions are taken into account, the mathematical model posed to express the

underlying relation between the response under study *y* and the seven factors considered is, in codified variables:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \varepsilon \quad (3.3.1)$$

where $\varepsilon$ is the experimental error and the variable $x_i$, $i$ = 1,...,7, equals either −1 or +1 whether the *i*-th factor is at either nominal or extreme level, respectively.

Since the model in Eq. (3.3.1) consists of eight coefficients and only eight runs are available, the effect of every factor will thus be calculated rather than estimated as follows:

$$\beta_i = \frac{\sum y(+1) - \sum y(-1)}{4} \quad (3.3.2)$$

where $\sum y(+1)$ and $\sum y(-1)$ represent the sums of the responses where factor $x_i$ is at (+1) and (−1) level, respectively. There will be no degrees of freedom left to estimate the experimental error $\varepsilon$ either, so the evaluation of the statistical significance of either the model in Eq. (3.3.1) or its coefficients by a hypothesis test will not be possible from the design. At this point, several alternatives are available.

### *3.3.3.1. Using an external variance for the estimation of the residual*

The first choice is to have an estimation of the experimental error as standard deviation *s*, which could have been calculated either from previous tests with the analytical procedure of interest or by performing more experiments (replicates of one or more design points). Both choices will make it possible to estimate the significance of the coefficients in Eq. (3.3.1) thanks to the following hypothesis test:

$H_0$: $\beta_i = 0$ (There is not effect of the factor *i*) (3.3.3)

$H_a$: $\beta_i \neq 0$ (There is effect of the factor *i*)

The critical region (*CR*) of this hypothesis test, at a 95% confidence level, is defined as

$$CR = \left\{ \frac{\beta_i}{s_{\beta_i}} < -1.96 \text{ or } \frac{\beta_i}{s_{\beta_i}} > 1.96 \right\} \; for \; i = 1, ..., 7 \quad (3.3.4)$$

where the value of the standard deviation $s_{\beta_i}$ is calculated from Eq. (3.3.5), being *n* the number of runs in the experimental design (*n* = 8 in this work) and *s* the experimental error externally estimated.

$$s_{\beta_i} = \frac{s}{\sqrt{n}} \quad (3.3.5)$$

#### *3.3.3.2. Lenth's method*

Another alternative means the use of statistical approaches that do not need a previous estimation of the residual error. The standard procedure for the analysis of an unreplicated two-level factorial design is the normal (or half-normal) plot of the estimated factor effects. However, unreplicated designs are so widely used in practice that many formal analysis procedures have been proposed to overcome the subjectivity of the normal probability plot [26]. One of these options is Lenth's method [20,26]. This is based on the idea that, if none of the factors is active, the coefficients of the model in Eq. (3.3.1), except for the constant term $\beta_0$, will have values around 0, whereas the coefficient of an active factor will be different from 0. So, if the absolute values of those coefficients are considered, an initial estimate of the coefficient standard deviation $s_\beta$ is obtained by arranging those absolute values in increasing order and multiplying their median by 1.5. Any coefficient whose absolute value is greater than $2.5\, s_\beta$ is then removed from the list and the whole procedure is repeated until no further coefficient is rejected. Let $c$ denote the number of coefficients that remain in the end. A coefficient will be active at a 5% significance level if it lies out of the interval $\left(-t_{0.025,c/3}\ s_\beta, t_{0.025,c/3}\ s_\beta\right)$, being $t_{0.025,c/3}\ s_\beta$ the value of a Student's $t$-distribution for a probability equal to 0.025 and $c/3$ degrees of freedom.

#### *3.3.3.3. Bayesian approach*

In line with Section 3.3.3.2, a second option when no estimation of the experimental variance is possible is the Bayesian analysis of the coefficients $\beta_i$ [20]. This procedure is based on calculating the "*a posteriori*" probability that the respective factor is active. This involves that: i) only a fraction $\alpha$ (between 0.10 and 0.40) of the factors is active; ii) inactive coefficients are assumed to follow a normal distribution with mean 0 and constant variance; and iii) the coefficients of the active factors are also supposed to be normally distributed, but now with a higher variance, so that the ratio $k$ of the variance of the active factors to that of the inactive ones ranges from 5.0 to 15.0. By changing $\alpha$ and $k$, it is possible to determine both the maximum and the minimum "*a posteriori*" probabilities of every factor being active.

### 3.3.4. Material and methods

#### *3.3.4.1. Reagents and chemicals*

HPLC-grade methanol, analytical-grade ethanol, HPLC-grade acetonitrile, formic acid (98-100% purity), 32% ammonia solution and analytical-grade potassium dihydrogen phosphate were all purchased from Merck (Darmstadt, Germany). Analytical-grade

orthophosphoric acid and potassium hydroxide (both 85% purity) were supplied by Panreac (Barcelona, Spain).

A solution of phosphate buffer (0.1 M, pH 6, 1 L) was prepared by dissolving potassium dihydrogen phosphate into Milli-Q water and adjusting pH to 6 with 10 M potassium hydroxide.

Azaperol (97% minimum purity) was supplied by Dr. Ehrenstorfer GmbH (Augsburg, Germany). Xylazine hydrochloride, haloperidol, chlorpromazine hydrochloride, propionylpromazine hydrochloride, azaperone and azaperone-$d_4$ (used as the internal standard (IS) for both azaperone and azaperol) were obtained from Sigma-Aldrich (Steinheim, Germany), all of them with a 98% minimum purity. Chlorpromazine-$d_3$ (IS for both chlorpromazine and propionylpromazine) and haloperidol-$d_4$ (IS for haloperidol, carazolol and xylazine) were purchased from LGC Standards (Teddington, UK), both with a 98% minimum purity. Carazolol (Suacron®, 98% minimum purity) was a generous gift from Divasa-Farmavic, S.A. (Barcelona, Spain). Whenever needed, ultra-pure water was obtained from a Milli-Q water purification system (Millipore, Bedford, MA, USA).

#### *3.3.4.2. Standard solutions*

Stock solutions of xylazine, azaperone, propionylpromazine, chlorpromazine, haloperidol, azaperol and carazolol at 100 mg $L^{-1}$ and individual stock solutions of azaperone-$d_4$ (100 mg $L^{-1}$), chlorpromazine-$d_3$ (1 mg $L^{-1}$) and haloperidol-$d_4$ (1 mg $L^{-1}$) were prepared in ethanol and stored in amber bottles at 4 ºC. Concentrations of these analytes must be referred to their free forms, not to the hydrochloride ones.

Two working solutions were prepared by further dilution in ethanol: one containing the seven non-deuterated sedatives at the following concentrations: 100 μg $L^{-1}$ for xylazine, 2500 μg $L^{-1}$ for azaperone, 150 μg $L^{-1}$ for propionylpromazine, 100 μg $L^{-1}$ for chlorpromazine, 50 μg $L^{-1}$ for haloperidol, 2000 μg $L^{-1}$ for azaperol and 125 μg $L^{-1}$ for carazolol; and the other with the three ISs at concentrations of 2500 μg $L^{-1}$ for azaperone-$d_4$, 200 μg $L^{-1}$ for chlorpromazine-$d_3$ and 40 μg $L^{-1}$ for haloperidol-$d_4$. Both solutions were stored in amber bottles at 4 ºC.

#### *3.3.4.3. Sample pretreatment and purification procedure*

For the evaluation of the seven influential factors listed in Table 3.3.1, eight muscle samples were prepared according to the experimental plan of the Youden design shown in Table 3.3.2 (in codified variables).

**Table 3.3.1.** Experimental domain: factors and selected levels for the evaluation of the robustness of the sample preparation step.

| | **Factor (Units)** | **Level** | |
|---|---|---|---|
| | | **−1** | **+1** |
| 1 | Analyst | Analyst A | Analyst B |
| 2 | SPE performance | Manual | Automated |
| 3 | Age of eluting solution | Newly prepared | 1 week |
| 4 | Age of phosphate buffer | Newly prepared | 1 month |
| 5 | Volume of phosphate buffer (mL) | 10 | 9 |
| 6 | Volume of eluting solution (mL) | 4 | 4.5 |
| 7 | Volume of washing solution (mL) | 3 | 3.5 |

**Table 3.3.2**. Experimental matrix (in codified variables) for the robustness test and estimated concentrations of the seven analytes at each experiment.

| Run | Level | | | | | | | Concentration (μg kg$^{-1}$) | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Factor 5 | Factor 6 | Factor 7 | XYL | AZA | PROP | CHLOR | HALO | AZOL | CAR |
| 1 | −1 | +1 | −1 | −1 | +1 | −1 | +1 | 2.18 | 49.70 | 2.56 | 1.73 | 0.88 | 41.34 | 2.80 |
| 2 | +1 | +1 | −1 | +1 | −1 | −1 | −1 | 2.13 | 51.13 | 2.20 | 1.81 | 1.04 | 42.09 | 2.46 |
| 3 | −1 | −1 | −1 | +1 | +1 | +1 | −1 | 1.74 | 48.97 | 2.80 | 1.55 | 0.92 | 39.18 | 2.32 |
| 4 | +1 | +1 | +1 | +1 | +1 | +1 | +1 | 2.04 | 49.45 | 2.40 | 1.69 | 0.95 | 40.86 | 2.90 |
| 5 | −1 | +1 | +1 | −1 | −1 | +1 | −1 | 2.09 | 53.02 | 2.85 | 2.03 | 1.05 | 43.91 | 3.00 |
| 6 | +1 | −1 | −1 | −1 | −1 | +1 | +1 | 2.21 | 50.57 | 2.65 | 1.73 | 0.96 | 39.84 | 2.69 |
| 7 | −1 | −1 | +1 | +1 | −1 | −1 | +1 | 1.65 | 48.77 | 2.58 | 1.70 | 0.91 | 38.65 | 2.30 |
| 8 | +1 | −1 | +1 | −1 | +1 | −1 | −1 | 2.25 | 49.99 | 2.69 | 1.76 | 0.90 | 38.81 | 2.77 |

In addition, six matrix-matched calibration standards including a blank underwent the same pretreatment process, being the 7 variables selected as factors set at their nominal conditions (level −1). In all cases, 5 g of homogenised animal muscle was weighed in a 50-mL polypropylene tube and then fortified with the three ISs by adding a 50-µL aliquot of the working internal standard solution. Next, 100 µL of the working standard solution of the sedatives was added to all robustness samples, while five of the six matrix-matched standards were spiked, respectively, with 50, 100, 200, 300 and 400 µL of that solution. Before the preparation of both the samples and the standards, it had been checked that the animal muscle tissue used as matrix was completely free of the analytes under study. This avoided performing additional experimentation for the calculation of recovery.

10 mL (level −1) or 9 mL (level +1) of a newly prepared (level −1) or a one-month-aged (level +1) phosphate buffer solution was pipetted to every sample and the mixture was stirred in a vortex mixer for 30 s. After 10 min in an ultrasonic bath, 3 mL of 1% orthophosphoric acid was added, and every sample was vortex mixed again for 30 s and then centrifuged for 10 min at 6800 rpm at room temperature. The supernatant of every homogenate was clarified using a Millex-AP prefilter coupled to a 20 mL syringe.

The purification step by solid-phase extraction (SPE) was performed either manually (level −1) or automatedly (level +1). SPE cartridges were first conditioned with 3 mL of methanol and then equilibrated with 3 mL of Milli-Q water. After every sample filtrate had been loaded, cartridges were rinsed with 3 mL of 1 M formic acid and next with 3 mL (level −1) or 3.5 mL (level +1) of a methanol/1 M formic acid (55/45 v/v) washing solution. After drying by passing an air stream under vacuum for 10 min, elution of the analytes and the internal standards was performed with 4 mL (level −1) or 4.5 mL (level +1) of a newly prepared (level −1) or one-week-aged (level +1) eluting solution (ammonium hydroxide/ethanol 2/98 v/v). Each eluate was collected in a 15 mL amber tube, evaporated to dryness under nitrogen at 38 ºC, reconstituted with 250 µL of 30% acetonitrile and placed in an ultrasonic bath for 5 min. Every final extract was clarified through a 0.22-µm nylon filter coupled to a 1 mL syringe, and the filtrate was transferred to an insert contained into a vial for LC-MS/MS analysis.

This sample pretreatment and purification procedure had been optimized prior to carrying out the robustness study described in this work, as it has been precisely detailed throughout Section 3.2. The thorough cleanup step performed before LC-MS/MS analysis, together with the use of three deuterated internal standards and matrix-matched calibration, helped to avoid the ionization-suppression/enhancement problem as far as possible.

#### *3.3.4.4. Instrumental*

Millex-AP prefilters were obtained from Millipore Corporation (Bedford MA, USA). A Jouan C3i centrifuge (Thermo Electron Corporation, San Jose, CA, USA) was used. The vacuum manifold used for the SPE step was purchased from Waters Corporation (Milford, MA, USA). SPE cartridges were 300 mg/6 mL Varian Bond Elut LRC Certify® (Agilent Technologies, Waldbronn, Germany).

#### *3.3.4.5. LC-MS/MS analysis*

Chromatographic separation was carried out on an Agilent 1200RR series liquid chromatograph (Agilent Technologies, Waldbronn, Germany) equipped with a binary pump, a degasser, an autosampler and a column heater, in which a Gemini NX C18 column (50 mm × 2 mm I.D.; particle size, 3 μm; Phenomenex, Macclesfield, UK) was placed at 30 ºC. Two mobile phases were used in a gradient program, in which the mobile phase A consisted of acetonitrile containing 0.5% ammonium hydroxide and mobile phase B consisted of water containing 0.5% ammonium hydroxide. The initial composition was 30% A and 70% B, which was increased linearly to 85% A in the first 7 min, held at these conditions for 2 min and returned to the initial composition from minute 9 to minute 11. The column was then equilibrated in these conditions for 4 min before the next injection, being the total run time 15 min. The flow rate was set at 0.3 mL $min^{-1}$ and the injection volume was 10 μL.

Tandem mass spectrometry was performed on an Agilent 6410B triple quadrupole mass analyzer with a multimode ionization source which was programmed to operate in mixed conditions of atmospheric pressure chemical ionization (APCI) and positive ion electrospray ($ESI^+$). The source temperature was held at 250 ºC and the capillary voltage was set at 4300 V. Nitrogen was used both as nebuliser and collision gas. Analytes were detected using the multiple reaction monitoring (MRM) mode with a dwell time of 70 ms per ion. Both Q1 and Q3 quadrupoles were maintained at unit resolution. As can be seen in Table 3.3.3, for the confirmation of every drug, its molecular ion was selected as the precursor ion and two transitions thereof were monitored, while quantification was carried out with the most intense transition. Regarding the three internal standards, only one transition from its corresponding molecular ion was considered in each case.

**Table 3.3.3.** Transitions monitored for the determination of the seven analytes and the three internal standards by LC-MS/MS.

| Compound | Most intense transition | Less intense transition |
|---|---|---|
| Xylazine | 221.2 > 90.0 | 221.2 > 164.1 |
| Azaperone | 328.1 > 165.0 | 328.1 > 121.0 |
| Propionylpromazine | 341.2 > 86.1 | 341.2 > 58.1 |
| Chlorpromazine | 319.1 > 86.2 | 319.1 > 58.1 |
| Haloperidol | 376.2 > 165.1 | 376.2 > 123.2 |
| Azaperol | 330.3 > 121.1 | 330.3 > 149.2 |
| Carazolol | 299.2 > 116.1 | 299.2 > 222.1 |
| Azaperone-$d_4$ | 332.0 > 169.0 | - |
| Chlorpromazine-$d_3$ | 322.2 > 61.0 | - |
| Haloperidol-$d_4$ | 380.0 > 169.0 | - |

#### *3.3.4.6. Software*

The experimental design performed was constructed and interpreted using NEMRODW [30]. All the regression lines were estimated and validated using STATGRAPHICS Centurion XVI [31].

### 3.3.5. Results and discussion

As one of the last stages in the development of a method for the multiresidue determination of seven drugs in animal muscle by LC-MS/MS, the robustness of the sample preparation step was assessed by means of the strategy proposed by [**¡Error! Marcador no definido.**] and recommended by the legislation in force in the European Union [14]. The seven experimental factors evaluated were: 1) Analyst; 2) Way of performing the SPE stage by using either a multistation vacuum manifold or an automated extraction system; 3) Time elapsed between the preparation of the eluting solution, whose expiry was set to 1 week, and the day of analysis; 4) Time elapsed between the preparation of the phosphate buffer solution, whose expiry was set to 1 month, and the day of analysis; 5) Volume of the phosphate buffer added to the samples; 6) Volume of the ammonia/ethanol eluent in SPE; 7) Volume of the methanol/formic acid rinsing solution in SPE. Table 3.3.1 shows these seven factors as well as the two levels chosen for each of them. As commented in Section 3.3.3 and in accordance with the procedures in

[1,2,13,14,29], the nominal operating conditions of the method were selected as level −1 for every factor.

Seven experimental responses were considered, namely the concentration of the seven drugs in a muscle sample that had been fortified with 100 μL of the working standard solution. The concentration of each sedative in every muscle sample was determined from the corresponding least squares (LS) linear regression of standardized peak area (calculated as the ratio between the absolute peak area of the analyte of interest and that of its internal standard) *versus* true concentration. These seven regression models were estimated from a calibration set of six matrix-matched standards, each of them corresponding, besides a blank, to the addition of 50, 100, 200, 300 and 400 μL of the working standard solution, respectively. The calibration range of every analyte was thus 0 – 6.92 μg $kg^{-1}$ for xylazine, 0 – 202.00 μg $kg^{-1}$ for azaperone, 0 – 11.15 μg $kg^{-1}$ for propionylpromazine, 0 – 7.26 μg $kg^{-1}$ for chlorpromazine, 0 – 4.00 μg $kg^{-1}$ for haloperidol, 0 – 160.00 μg $kg^{-1}$ for azaperol, 0 – 10.03 μg $kg^{-1}$ for carazolol. The matrix-matched standards were pretreated and analysed under the nominal operating conditions of the method (see Sections 3.3.4.3 and 3.3.4.5).

The influence of the seven factors considered on the final concentration of every drug in a muscle sample was studied through a $2_{III}^{7-4}$ fractional factorial design consisting of eight experiments. As can be seen in Table 3.3.2, the resulting experimental matrix matches up with that of the Youden proposal by multiplying by (−1) every element of the matrix. The suitability and great quality of this experimental design to be performed is well-known. As its information matrix (and thus its dispersion matrix) is a diagonal matrix, the orthogonality of the design is ensured. Because of this orthogonal design structure, all main effects would be estimated independently, and the variance inflation factors (VIFs) of all the coefficients in the model of Eq. (3.3.1), which describes the relationship between every response and the seven factors monitored, were equal to 1. As this is the best value for a VIF, the estimates of those coefficients would be the most precise possible ones.

Once the whole experimentation had been carried out, the presence of both all the analytes and the internal standards in every sample was confirmed from their precursor-to-product transitions, listed in Table 3.3.3. Then, the standardized peak area of every analyte in every run was calculated from its most intense transition, and the seven LS calibration models necessary to determine the response (concentration of every drug) for every experiment of the design were estimated. The equations of these linear regression models "*Standardized peak area versus True concentration*" figure in Table 3.3.4. A LS linear regression "*Predicted concentration versus True concentration*" was also performed for every analyte in order to assess the trueness of the analytical method; the equations of these seven accuracy lines appear in the last column of Table 3.3.4.

**Table 3.3.4.** Calibration models for the quantification of the seven drugs analysed.

| Analyte | Calibration range[a] (µg kg$^{-1}$) | $A_{analyte}/A_{IS}$[b] = f($c_{true}$) ($R^2$; $s_{yx}$) | $c_{pred} = b_0 + b_1 c_{true}$ ($R^2$; $s_{yx}$) |
|---|---|---|---|
| XYL | 0 – 6.92 | $y$ = 1.04 + 4.68$x$ (98.41%; 1.77) | $y$ = 5.6 10$^{-5}$ + 1.00$x$ (98.41%; 0.38) |
| AZA | 0 – 202.00 | $y$ = 0.09 + 0.06$x$ (99.89%; 0.17) | $y$ = -1.3 10$^{-5}$ + 1.00$x$ (99.89%; 2.84) |
| PROP | 0 – 11.15 | $y$ = -0.64 + 2.37$x$ (99.73%; 0.59) | $y$ = -5.6 10$^{-6}$ + 1.00$x$ (99.73%; 0.25) |
| CHLOR | 0 – 7.26 | $y$ = -0.06 + 0.67$x$ (99.93%; 0.06) | $y$ = 1.5 10$^{-5}$ + 1.00$x$ (99.93%; 0.08) |
| HALO | 0 – 4.00 | $y$ = -2.5 10$^{-3}$ + 3.99$x$ (99.38%; 0.54) | $y$ = 2.9 10$^{-5}$ + 1.00$x$ (99.38%; 0.14) |
| AZOL | 0 – 160.00 | $y$ = 4.9 10$^{-3}$ + 0.07$x$ (99.93%; 0.13) | $y$ = -1.5 10$^{-4}$ + 1.00$x$ (99.93%; 1.86) |
| CAR | 0 – 10.03 | $y$ = 0.08 + 2.61$x$ (98.75%; 1.27) | $y$ = 7.4 10$^{-5}$ + 1.00$x$ (98.75%; 0.49) |

(a) $n$ = 6 calibration standards

(b) Standardized peak area of the analyte, calculated as *Absolute peak area of the analyte/Absolute peak area of its internal standard*

As the estimates of the intercept ($b_0$) and the slope ($b_1$) of a linear regression are correlated, the joint confidence region for both parameters of every accuracy model was estimated at $\alpha$ = 0.05. The seven resulting ellipses are plotted together in Figure 3.3.2. As can be seen there, the size of the seven ellipses was clearly different; this is due to the fact that the area of a joint confidence region is directly related to the residual variance of the model: the lower the value of the residual standard deviation (see Table 3.3.4, last column), the greater the quality of the precision on the predictions and the smaller the area within the confidence contour. The orientation of the ellipses for azaperone (in dark blue) and azaperol (in magenta), both standardized by the same IS (azaperone-$d_4$), differed from that of the confidence regions for the rest of the analytes. The reason was the use of a wider calibration range around the MRL of 100 µg kg$^{-1}$ established in [23] for the sum of azaperone and azaperol, so higher values of the standard deviation of $b_0$ were obtained in both cases. Anyway, as the centroid of all the ellipses, represented by the least squares

estimator ($b_0$,$b_1$), was statistically equal to (0,1), it could be asserted that the trueness of the analytical method was ensured jointly for the seven drugs under study at a 95% confidence level.

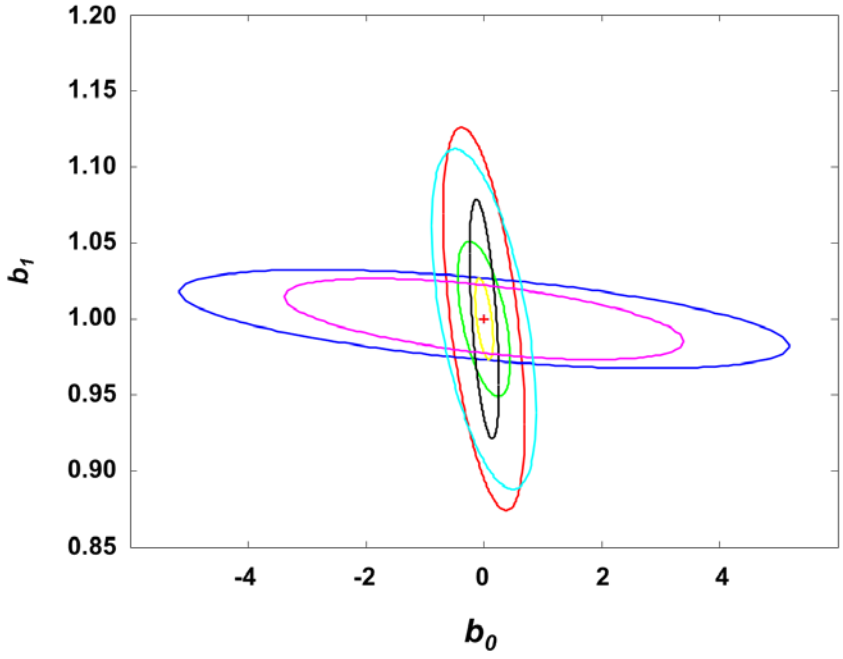


**Figure 3.3.2.** Joint confidence region for the intercept ($b_0$) and the slope ($b_1$) of the accuracy line estimated for every analyte at a 5% significance level.
**(a)** Xylazine: red line; **(b)** Azaperone: dark blue line; **(c)** Propionylpromazine: green line; **(d)** Chlorpromazine: yellow line; **(e)** Haloperidol: black line;
**(f)** Azaperol: magenta line; **(g)** Carazolol: light blue line. The target mean vector (0, 1) of all ellipses is represented by +.

The decision limit ($CC\alpha$) and the detection capability ($CC\beta$) of the analytical method for azaperone, azaperol and carazolol at its respective MRL with $\alpha = \beta = 0.05$ were estimated from the corresponding calibration model as it is laid down in [14**¡Error! Marcador no definido.**]. The values of $CC\alpha$ and $CC\beta$ achieved for these three analytes when 1 replicate is performed were, respectively, 106.5 μg $kg^{-1}$ and 112.5 μg $kg^{-1}$ for azaperone, 104.4 μg $kg^{-1}$ and 108.3 μg $kg^{-1}$ for azaperol, and 6.12 μg $kg^{-1}$ and 7.14 μg $kg^{-1}$ for carazolol.

The eight runs that made up the experimental matrix together with the values of the seven responses (concentration of every drug) are gathered in Table 3.3.2. As the experimental design conducted was saturated, the eight coefficients $\beta_i$ of each of the seven models could be exactly determined from Eq. (3.3.2); the values of $\beta_1$ to $\beta_7$, related to the effect of the factors 1 to 7 on every response, are shown in Table 3.3.5.

**Table 3.3.5.** Evaluation of the significance of the effects of the seven factors from a hypothesis test using an external variance and from Lenth's method.

| Analyte | External reproducibility variance | | | | | | | | | | Lenth's method |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***c*** **(μg kg$^{-1}$)** | ***s*** **(μg kg$^{-1}$)** | **Coefficients and statistic** | **Factor** | | | | | | | **Confidence interval** |
| | | | | **1** | **2** | **3** | **4** | **5** | **6** | **7** | |
| XYL | 1.73 | 0.36 | $\beta_i$ | 0.12 | 0.07 | -0.03 | -0.15 | 0.02 | -0.02 | -0.02 | (-0.25, 0.25) |
| | | | $\beta_i / s_{\beta_i}$ | 0.96 | 0.57 | -0.23 | -1.13 | 0.12 | -0.14 | -0.13 | |
| AZA | 50.50 | 2.02 | $\beta_i$ | 0.09 | 0.62 | 0.11 | -0.62 | -0.67 | 0.30 | -0.58 | (-3.40, 3.40) |
| | | | $\beta_i / s_{\beta_i}$ | 0.12 | 0.87 | 0.15 | -0.87 | -0.94 | 0.42 | -0.81 | |
| PROP | 2.79 | 0.42 | $\beta_i$ | -0.10 | -0.09 | 0.04 | -0.10 | 0.02 | 0.08 | -0.04 | (-0.49, 0.49) |
| | | | $\beta_i / s_{\beta_i}$ | -0.70 | -0.61 | 0.26 | -0.65 | 0.14 | 0.56 | -0.30 | |
| CHLOR | 1.81 | 0.18 | $\beta_i$ | -0.003 | 0.06 | 0.05 | -0.06 | -0.07 | 0.001 | -0.04 | (-0.27, 0.27) |
| | | | $\beta_i / s_{\beta_i}$ | -0.05 | 1.00 | 0.72 | -0.97 | -1.05 | 0.01 | -0.58 | |
| HALO | 1.00 | 0.07 | $\beta_i$ | 0.01 | 0.03 | 0.002 | 0.005 | -0.04 | 0.02 | -0.03 | (-0.11, 0.11) |
| | | | $\beta_i / s_{\beta_i}$ | 0.39 | 1.12 | 0.08 | 0.19 | -1.53 | 0.75 | -1.04 | |
| AZOL | 40.00 | 4.40 | $\beta_i$ | -0.19 | 1.47 | -0.03 | -0.39 | 0.54 | 0.36 | -0.41 | (-2.44, 2.44) |
| | | | $\beta_i / s_{\beta_i}$ | -0.12 | 0.94 | -0.02 | -0.25 | -0.34 | 0.23 | -0.27 | |
| CAR | 2.51 | 0.38 | $\beta_i$ | 0.05 | 0.14 | 0.09 | -0.16 | 0.04 | 0.07 | 0.02 | (-0.41, 0.41) |
| | | | $\beta_i / s_{\beta_i}$ | 0.36 | 1.03 | 0.66 | -1.20 | 0.30 | 0.53 | 0.13 | |

However, an estimate of the variance of the random term $\varepsilon$ in Eq. (3.3.1) could not be achieved, so the statistical significance of neither the model nor that of its coefficients could be evaluated through the corresponding hypothesis test. For this work, the reproducibility standard deviation at a concentration $c$ had been previously estimated for every analyte at an earlier stage of the validation process of the method. This standard deviation $s$, external to the experimental design, made it possible to determine whether the variations considered for the seven factors had a significant effect on every response thanks to the hypothesis test posed in Section 3.3.3.1.

For every analyte, the reproducibility standard deviation $s$ at a specific concentration $c$, together with the value of the statistic $\beta_i / s_{\beta_i}$ for the significance of each coefficient $\beta_i$, is shown in Table 3.3.5. For xylazine, propionylpromazine and haloperidol, the value of $c$ corresponded to the final concentration of every analyte in a muscle sample fortified with 100 µL of the working standard solution of these sedatives; for chlorpromazine, which is a banned substance as tranquillizer in the European Union, and azaperone, azaperol and carazolol, with values of MRL established in porcine muscle [23**¡Error! Marcador no definido.**], the value of $c$ was exactly twice the decision limit ($CC\alpha$) of the method.

Taking the hypothesis test posed in Eq. (3.3.3) into account, there was not enough experimental evidence to reject the null hypothesis ($H_0$) in any case since no statistic value lay into the corresponding critical region (see Eq. (3.3.4)). Therefore, it could be concluded that the performance of the method was not influenced by the changes in the seven variables under study and the sample preparation procedure was robust.

Other alternatives for the evaluation of potentially influential factors when an estimation of the experimental variance is not used are Lenth's method and Bayesian analysis of the coefficients $\beta_i$, both of which were also applied in this work. These two strategies, which are explained in Sections 3.3.3.2 and 3.3.3.3, respectively, are especially useful for the interpretation of saturated designs if an estimate of the residual error ε is not available.

The results of the application of Lenth's method are reflected in the last column of Table 3.3.5, where the confidence interval $\left(-t_{0.025,c/3}\ s_\beta, t_{0.025,c/3}\ s_\beta\right)$ estimated through this approach for every analyte is collected. It is clear that, in all cases, none of the seven coefficients $\beta_1$ to $\beta_7$ (also in Table 3.3.5) lay out of the corresponding interval, so all factors were non-significant and the sample preparation step was considered to be acceptably robust.

The second option applied for testing the significance of the effects was the Bayesian analysis of the coefficients of every model. As a result, the probabilities of each effect being active on every response (analyte) were calculated "*a posteriori*" for different values of $k$ and

$\alpha$. Figure 3.3.3 shows the range of "*a posteriori*" probabilities of significance plotted as a solid box between the maximum and minimum values of this probability. This range was determined for every analyte and every coefficient in Eq. (3.3.1) estimating the effect of a factor ($\beta_1$, …, $\beta_7$), so 49 ranges of "*a posteriori*" probabilities of significance are depicted in Figure.3.3.3.

The maximum was below 60% in 35 cases (coefficients), between 64% and 80% in another 5, and higher than 80% for the remaining 9 cases. However, as the "*a posteriori*" probability of significance could have any value between those minimum and maximum, no factor gave clear proof of being active on the determination of none of the seven analytes. Only the way of performing the SPE step (Factor 2) could be considered the most likely to be significant, since its "*a posteriori*" probability varied from nearly 45% to 95% in the case of azaperol, although it must be borne in mind that its real probability of significance might have any value within this range. As no effect had a minimum "*a posteriori*" probability of being significant close to 100%, no evidence of clearly influencing factors was provided, so the robustness of the sample preparation step could not be rejected. This conclusion exactly coincides with that drawn from each of the other two methodologies employed for the interpretation of the results of the Youden experimental design performed in this work (hypothesis testing and Lenth's method).

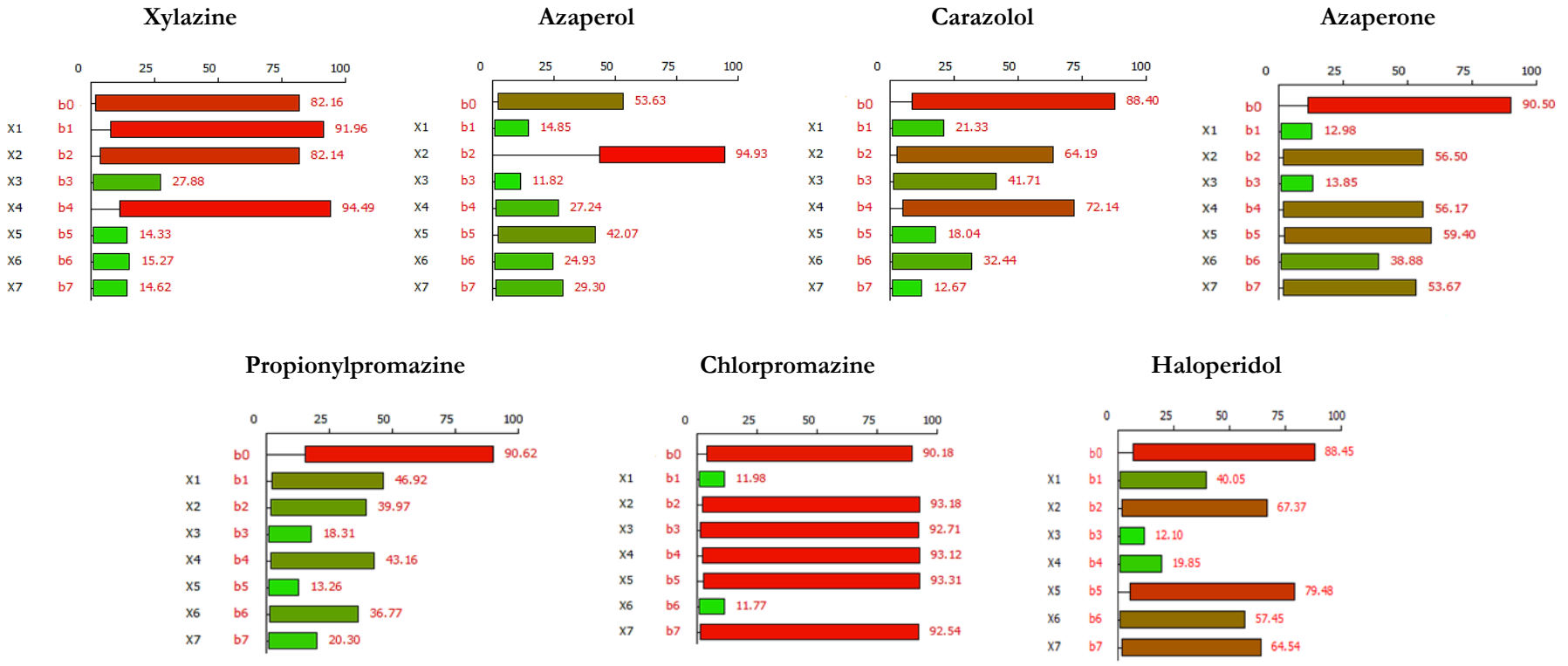


**Figure 3.3.3.** Graphic representation of the results obtained, for each analyte, from the Bayesian analysis of the statistical significance of the seven factors evaluated ($\alpha$ belongs to the interval [0.10,0.40], while $k$ does to [5.0,15.0]).

### 3.3.6. Conclusions

The robustness of the sample preparation stage for the determination of xylazine, azaperone, propionylpromazine, chlorpromazine, haloperidol, azaperol and carazolol in muscle tissue by LC-MS/MS has been verified by means of a $2_{III}^{7-4}$ fractional factorial design, following the procedure recommended by Commission Decision No 2002/657/EC. After carrying out the eight runs that make up the experimental plan, the influence of seven factors involved in the sample preparation prior to analysis (namely, analyst, way of performance of the SPE step, age of both the phosphate buffer and eluent, volumes needed of the buffer, SPE washing solution and SPE eluent) has been evaluated by means of different statistical strategies, such as hypothesis testing, Lenth's method and Bayes' analysis, the latter two allowing to draw conclusions on which factors are active even though no estimate of the residual error is available. Whatever the approach used to interpret the results of the experimental design, the sample preparation procedure under study has proved to be robust to the small variations considered, which can be expected throughout its routine performance.

### 3.3.7. Acknowledgements

The authors thank the financial support provided by Ministerio de Ciencia e Innovación (CTQ2011-26022) and MINECO (CTQ2014-53157-R). M.L. Oca and L. Rubio are particularly grateful to University of Burgos for their FPI grants.

### 3.3.8. References

[1] B. Magnusson, U. Ornemark (Eds.), *Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics*, 2nd ed., 2014. ISBN 978-91-87461-59-0. Available from http://www.eurachem.org.

[2] M. Thompson, S.L.R. Ellison, R. Wood, *Harmonized guidelines for single-laboratory validation of methods of analysis (IUPAC technical report)*, Pure and Applied Chemistry 74 (2002) 835-855.

[3] Y. Vander Heyden, A. Nijhuis, J. Smeyers-Verbeke, B.G.M. Vandeginste, D.L. Massart, *Guidance for robustness/ruggedness tests in method validation*, Journal of Pharmaceutical and Biomedical Analysis 24 (2001) 723-753.

[4] Y. Vander Heyden, M. Jimidar, E. Hund, N. Niemeijer, R. Peeters, J. Smeyers-Verbeke, D.L. Massart, J. Hoogmartens, *Determination of system suitability limits with a robustness test*, Journal of Chromatography A 845 (1999) 145-154.

[5] I. Taverniers, M. De Loose, E. Van Bockstaele, *Trends in quality in the analytical laboratory. II. Analytical method validation and quality assurance*, Trends in Analytical Chemistry 23 (2004) 535-552.

[6] E. Karageorgou, V. Samanidou, *Youden test application in robustness assays during method validation (Review)*, Journal of Chromatography A 1353 (2014) 131-139.

[7] E. Aguilera, R. Lucena, S. Cárdenas, M. Valcárcel, E. Trullols, I. Ruisánchez, *Robustness in qualitative analysis: a practical approach*, Trends in Analytical Chemistry 25 (2006) 621-627.

[8] L. Cuadros-Rodríguez, R. Romero, J.M. Bosque-Sendra, *The role of the robustness/ruggedness and inertia studies in research and development of analytical processes*, Crit. Reviews in Analytical Chemistry 35 (2005) 57-69.

[9] ICH harmonised tripartite guideline – Validation of analytical procedures: text and methodology, Q2 (R1), 2005.

[10] The United States Pharmacopeia 23/The National Formulary 18, United States Pharmacopeial Convention Inc., Rockville, 1995.

[11] D.T. Burns, K. Danzer, A. Townshend, *A tutorial discussion of the use of the terms "robust" and "rugged" and the associated characteristics of "robustness" and "ruggedness" as used in*

*descriptions of analytical procedures*, Journal of the Association of Public Analysts 37 (2009) 40-60.

[12] R.L. Plackett, J.P. Burman, *The design of optimum multifactorial experiments*, Biometrika 33 (1946) 305-325.

[13] W.J. Youden, E.H. Steiner, *Statistical Manual of the AOAC – Association of the Official Analytical Chemists, AOAC-I*, Washington, D.C., 1975.

[14] Commission Decision (EC) No 2002/657/EC of 12 August 2002, implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Official Journal of the European Communities, L 221 (2002) 8-36.

[15] I. García, L. Sarabia, M.C. Ortiz, J.M. Aldama, *Robustness of the extraction step when parallel factor analysis (PARAFAC) is used to quantify sulfonamides in kidney by high performance liquid chromatography-diode array detection (HPLC-DAD)*, Analyst 129 (2004) 766-771.

[16] V. Gamba, C. Terzano, L. Fioroni, S. Moretti, G. Dusia, R. Galarini, *Development and validation of a confirmatory method for the determination of sulphonamides in milk by liquid chromatography with diode array detection*, Analytica Chimica Acta 637 (2009) 18-23.

[17] N. Rodríguez, M.C. Ortiz, L.A. Sarabia, *Study of robustness based on n-way models in the spectrofluorimetric determination of tetracyclines in milk when quenching exists*, Analytica Chimica Acta 651 (2009) 149-158.

[18] B. Solomun Kolanović, N. Bilandžić, I. Varenina, *Validation of a multi-residue enzyme-linked immunosorbent assay for qualitative screening of corticosteroids in liver, urine and milk*, Food Additives & Contaminants: Part A 28 (2011) 1175-1186.

[19] M.R. Coleman, J.M. Rodewald, S.L. Brunelle, M. Nelson, L. Bailey, T.J. Burnett, *Determination and confirmation of nicarbazin, measured as 4,4-dinitrocarbanilide (DNC), in chicken tissues by liquid chromatography with tandem mass spectrometry: First Action 2013.07.*, Journal of AOAC INTERNATIONAL 97 (2014) 630-640.

[20] G.A. Lewis, D. Mathieu, R. Phan-Tan-Luu, *Pharmaceutical Experimental Design*, Marcel Dekker, Inc., New York, 1999.

[21] N.A. Botsoglou, D.J. Fletouris, *Drug Residues in Foods: Pharmacology, Food safety and Analysis*, Marcel Dekker, Inc. New York, 2001.

[22] H.F. De Brabander, H. Noppe, K. Verheyden, J. Vanden Bussche, K.Wille, L. Okerman, L. Vanhaecke, W. Reybroeck, S. Ooghe, S. Croubels, *Residue analysis: Future trends from a historical perspective (Review)*, Journal of Chromatography A 1216 (2009) 7964-7976.

[23] Commission Regulation (EU) No 37/2010 of 22 December 2009 on pharmacologically active substances and their classification regarding maximum residue limits in foodstuffs of animal origin, Official Journal of the European Union, L 15 (2010) 1-72.

[24] Regulation (EC) No 470/2009 of the European Parliament and of the Council of 6 May 2009 laying down Community procedures for the establishment of residue limits of pharmacologically active substances in foodstuffs of animal origin, Official Journal of the European Communities, L 152 (2009) 11-22.

[25] R. Romero González, J.L. Fernández Moreno, P. Plaza Bolaños, A. Garrido Frenich, J.L. Martínez Vidal, *Using mass spectrometry as a tool for testing for toxic substances in foods: toward food safety*, Revista Española de Salud Pública 81 (2007) 461-474.

[26] D.C. Montgomery, *Design and Analysis of Experiments*, eighth ed., John Wiley & Sons, Inc., New York, 2012.

[27] R. Cela, M. Claeys-Bruno, R. Phan-Tan-Luu, in: S.D. Brown, R. Tauler, B. Walczak (Eds.), *Comprehensive Chemometrics*, Elsevier, Amsterdam, 2009, pp. 251-300.

[28] F. Yates, *The Design and Analysis of Factorial Experiments*, Imperial Bureau of Soil Science, Harpenden, 1937.

[29] S. Bratinova, B. Raffael, C. Simoneau, *Guidelines for performance criteria and validation procedures of analytical methods used in controls of food contact materials (EUR 24105 EN – Joint Research Centre – Institute for Health and Consumer Protection)*, Office for Official Publications of the European Communities, Luxembourg, 2009.

[30] D. Mathieu, J. Nony, R. Phan-Tan-Luu, *NemrodW (Versión 2007_03)*, L.P.R.A.I. Marseille, France, 2007.

[31] *STATGRAPHICS Centurion XVI (Version 16.1.11)*, StatPoint Technologies, Inc., Herndon, VA, 2010.

# CHAPTER 4

## ENSURING FOOD SAFETY FROM PACKAGING THROUGH MULTIWAY TECHNIQUES: DEVELOPMENT OF AN ANALYTICAL METHOD TO DETERMINE BISPHENOLS AND THEIR DERIVATIVES BY GAS CHROMATOGRAPHY-MASS SPECTROMETRY AND MODELLING OF THEIR MIGRATION INTO LIPOPHILIC/ALCOHOLIC FOODSTUFFS

## 4.1. Introducción

Dada la protección que los actuales materiales y artículos de envasado ofrecen a los alimentos contra el daño físico, el deterioro químico o microbiológico y la suciedad, es imposible considerar separadamente a los primeros de los segundos y no tener en cuenta su interacción y la migración con el tiempo de sustancias presentes en los envases a los alimentos que contienen. Además, puesto que alguna de estas sustancias migrantes puede impactor negativamente en la salud humana, con el fin de garantizar la Seguridad Alimentaria, su inequívoca determinación resulta obligatoria.

Los plásticos son la principal fuente de suministro para la producción de materiales y artículos de envasado destinados a entrar en contacto con alimentos en la industria moderna. Dentro de este colectivo tan amplio, las sustancias pertenecientes a la familia química de los bisfenoles son empleadas en el marco de la industria alimenticia en la fabricación de materiales como el cloruro de polivinilo (PVC), el policarbonato (PC) y las resinas epoxi, como las utilizadas para revestir el interior de las latas que contienen alimentos sólidos y líquidos. La más que potencial presencia de unidades de bisfenol (como bisfenol F (BPF) y bisfenol A (BPA)), que quedan como monómeros libres en productos de polimerización como PC o no reaccionantes en la síntesis de resinas epoxi y otros barnices y recubrimientos, y sus derivados (como sus diglicidiléteres BFDGE y BADGE, respectivamente, principales integrantes de las resinas epoxi), es ampliamente conocida, así como su potencial peligrosidad para la salud, que comprende, entre otros riesgos, efectos disruptores endocrinos por acumulación en el organismo tras su exposición a alimentos que contienen estos compuestos tras un proceso de migración desde sus envases. A este respecto y como se detalla en el apartado 1.2.2.1.3 (*"Food contamination from food contact materials (FCMs)"*) del capítulo 1 de esta tesis, la Unión Europea ha definido un marco legislativo en constante revisión y actualización de aplicación a las sustancias químicas potencialmente presentes en alimentos (Reglamento (CE) Nº 1935/2004 del Parlamento Europeo y del Consejo, Reglamento (UE) Nº 10/2011 de la Comisión, Reglamento de Ejecución (UE) Nº 321/2011 de la Comisión, Reglamento (CE) Nº 1895/2005 de la Comisión, Reglamento (UE) 2018/213 de la Comisión), por el que se establecen restricciones, límites de migración e incluso prohibiciones de uso de estas sustancias. Por todo ello, cualquier método de análisis de estas sustancias en alimentos y bebidas ha de garantizar tanto una detección inequívoca como una cuantificación lo más exacta posible de las mismas.

A lo largo del capítulo que se expone a continuación, se describe el desarrollo y la optimización de un método analítico para determinar BPA, BPF, BADGE y BFDGE mediante un sistema PTV-GC-MS con BPA-$d_{16}$ como estándar interno teniendo en cuenta:

a. Las figuras de mérito que dan cuenta de la calidad de un proceso analítico. Estos índices pueden ser optimizados directamente a través de técnicas multivía, que manejan sin problema la naturaleza bidimensional (matricial) de la señal generada para cada muestra analizada por un equipo de GC-MS, lo que supone una ventaja considerable sobre la estrategia tradicional basada en la obtención de picos cromatográficos lo más separados e intensos posible para conseguir calibrados más exactos y sensibles. Así, tal y como se detalla en el apartado 4.2 del presente capítulo, mediante la ejecución de un diseño de experimentos, se estudió el efecto provocado por los cambios de determinados parámetros del programa de temperatura del horno del cromatógrafo sobre la calidad de la determinación de BPF, BPA, BFDGE y BADGE Los datos experimentales obtenidos fueron descompuestos a través de la técnica de análisis de factores paralelos (*Parallel Factor Analysis* (PARAFAC)), y se calcularon cinco figuras de método a partir de calibrados no lineales basados en los modelos PARAFAC estimados. A continuación y teniendo presentes las directrices y restricciones de la legislación europea en vigor durante el desarrollo de esta investigación, las figuras de mérito resultantes para los cuatro analitos de interés fueron evaluadas conjuntamente mediante un análisis de componentes principales (*Principal Component Analysis* (PCA)).

b. Las propiedades físico-químicas de todos los analitos de interés con el objetivo de evitar escoger condiciones experimentales desfavorables para alguno/s de ellos. A raíz de los bajos valores de recuperación alcanzados para BFDGE y BADGE por el método resultante del estudio recogido en el apartado 4.2, el apartado 4.3 expone en consecuencia la evaluación de la influencia del pH del medio sobre la determinación de BPF, BPA, BFDGE y BADGE, dada la distinta naturaleza química de los sustituyentes de la estructura aromática de estos cuatro compuestos, manteniéndose tanto la etapa de extracción en fase sólida como el uso de la descomposición PARAFAC como herramienta de garantía de consecución de una identificación y cuantificación correctas ya integrantes del método analítico producto del estudio reflejado en el apartado 4.2.

Por último, la parte final del apartado 4.3 explica detenidamente la monitorización de la liberación de BPA desde vasos de PC al simulante alimentario *D1* (50% de etanol (v/v)) a lo largo del tiempo a través de 7 ensayos de migración, evaluando al mismo tiempo las diferencias en la cinética de migración provocadas por la temperatura de incubación (50 y 70 ºC). Como resultado, se postuló una relación exponencialmente decreciente entre la concentración de BPA transferida desde cada vaso de PC y el tiempo de uso. En cualquier caso, independientemente de la temperatura de incubación y el tiempo transcurrido, todos

los vasos analizados cumplían con el límite específico de migración de 600 μg $L^{-1}$ de BPA fijado para esta sustancia en la legislación europea en vigor en el momento de realización de este estudio (Reglamento (UE) Nº 10/2011 de la Comisión y Reglamento de Ejecución (UE) Nº 321/2011 de la Comisión), haciéndolo igualmente si se considera el límite específico de migración actual de 50 μg $L^{-1}$ de BPA, definido en el Reglamento (UE) 2018/213 de la Comisión, que enmienda al Reglamento (UE) Nº 10/2011 de la Comisión en lo que a BPA se refiere.

## 4.2. Optimization of a GC-MS procedure that uses Parallel Factor Analysis for the determination of bisphenols and their diglycidyl ethers after migration from polycarbonate tableware[1]

### 4.2.1. Abstract

Bisphenol A (BPA), bisphenol F (BPF) and their corresponding diglycidyl ethers (BADGE and BFDGE) are simultaneously determined using a programmed-temperature vaporizer–gas chromatography-mass spectrometry (PTV-GC-MS) system. BPA is used in the production of polycarbonate (PC), whereas BADGE and BFDGE are for manufacturing epoxy resins. Several food alerts caused by the migration of this kind of substances from food contact materials have led to the harmonization of the European legislation in Commission Regulation (EU) No 10/2011, in force from 14 January 2011. In consequence, the use of BPA has been prohibited in the manufacture of plastic infant feeding bottles from 1 May 2011 and from 1 June 2011 regarding their placing on the market and importation into the European Union. Recently, the French Parliament has decreed that the presence of BPA in any food containers is banned. Similarly, the use and/or presence of BFDGE are not allowed.

In this work, a GC-MS method has been developed for the simultaneous determination of BPF, BPA, BFDGE and BADGE. For each one of the $I$ samples that are analysed, the abundance of $J$ characteristic $m/z$ ratios is recorded at $K$ times around the retention time of each peak, so a data tensor of dimension $I \times J \times K$ is obtained for every analyte. The decomposition of this tensor by means of Parallel Factor Analysis (PARAFAC) enables to: (a) identify unequivocally each analyte according to the maximum permitted tolerances for relative ion intensities, and (b) quantify each analyte, even in the presence of coeluents. This identification, based on the mass spectrum and the retention time, guarantees the specificity of the analysis. This specificity can fail if the Total Ion Chromatogram (TIC) is considered when there is poor resolution between some peaks or whether interferents coelute.

With the aim of studying the effect of shortening the time of the analysis on the quality of the determinations while maintaining the specificity of the identifications, two of the

[1] This section is published as *M.L. Oca et al. / Talanta 106 (2013) 266–280* (34 citations up to August 30, 2021 according to *Scopus*). Furthermore, and despite its post-2012 publication date, EFSA decided to include this research study in *EFSA CEF Panel (EFSA Panel on Food Contact Materials, Enzymes, Flavourings and Processing Aids), 2015. "Scientific Opinion on the risks to public health related to the presence of bisphenol A (BPA) in foodstuffs: Executive summary". EFSA Journal 2015;13(1):3978, 23 pp. doi: 10.2903/j.efsa.2015.3978* since it was the only paper back in that time that reported migration data from PC tableware. In addition, some of the findings collected in this section are contained as Case 1 (Section 4.1) in *M.C. Ortiz et al. / Chemometrics and Intelligent Laboratory Systems 200 (2020) 104003* (7 citations up to August 30, 2021 according to *Scopus*).

heating ramps in the oven temperature program are changed according to a two-level factorial design. Each analyte is identified by means of a PARAFAC decomposition of a data tensor obtained from several concentration levels, in such a way that five figures of merit are calculated for each experiment of the design. The analysis of these figures of merit for the 16 objects (4 compounds×4 heating ramps) using Principal Component Analysis (PCA) shows that the shortest temperature program should be considered, since this is the one the best figures of merit for BPA and BFDGE (both banned) are achieved with. At these conditions and with probabilities of false positive and false negative fixed at 0.05, values of detection capability (*CCβ*) between 2.65 and 4.71 μg $L^{-1}$ when acetonitrile is the injection solvent, and between 1.97 and 5.53 μg $L^{-1}$ for acetone, are obtained.

This GC-MS method has been applied to the simultaneous determination of BPF, BPA, BFDGE and BADGE in food simulant *D1* (ethanol:$H_2O$, 1:1 v/v), which had been previously in contact with PC tableware for 24 h at 70 ºC and then pretreated by a solid-phase extraction (SPE) step. The migration of BPA from the new PC containers analysed is confirmed, and values between 104.67 and 181.46 μg $L^{-1}$ (0.73 and 1.27 μg $L^{-1}$ after correction) of BPA have been estimated. None of the results obtained exceeds the specific migration limit of 600 μg $L^{-1}$ established by law for BPA in plastic food materials different from PC infant feeding bottles. Severe problems of coelution of interferents have been overcome using PARAFAC decompositions in the analysis of these food simulant samples.

### 4.2.2. Introduction

The optimization of GC-MS procedures is usually aimed to achieve the most intense and resolved peaks so as to yield determinations of the best analytical quality. Somehow, if the structure of the signals provided by the GC-MS instrumentation is considered, the figures of merit that determine the quality of an analytical procedure can be directly optimized [1] by means of multiway techniques.

Calibrations based on *n*-way data provided by modern chromatographic instruments are increasingly used in chemical analysis due to their applicability to numerous analytical techniques and their so-called "second-order advantage". This property makes these methods especially useful for quantifying and identifying analytes in complex samples where unknown interferents are present. This is of great interest in identifying and quantifying either banned substances or those for which a permitted limit has been established. Several reviews on this topic can be found in [2,3,4,5].

Among multiway techniques, parallel factor analysis (PARAFAC), which is a generalization of principal component analysis (PCA) for three-way data, has been applied in many occasions on GC-MS data. The use of PARAFAC has successfully helped to solve frequent

problems such as coelution of either compounds whose mass spectrum is quite similar to that of the analyte of interest, or of unknown substances [6]; simultaneous optimization of solid-phase microextraction and derivatization stages [1,7]; or building overall calibration models to obtain robust analytical procedures to identify and quantify several compounds in a single analysis [8].

BPA and BPF are widely used in the synthesis of resins and plastics [9,10] with multiple industrial applications, such as the production of lacquers, the inner coating of food cans and thermal paper [11]. Furthermore, BPA is used in the production of polycarbonate (PC), the chosen polymer to manufacture food containers, such as non-returnable bottles, feeding bottles, containers, etc. On the other hand, BADGE and BFDGE are intermediates in the manufacturing of epoxy resins to produce lacquers in food cans [12].

It has been reported that trace amounts of these compounds can be released into food from plastic materials and articles intended to come into contact with food [13,14,15,16,17]. In consequence, the European Union has established specifications for these substances and restrictions of use, in addition to migration limits to ensure the safety of the food contact materials. The specific migration limit applicable in particular for BPA, expressed in mg BPA per kg food, is 0.6 (Commission Regulation (EU) No 10/2011 [18]). However, the use of BPA has been prohibited in the manufacture of plastic infant feeding bottles from 1 May 2011 and from 1 June 2011 regarding their placing on the market and importation into the European Union, respectively (Commission Implementing Regulation (EU) No 321/2011). Recently, the French Parliament has decreed that the presence of BPA in any food containers is banned from 1 January 2014. Similarly, the use and/or presence of BFDGE in the production of materials and articles are also prohibited (Commission Regulation (EC) 1895/2005).

Identification and quantification of these compounds are usually carried out by means of chromatography coupled to mass spectrometry, both LC-MS and GC-MS. But other methods using LC coupled to fluorescence or electrochemical detection, as well as immunochemical methods, are also used [17,19]; however, the identification of compounds in these cases is only based on retention times, so the possibility of interference from other electrochemically active, absorbent or fluorescent substances may produce false positives; in this situation, a different method for confirmation would be necessary. The LC-MS analysis of BPA in food is exclusively performed using electrospray ionization and atmospheric pressure chemical ionization, both in negative mode [19,20,21]. Extracts of BPA in water and in different concentrations of ethanol in water [22] and in acetic acid [23] used as food simulants were also directly analysed by reverse-phase HPLC. BPA can be detected underivatized in water by GC-MS [24,25], but sensitivity can be improved using

preconcentration, liquid-liquid extraction and derivatization [26,27,28]. Derivatization leads to sharper peaks for BPA and consequently to a better separation from other analytes and coextracted matrix components, so higher sensitivity is achieved. Silylation and acetylation are the most used derivatization procedures, among other possibilities [29,30]. Large volume injection (LVI) techniques [31,32] can be an alternative to derivatization stages, which make analyses difficult and time-consuming. Chromatographic systems equipped with a programmed-temperature vaporizer (PTV) inlet permit this kind of sample injection, which improves the sensitivity of analytical procedures.

In this work, an optimization procedure applied when the PARAFAC technique is used for multiresidue analyses is described. This procedure was applied to the determination of BPA, BPF and their diglycidyl ethers (BADGE and BFDGE). With the aim of studying the effect of some operating chromatographic conditions on the quality of the determinations and identifications of several compounds, various parameters in the oven temperature program were changed according to a two-level (30 ºC $min^{-1}$ and 6 ºC $min^{-1}$) factorial design. In order to compare the results obtained at the different operating conditions, five figures of merit were calculated for each analyte from a calibration curve based on a PARAFAC decomposition. Next, PCA was used to analyse the figures of merit for the four compounds altogether.

This procedure fulfilled the specifications set by the regulatory body of the European Union for the studied migrants without any derivatization step, saving time and reagents. PARAFAC decomposition made it possible to identify unequivocally (according to the maximum permitted tolerances for relative ion abundances) and quantify each analyte. This identification, based on mass spectra and retention times, guaranteed the specificity of the analysis.

In order to test the applicability of the procedure developed, the migration of BPA from some PC cups and glasses was studied. It is established in [18] that, for the demonstration of compliance for plastic materials and articles not yet in contact with food, like those evaluated in this work, it is possible to use food simulants representing the major physico-chemical properties exhibited by food. To be precise, food simulant *D1* (50% ethanol) is assigned for milk and milk-based drinks by [18], which are some of the foodstuffs both cups and glasses are intended to come into contact with. After the migration test (24 h at 70 ºC), food simulant samples were cleaned up and preconcentrated by solid-phase extraction (SPE) and analysed following the GC-MS method optimized previously.

## 4.2.3. Methodology

### *4.2.3.1. PARAFAC model*

When, for each one of the $I$ analysed samples, the abundance of $J$ characteristic ions is recorded at $K$ times around the retention time of each peak using a GC-MS system, $I \times J \times K$ abundances ($x_{ijk}$) are obtained and arranged into three dimensions for every peak. This cube of experimental data is known as tensor $\underline{\mathbf{X}}$ with dimension $I \times J \times K$. The three ways (profiles or modes) of this tensor are named by their chemical meaning; in this case, the first one is the sample way, the second one is the spectral way, and the third one is the chromatographic way. In other words, the element $x_{ijk}$ of the tensor represents the abundance of the $j$-th ion at the $k$-th time recorded for the $i$-th sample. Specific data analysis techniques are used to analyse the structured information contained in three-way tensors.

Considering an analyte, the abundance $x_j$ of the $j$-th $m/z$ ratio acquired by a mass spectrometer is

$$x_j = \chi \varepsilon_j, \qquad j = 1, 2, ..., J \tag{4.2.1}$$

where $\varepsilon_j$ is a coefficient of proportionality between the concentration of analyte, $\chi$, and the abundance, $x_j$. $\varepsilon_j$ depends on the $j$-th $m/z$ ratio. The vector consisting of these coefficients constitutes the spectral profile (the mass spectrum).

As the mass spectrometer is coupled to a chromatograph, the fraction of analyte eluting from the chromatographic column into the mass spectrometer changes over time. So the abundance recorded at the $k$-th time becomes

$$x_{jk} = \chi \varepsilon_j \tau_k, \qquad j = 1, 2, ..., J; \;\; k = 1, 2, ..., K \tag{4.2.2}$$

where $\tau_k$ can be considered as the fraction of analyte reaching the mass spectrometer at the time $k$. The vector of all $\tau_k$'s forms the chromatographic profile (chromatographic peak).

If $F$ substances coelute, the abundance recorded at that time is the sum of the contributions of these $F$ different compounds

$$x_{jk} \cong \sum_{f=1}^{F} \chi_f \varepsilon_{jf} \tau_{kf}, \qquad j = 1, 2, ..., J; \;\; k = 1, 2, ..., K \tag{4.2.3}$$

Finally, assuming that the $f$-th analyte appears at the same retention time in every chromatographic run, the abundance of the $j$-th $m/z$ ratio in the $i$-th sample registered at the $k$-th retention time can be expressed as

$$x_{ijk} \cong \sum_{f=1}^{F} \chi_{if}\, \varepsilon_{jf}\, \tau_{kf}\,, \qquad i = 1, 2, ..., I;\;\; j = 1, 2, ..., J;\;\; k = 1, 2, ..., K \tag{4.2.4}$$

The set of all $x_{ijk}$ experimental data forms the three-way array or tensor $\underline{\mathbf{X}}$.

A PARAFAC model of rank $F$ for a data tensor $\underline{\mathbf{X}}$ can be expressed [33,34] as

$$x_{ijk} = \sum_{f=1}^{F} a_{if}\, b_{jf}\, c_{kf} + e_{ijk}\,, \qquad i = 1, 2, ..., I;\;\; j = 1, 2, ..., J;\;\; k = 1, 2, ..., K \tag{4.2.5}$$

where $e_{ijk}$ are the residual errors of the model. As it can be observed, the PARAFAC model of Eq. (4.2.5) corresponds to the physical model of Eq. (4.2.4). PARAFAC is a generalization of bilinear PCA to three-way data first proposed by Cattell [35]. Assuming that the model of Eq. (4.2.5) is adequate for the acquired data, then the least squares solution is unique. In this case, it must coincide with the true underlying physical model of Eq. (4.2.4).

If the acquired data are trilinear, i.e. spectral and chromatographic profiles for every factor should be the same in all samples (differing only in their sizes), the least squares estimates for the unknown coefficients of Eq. (4.2.5) are unique [36]. This is the uniqueness property of PARAFAC, which means that each *f*-th factor is unequivocally characterized by its own sample, spectral and chromatographic profiles. Experimental trilinear data are suitable to be analysed using PARAFAC, which makes it possible to identify compounds unequivocally by their spectral and chromatographic profiles. To measure the trilinearity degree of the experimental data tensor, the core consistency diagnostic index (CORCONDIA) has been developed [37].

Uniqueness is known in chemical analysis as the "second-order advantage". In this case, PARAFAC enables to estimate the mass spectrum and the chromatographic profile (both normalized) of an analyte even if another non-modelled substance coelutes. Therefore, it is possible to identify each analyte unequivocally using its own spectral profile, as established in some official regulations [38,39]. This property is useful for optimizing a GC-MS procedure because it makes it feasible to consider the effect on the identification and quantification of the analyte caused by changes in the experimental conditions of the procedure.

#### *4.2.3.2. Capability of detection and decision limit*

According to the ISO norm 11843-1:1997 [40], the detection capability or minimum detectable net concentration, $x_d$, is the true net concentration of the analyte in the material to be analysed which will lead, with probability $1-\beta$, to the correct conclusion that the

concentration in the analysed material is different from that in the blank. The need of evaluating the probabilities of false positive, $\alpha$, and of false negative, $\beta$, has also been recognised by IUPAC [41], and is mandatory in the European Union [39] for the identification and quantification of residues that either are toxic or come from veterinary treatments in foodstuffs. Recently, the International Vocabulary of Basic and General Terms in Metrology [42] has adopted the same criterion to define the limit of detection. The classic definition for the detection limit as $k$ times the standard deviation in the blank sample [43] only evaluates the probability of a false positive, but does not quantify the probability of a false negative explicitly; furthermore, it is independent of the slope of the calibration line. This in particular makes no sense, because the capability of detection of a method depends on the amount of the increase of the signal when the concentration of the analyte is increasing near the null concentration. A good revision of the evolution of the viewpoint of IUPAC can be found in the work by Currie [44]. IUPAC also advises against using a multiple of the standard deviation of a blank as quantification or detection limit, as indicated in Note 3, page 37 in Chapter 18 [41].

To consider the two probabilities of error correctly when evaluating the detection limit, the following one-tailed hypothesis test is posed about the presence of the analyte in the test sample:

$$\begin{aligned} &H_0\text{: } x = 0 \text{ (there is not analyte in the sample, i.e. its concentration is null).} \\ &H_1\text{: } x > 0 \text{ (there is analyte in the sample).} \end{aligned} \tag{4.2.6}$$

The probability of type I error, $\alpha$, is the probability of affirming that there is analyte when there is not actually. That is, $\alpha$ is the probability of a false positive. The probability of type II error, $\beta$, is the probability of wrongly deciding from the experimental results that there is not analyte. Therefore, $\beta$ is the probability of a false negative.

The definitions of decision limit and detection capability are generalized to decide if a sample is compliant or non-compliant with regard to a given reference value $x_0$. In this case, the decision limit is the concentration above which it can be decided with a statistical certainty of $1-\alpha$ that $x_0$ has been truly exceeded. The detection capability ($CC\beta$) means the smallest content of the substance that may be detected, identified and/or quantified in a sample with an error probability of $\beta$ (for a fixed error $\alpha$) [39]. Therefore, for this case, the decision is formally defined by the following one-tailed hypothesis test:

$$\begin{aligned} &H_0\text{: } x = x_0 \text{ (the concentration of analyte is equal to or less than } x_0\text{).} \\ &H_1\text{: } x > x_0 \text{ (the concentration of analyte is greater than } x_0\text{).} \end{aligned} \tag{4.2.7}$$

Now the probabilities $\alpha$ and $\beta$ are the probabilities of false non-compliance and false compliance, respectively. It is clear that the hypothesis test of Eq. (4.2.6) is a particular case of that in Eq. (4.2.7) when $x_0 = 0$.

The detection limit is a concentration, but analytical methods provide signals; thus, the calibration curve must be taken into account, because this is the way of transforming the detection signal into concentration. The solution in the univariate case when the linear calibration model is used appears in [45,46].

The increasing use of multivariate calibrations based on linear (Partial Least Squares, PLS) or non-linear (polynomial, neural network) models, as well as of multiway calibrations such as those based on PARAFAC, makes a generalization of both concepts mentioned above necessary. The key is to use the regression "*Calculated concentration ($c_{calc}$) versus True concentration ($c_{true}$)*"

$$y = \beta_0 + \beta_1 x + \varepsilon \tag{4.2.8}$$

The data ($x_i$, $y_i$), $i$ = 1,..., $N$, are now the true values ($x_i$) of concentration and the corresponding values ($y_i$) calculated from the calibration model. A detailed discussion about univariate calibration can be found in [47].

In the univariate case, the results are mathematically equivalent to those obtained with the regression "*Signal versus $c_{true}$*" [48]. But the calculation of the capability of detection and the limit of decision by means of the regression "*$c_{calc}$ versus $c_{true}$*" can be generalized to any calibrations performed with signals of any order. This approximation is detailed in the tutorial of [49].

## 4.2.4. Experimental

### *4.2.4.1. Reagents and standard solutions*

BPF (CAS no. 000620-92-8), BPA (CAS no. 000080-05-7), BPA-$d_{16}$ (CAS no. 096210-87-6), BFDGE (CAS no. 002095-03-6) and BADGE (CAS no. 001675-54-3), all 97% or higher purity, were obtained from Sigma–Aldrich (Steinheim, Germany). Isocratic-grade-for-HPLC acetonitrile (ACN) and both gradient-grade-for-liquid-chromatography acetone and methanol were purchased from Merck (Darmstadt, Germany), while hydrochloric acid 37% AnalaR NORMAPUR and ethanol absolute GPR RECTAPUR® were from VWR International, LLC (Radnor, Pennsylvania, USA). Deionised water was obtained by using the Milli-Q gradient A10 water purification system from Millipore (Bedford, MA, USA).

Stock solutions of 100 mg $L^{-1}$ in ACN and of 10 mg $L^{-1}$ in acetone were prepared.

Optimization samples: a set of 8 standard solutions of BPF, BPA, BFDGE and BADGE of 0, 100, 200 (replicated), 400, 600 (replicated) and 800 μg $L^{-1}$ (with 100 μg $L^{-1}$ of BPA-$d_{16}$ as internal standard) in ACN.

Validation samples: another set of 16 calibration standards of BPF, BPA, BFDGE and BADGE of 0, 2, 4, 6, 8, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55 and 60 μg $L^{-1}$ (with 8 μg $L^{-1}$ of BPA-$d_{16}$ as internal standard) in ACN.

For the analysis of simulant samples, calibration standards ranging from 0 to 90 μg $L^{-1}$ of BPF, BPA, BFDGE and BADGE (with 10 μg $L^{-1}$ of BPA-$d_{16}$ as internal standard) were prepared in acetone. BPA standards with concentrations between 100 and 650 μg $L^{-1}$ were also prepared for the quantification of the amount of this analyte migrated from the PC tableware.

#### *4.2.4.2. Migration testing*

The new PC cups (Cup 1 and Cup 2) and glasses (Glass 1 and Glass 2) were rinsed with deionised water several times before the migration test. After being filled with the simulant *D1* (ethanol:$H_2O$, 1:1 v/v) at room temperature (up to 250 mL in the case of the two cups and 200 mL for the two glasses), the containers were placed into a thermostatic water bath at 70 ºC for 24 h. This interval of time would simulate a two-year period of use if hot milk was considered to stay an average of 2 min in one of these containers once a day. At the end of the migration period, the cups and glasses were removed from the bath and allowed to reach room temperature for 12 h.

#### *4.2.4.3. Solid-phase extraction procedure*

The pH value of the samples was adjusted to 2.4 using hydrochloric acid 37% before they were pretreated by SPE. The final volume of each simulant sample was divided into 50-mL portions and kept in amber bottles, so subsamples spiked with BPF, BPA, BFDGE and BADGE at a concentration of either 200 or 400 ng $L^{-1}$ and non-spiked subsamples were available from every container. Procedural blanks as well as five simulant samples at the 200 ng $L^{-1}$ fortification level for assessing the repeatability of the method were also prepared in 50-mL glass flasks and acidified until pH 2.4.

SPE cartridges were conditioned without vacuum by adding 15 mL of methanol and, once 1 min had elapsed, equilibrated with 6 mL of Milli-Q water. After 2 min, the whole volume (50 mL) of every sample passed through the SPE cartridge at a flow rate of approximately 3 mL $min^{-1}$ by adjusting vacuum. After every sample had been loaded, cartridges were rinsed with 3 mL of a methanol-water (10:90 v/v) mixture and dried under vacuum for 5

min. Elution of the analytes was performed using 6 mL of methanol, and then SPE cartridges were left to dry under vacuum for 30 min. Each eluate was collected in a 10-mL glass tube and evaporated to dryness at 40 ºC using the miVac evaporator. Samples were redissolved in 400 µL of a solution containing 10 µg $L^{-1}$ of BPA-$d_{16}$ (internal standard), vortex-mixed for 35 s and transferred into an insert contained in a vial for GC-MS analysis.

#### *4.2.4.4. GC-MS analysis*

A volume of 10 µL was injected at a controlled speed of 3.5 µL $s^{-1}$ (after each injection the syringe was washed firstly with acetone and next with ACN at the first part of the work, while two washings with acetone were performed throughout the analyses of simulant samples). The PTV inlet was operated in the solvent vent mode. During injection, the inlet temperature was held at 50 ºC for 0.8 min, while the column head pressure was fixed at 7 kPa and the flow rate through the split vent was set at 20 mL $min^{-1}$ (at a vent time of 0.5 min, the split valve was closed). Next, the inlet temperature was ramped at 10.8 ºC $s^{-1}$ up to 280 ºC, which was held for 15 min. The split valve was re-opened 2 min after injection to purge the inlet at a vent flow of 47.4 mL $min^{-1}$. The oven temperature was maintained at 40 ºC for 2 min, and then ramped at the corresponding heating rates set by an experimental design detailed in Section 4.2.5.3 up to the end temperature of 280 ºC. The carrier gas was maintained at a constant flow rate of 1.1 mL $min^{-1}$.

Analyses were performed in the electron ionization mode at 70 eV operating either in full scan ($m/z$ ratios from 50 to 350) or in SIM mode. The SIM program used had four acquisition windows: window 1, for BPF peak, with an ion dwell time of 75 ms and 2.51 cycles $s^{-1}$; window 2, for BPA-$d_{16}$ and BPA peaks, with an ion dwell time of 40 ms and 2.58 cycles $s^{-1}$; window 3, for BFDGE peak, with an ion dwell time of 75 ms and 2.50 cycles $s^{-1}$; and window 4, for BADGE peak, with an ion dwell time of 75 ms and 2.50 cycles $s^{-1}$ (during the first stages of this work, the previous windows 3 and 4 and their respective $m/z$ ratios were combined into just one for the detection of BFDGE and BADGE, with an ion dwell time of 35 ms and 2.64 cycles $s^{-1}$). The electron multiplier was fixed at 1200 V, the ion source temperature at 230 °C, the GC-MS interface temperature at 300 °C, and the source vacuum at $10^{-5}$ Torr.

#### *4.2.4.5. Instrumental*

To control the temperature in the migration testing, a water bath equipped with an immersion thermostat Digiterm 200 (JP Selecta S.A., Barcelona, Spain) was employed. The vacuum manifold used for the SPE step was purchased from Waters Corporation (Milford, MA, USA). In order to reach the negative pressure required for the SPE, a laboratory

vacuum pump for manifolds (Sartorius AG, Goettingen, Germany) was used. SPE cartridges were Oasis® HLB 200 mg/6 cc (Waters Corporation, Milford, MA, USA). The evaporation of the solvent in the eluates from the SPE stage was performed in a miVac Modular Concentrator (GeneVac Limited, Ipswich, UK), which consisted of a miVac Duo concentrator, a SpeedTrapTM and a Quattro pump. A ZX3 vortex mixer (VELP Scientifica, Usmate (MB), Italy) was used for homogenizing samples.

Analyses were carried out on an Agilent 7890A gas chromatograph coupled to an Agilent 5975 Mass Selective Detector (MSD). The injection system consisted of a septumless head and a PTV inlet (CIS 6 from Gerstel GERSTEL GmbH & Co. KG, Germany) equipped with an empty multi-baffled deactivated glass liner. LVI was carried out using the MultiPurpose Sampler (MPS 2XL from Gerstel GERSTEL GmbH & Co. KG, Germany) with a 10-μL syringe. Analytical separations were performed on an Agilent DB HP-5MS Ultra Inert (30 m × 0.25 mm i.d., 0.25 μm film thickness) column. Helium was used as the carrier gas.

#### *4.2.4.6. Software*

MSD ChemStation E.02.01.1177 (Agilent Technologies, Inc.) and GERSTEL MAESTRO 1 (version 1.3.20.41/3.5) were used for data acquisition and processing with the help of the NIST mass spectral library [50]. A home-designed program enabled the transformation of the data exported from the GC-MS software into the appropriate format to build data tensors with MATLAB (version 7.9.0.529 (R2009b)). PCA and PARAFAC models were performed with the PLS_Toolbox 5.0 [51] for use with MATLAB. The least squares regressions were built and validated with STATGRAPHICS Centurion XVI [52]. Least median of squares regressions for the detection of outliers were carried out with PROGRESS [53]. Decision limit, $CC\alpha$, and capability of detection, $CC\beta$, were determined using the DETARCHI program [46], and $CC\alpha$ and $CC\beta$ at the maximum residue limit were estimated using NWAYDET (a program written in-house that evaluates the probabilities of false non-compliance and false compliance for *n*-way data).

### 4.2.5. Results and discussion

#### *4.2.5.1. Initial approach. Full scan mode*

In order to develop a procedure to determine BPF, BPA, BFDGE and BADGE by means of GC-MS, initial analyses were carried out. At first, the following oven temperature program was set: the initial temperature (40 ºC) was held for 2 min; then a 10 ºC $min^{-1}$ ramp to 175 ºC, which was held for 2 min, followed by a 6 ºC $min^{-1}$ ramp to 280 ºC,

temperature maintained for 3 min. The total run time was 37.0 min. The mass spectrometer detector (MSD) was operated in full scan mode to permit the identification of compounds by comparing the acquired spectra with those in the NIST mass spectral library.

The chromatogram shown in Figure 4.2.1.a was obtained from the injection of a solution containing 500 μg $L^{-1}$ of all these compounds along with the internal standard BPA-$d_{16}$ and its subsequent analysis with this temperature program.

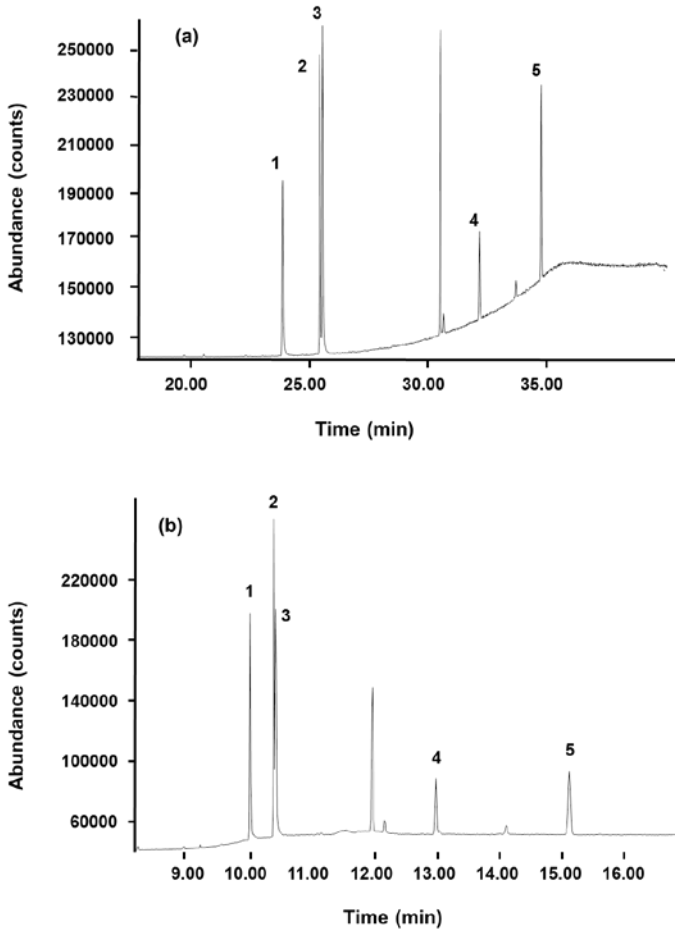


**Figure 4.2.1.** Total ion chromatograms obtained in full scan mode from the injection of 500 μg $L^{-1}$ of BPF, BPA-$d_{16}$, BPA, BFDGE and BADGE. Oven temperature programs were: **(a)** 40 ºC (for 2 min), 10 ºC $min^{-1}$ to 175 ºC (for 2 min), 6 ºC $min^{-1}$ to 280 ºC (3 min); and **(b)** 40 ºC (for 2 min), 60 ºC $min^{-1}$ to 150 ºC (2 min), 30 ºC $min^{-1}$ to 280 ºC (7 min). Peak labels: **1**, BPF; **2**, BPA-$d_{16}$; **3**, BPA; **4**, BFDGE; and **5**, BADGE.

In addition to the five expected peaks (**1** to **5** in Figure 4.2.1), other peaks appeared; one of them was identified as an isomer of BFDGE, while another as some phthalate compound, which did not interfere in the analysis of the target compounds. In this sense, the

chromatographic separation of BFDGE is harder than for the rest of the compounds, because the condensation reaction between phenol and formaldehyde, which may occur both in the *orto* and the *para* positions of the phenolic ring, results in the final formation of three isomers of BFDGE [54]. At these operating conditions, peak **5**, corresponding to BADGE, appeared almost at 35 min. The total run time was too long.

Next, the oven temperature program was modified in the following way: 40 ºC (maintained for 2 min); temperature increased at 60 ºC $min^{-1}$ to 150 ºC (for 2 min), then at 30 ºC $min^{-1}$ to 280 ºC (for 7 min). The total run time was 17.2 min. At these operating conditions, the chromatogram shown in Figure 4.2.1.b was obtained, where peaks appeared closer, as expected.

From a chromatographic point of view, the chosen temperature program should have been the last one, since it meant the shortest run. However, this is not necessarily linked to obtaining the best determinations from that temperature program, as the effect on the figures of merit of the analytes caused by changing the heating ramps in the oven program is unknown when PARAFAC is used. Finding this would be the main object of the first part of this work.

#### *4.2.5.2. Initial approach. SIM mode*

In order to achieve a higher sensitivity, the MSD no longer operated in full scan mode, but in SIM. The ions shown in Table 4.2.1 were chosen to identify the compounds from the spectra recorded previously. In the case of BFDGE, the characterization and quantification of the compound were carried out considering only the *p-p'* isomer (peak **4** in Figures 4.2.1.a and 4.2.1.b).

At the sight of the TICs in Figures 4.2.1.a and 4.2.1.b, it could be concluded that the peaks **2** and **3** were overlapped. As a consequence, as it is shown in Table 4.2.1, the diagnostic ions considered for identifying and quantifying the compound corresponding to the peak **2** (BPA-$d_{16}$) were singular and different from those chosen for BPA (peak **3**), so both peaks were not actually overlapped in the spectral mode, and a real interference was thus not reflected.

The presence of the analytes was confirmed in accordance with the criteria collected in the Annex 2 of [55], where it is set, among other requirements, that the relative retention times of an analyte shall be the same as that of the calibration standard within a margin of ± 0.5%. The relative retention time of every compound was calculated in relation to the absolute retention time of BPA-$d_{16}$, used as internal standard. In addition, the relative

abundances of the detected ions shall correspond to those of the calibration standard within permitted maximum tolerances.

To establish the tolerance intervals, a sample with 400 μg $L^{-1}$ of the five analytes was used as reference. So, after the recording of the mass spectrum of the analyte at its retention time, the percentage relative abundance of every *m/z* ratio in relation to the base peak (shown in bold in Table 4.2.1) was calculated. All the relative ion abundances calculated for the different compounds and their relative retention times were within the estimated tolerance intervals, so all of them were unequivocally identified.

**Table 4.2.1.** Acquired ions (the most intense ones are in bold), abundance, relative abundance and tolerance intervals for the reference sample of 400 μg $L^{-1}$ of the five compounds

| Analyte | *m/z* ratio | Abundance | Relative abundance (%) | Tolerance interval (%) |
|---|---|---|---|---|
| BPF | 94 | 4018 | 15.83 | (12.66 – 19.00) |
| | 107 | 23760 | 93.60 | (84.24 – 102.96) |
| | 152 | 3948 | 15.55 | (12.44 – 18.66) |
| | 183 | 5616 | 22.12 | (18.81 – 25.44) |
| | **200** | 25384 | 100 | – |
| BPA-$d_{16}$ | 125 | 14313 | 10.97 | (8.78 – 13.17) |
| | **224** | 130424 | 100 | – |
| | 242 | 26384 | 20.23 | (17.20 – 23.26) |
| | 244 | 422 | 0.32 | (0.16 – 0.49) |
| BPA | 91 | 11897 | 9.51 | (4.76 – 14.27) |
| | 107 | 5590 | 4.47 | (2.23 - 6.70) |
| | 119 | 21080 | 16.85 | (13.48 – 20.22) |
| | **213** | 125096 | 100 | – |
| | 228 | 22608 | 18.07 | (14.46 – 21.69) |
| BFDGE (first peak) | 107 | 2135 | 25.55 | (21.72 – 29.38) |
| | 181 | 5502 | 65.84 | (59.26 – 72.43) |
| | **197** | 8356 | 100 | – |
| | 279 | 344 | 4.12 | (2.06 – 6.18) |
| | 312 | 6438 | 77.05 | (69.35 – 84.76) |
| BADGE | 91 | 1313 | 2.62 | (1.31 – 3.93) |
| | 119 | 2006 | 4.00 | (2.00 – 6.00) |
| | 169 | 960 | 1.91 | (0.96 – 2.87) |
| | **325** | 50184 | 100 | – |
| | 340 | 7972 | 15.89 | (12.71 – 19.06) |

### *4.2.5.3. Optimization. Experimental design*

From the chromatograms shown in Figures 4.2.1.a and 4.2.1.b, it was not possible to know which operating conditions led to the best figures of merit of the analytical procedure for all the analytes. This is a practical problem in multiresidue analysis, since the set of operating conditions that optimize the figures of merit of an analyte does not necessarily lead to the simultaneous optimization of those of another one. This question could be approached by taking the structure of the acquired analytical signal into account: if the concentration of the compounds was varied, a data tensor would be available, which could be analysed through the multiway techniques described previously. Moreover, it was of interest to study if the PARAFAC decomposition of that data tensor succeeded in extracting the chromatographic, spectral and sample profiles related to every analyte.

The aim of the design carried out was to find the optimum oven temperature program according to three criteria: (1) the capability of PARAFAC decomposition to resolve chromatographic peaks; (2) the quality of some figures of merit (*CCα*, *CCβ*, trueness, precision and relative errors) of the determinations; and (3) the total run time, which should be as short as possible.

Bearing in mind these considerations, two heating ramps ($\phi_3$ and $\phi_4$ in Figure 4.2.2) in the oven temperature program were changed according to a two-level factorial design. Each ramp took the values shown in Table 4.2.2, which summarizes the four experiments in the experimental plan. The rest of the parameters in the oven temperature program were maintained as in the previous analyses. Figure 4.2.2 shows the parameters in the oven temperature program related to these four runs (named Experiments 1, 2, 3 and 4 throughout the text).

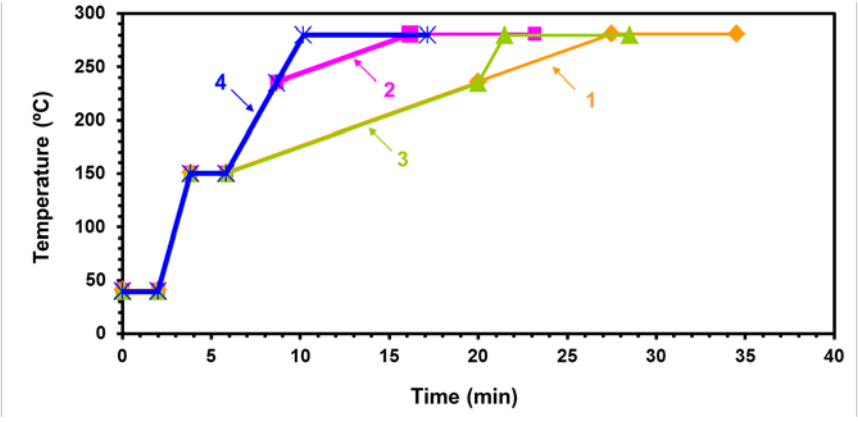


**Figure 4.2.2.** Diagram of the studied oven temperature programs. Experiment 1, orange diamonds ($\Phi_3$ = 6 ºC min$^{-1}$; $\Phi_4$ = 6 ºC min$^{-1}$); Experiment 2, pink squares ($\Phi_3$ = 30 ºC min$^{-1}$; $\Phi_4$ = 6 ºC min$^{-1}$); Experiment 3, green triangles ($\Phi_3$ = 6 ºC min$^{-1}$; $\Phi_4$ = 30 ºC min$^{-1}$); and Experiment 4, blue asterisks ($\Phi_3$ = 30 ºC min$^{-1}$; $\Phi_4$ = 30 ºC min$^{-1}$).

**Table 4.2.2.** Experimental plan and figures of merit studied by means of PCA: residual standard deviation of the polynomial curves ($s_{yx\_load}$), residual standard deviation of the accuracy line ($s_{yx\_acc}$), decision limit for a nominal value of 600 μg $L^{-1}$ and a probability of false non-compliance equal to 0.05 ($CC\alpha$), mean (*ERRORm*) and standard deviation (*ERRORsd*) of the absolute values of relative errors.

| Experimental plan | | | Object label | Figures of merit | | | | |
|---|---|---|---|---|---|---|---|---|
| Exp. | Factors | | | $s_{yx\_load}$ | $s_{yx\_acc}$ | $CC\alpha$ (μg $L^{-1}$) | *ERRORm* | *ERRORsd* |
| | $\Phi_3$ (°C $min^{-1}$) | $\Phi_4$ (°C $min^{-1}$) | | | | | | |
| 1 | 6 | 6 | BPF_1 | 2.26 $10^{-2}$ | 13.74 | 621.2 | 3.68 | 3.94 |
| | | | BPA_1 | 2.41 $10^{-2}$ | 12.91 | 619.9 | 3.57 | 3.43 |
| | | | BFDGE_1 | 2.75 $10^{-2}$ | 10.73 | 616.5 | 1.75 | 0.88 |
| | | | BADGE_1 | 5.82 $10^{-2}$ | 31.78 | 650.2 | 13.07 | 21.43 |
| 2 | 30 | 6 | BPF_2 | 3.16 $10^{-2}$ | 12.28 | 620.3 | 3.75 | 5.06 |
| | | | BPA_2 | 2.88 $10^{-2}$ | 11.55 | 619.1 | 2.55 | 2.48 |
| | | | BFDGE_2 | 3.92 $10^{-2}$ | 16.83 | 627.7 | 3.26 | 1.79 |
| | | | BADGE_2 | 3.75 $10^{-2}$ | 16.47 | 627.3 | 6.15 | 8.09 |
| 3 | 6 | 30 | BPF_3 | 3.48 $10^{-2}$ | 15.01 | 623.0 | 3.67 | 4.24 |
| | | | BPA_3 | 1.50 $10^{-2}$ | 8.56 | 613.2 | 2.48 | 2.95 |
| | | | BFDGE_3 | 2.63 $10^{-2}$ | 16.49 | 625.1 | 2.60 | 1.58 |
| | | | BADGE_3 | 7.11 $10^{-2}$ | 35.05 | 654.0 | 9.16 | 15.29 |
| 4 | 30 | 30 | BPF_4 | 3.07 $10^{-2}$ | 13.58 | 622.5 | 3.98 | 5.30 |
| | | | BPA_4 | 1.77 $10^{-2}$ | 8.74 | 614.5 | 2.67 | 3.04 |
| | | | BFDGE_4 | 1.29 $10^{-2}$ | 6.96 | 611.5 | 1.76 | 2.57 |
| | | | BADGE_4 | 4.93 $10^{-2}$ | 22.16 | 636.7 | 8.01 | 8.05 |

At the oven temperature conditions of each experiment, the chromatograms corresponding to the samples from 0 to 800 μg $L^{-1}$ detailed in Section 4.2.4.1 were acquired. At Experiments 1 and 3, a replicate of the 800-μg $L^{-1}$ sample was also analysed, so 8 or 9 analyses were carried out at every point of the design. After baseline correction, all the data matrices for every run were arranged together into a data tensor. Since for each of the $I$ analysed samples, the abundance of $J$ characteristic $m/z$ ratios was recorded at $K$ times around the retention time of each peak, a data tensor of dimension $I \times J \times K$ was obtained for every analyte, except for BPA-$d_{16}$ and BPA peaks, for which a common tensor was achieved. Therefore, four data tensors were available for each experiment.

#### *4.2.5.4. Optimization. PARAFAC decomposition*

PARAFAC decompositions of these data tensors were performed with the non-negativity constraint on two ways, as both chromatograms and spectra must always be positive. For BFDGE, the PARAFAC models had two factors for Experiments 1, 2 and 3, with CORCONDIA values higher than 99%. The first of these factors was related to the *p-p'* isomer of BFDGE; the second factor was associated with a third isomer of BFDGE. That is, the use of this multiway technique made it possible to find a compound that could not be detected because it coeluted very close to the *p-p'* isomer. On the other hand, for BPF and BADGE, one-factor models were fitted for the four experiments of the design. For BPA-$d_{16}$ and BPA, each PARAFAC model needed two factors, as expected, with CORCONDIA values of 100%. In all the cases, the loadings of the chromatographic, spectral and sample profiles were coherent.

By way of example, Figure 4.2.3 shows the loadings of the PARAFAC model for BPA-$d_{16}$ and BPA corresponding to Experiment 4. The PARAFAC decomposition found two factors related to two different compounds, which were characterized by means of their respective chromatographic (Figure 4.2.3.a), spectral (Figure 4.2.3.b) and sample (Figure 4.2.3.c) profiles. The loadings of the chromatographic profile showed two close chromatographic peaks. It was confirmed that the relative retention time of both peaks corresponded to those in the reference sample within a tolerance of ± 0.5%, as [55] states.

In addition, the loadings of the spectral profile matched the spectra obtained from the 400-μg $L^{-1}$ sample used as reference, which helped to identify both BPA-$d_{16}$ and BPA peaks. The unequivocal identification of both compounds was carried out by verifying that the relative abundances obtained from the loadings of the spectral profiles were within the estimated tolerance intervals shown in Table 4.2.1. This meant that the PARAFAC decomposition was capable of extracting and differentiating the information related to each of these two analytes. The loadings of the sample profile for BPA increased throughout every experiment, whereas those for the internal standard BPA-$d_{16}$ remained nearly constant.

The unequivocal identification of all the analysed compounds was performed in the same way for the four experiments of the design. It was also confirmed that the relative retention times of the rest of peaks corresponded to those in the reference sample (data not shown in this work).

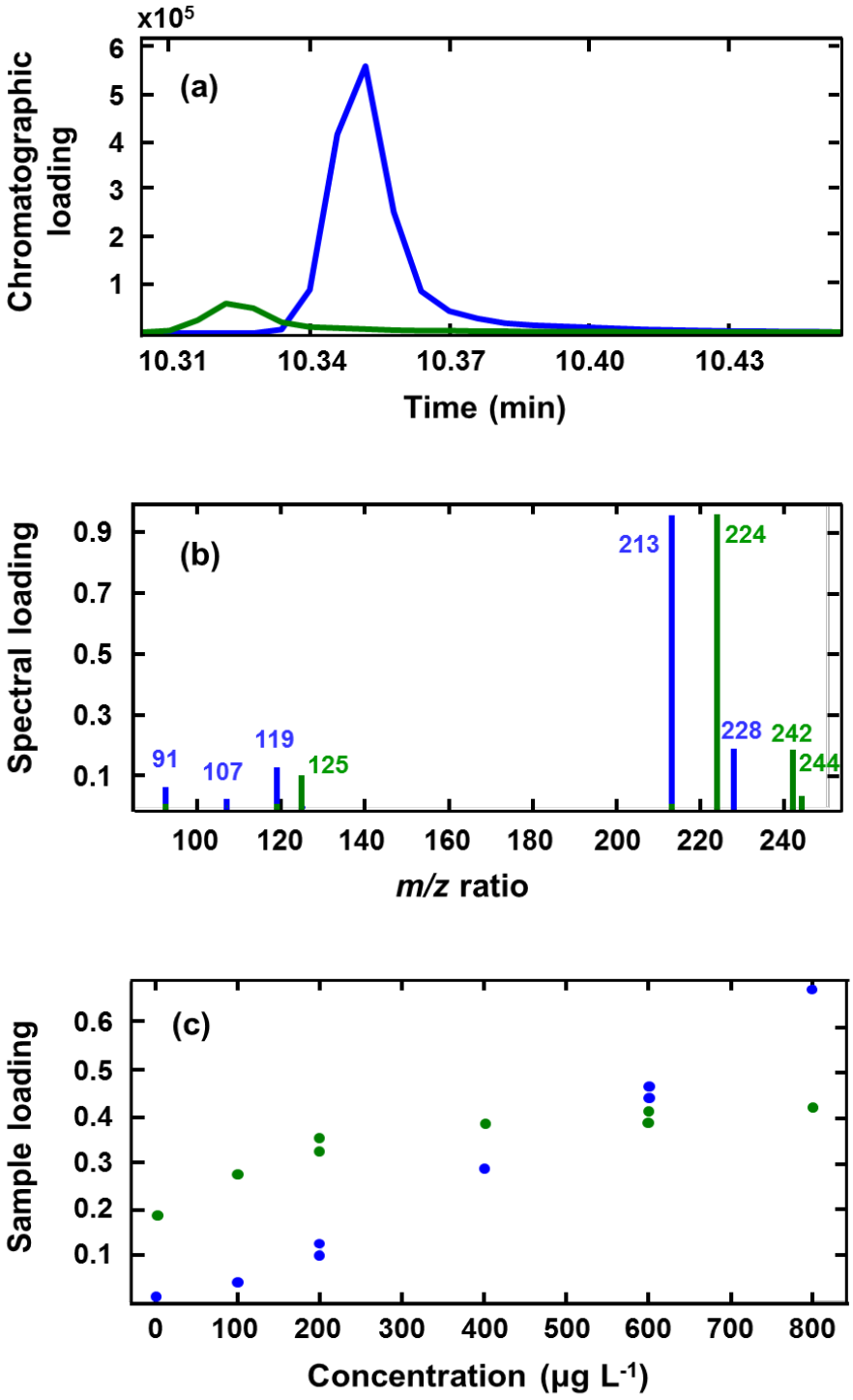


**Figure 4.2.3.** Loadings of the **(a)** chromatographic, **(b)** spectral and **(c)** sample profiles of the PARAFAC model for BPA-$d_{16}$ and BPA corresponding to Experiment 4. Factor 1 (BPA): dark blue; and Factor 2 (BPA-$d_{16}$): light green.

#### *4.2.5.5. Optimization. Figures of merit*

Once PARAFAC decomposition was carried out for every analyte, the loadings of the sample profile of the factor related to the analyte were standardized by dividing each of them by the corresponding loading of the internal standard. Next, a regression model was built for each analyte at each experiment, so 16 calibration models were performed.

A significant lack of fit was concluded in 10 out of 16 models when a simple regression "*Standardized loading (y) versus True concentration (x)*" was fitted. This drawback improved when a second-degree polynomial model was considered. In addition, in all cases, for the corresponding accuracy lines ("*Estimated concentration ($c_{calc}$) versus True concentration ($c_{true}$)*") derived from those quadratic models, slope equal to 1 and intercept equal to 0 were obtained, i.e. at a 5% signification level, there was neither proportional nor constant bias, respectively, so trueness was verified. Table 4.2.3 collects the results from this approach.

**Table 4.2.3.** Second-order polynomial calibration curves obtained from the PARAFAC decompositions ( $y = c_0 + c_1 x + c_2 x^2$ ), determination coefficient, $R^2$, and standard deviation of residuals, $s_{y/x}$. Accuracy lines ( $c_{calc} = b_0 + b_1 c_{true}$ ) and the *p*-value of the hypothesis tests with $H_0$ is stated, respectively, intercept equals 0 and slope equals 1.

| | **Calibration curve** | | | | | **Accuracy line** | |
|---|---|---|---|---|---|---|---|
| | $c_0$ | $c_1$ | $c_2$ | $R^2$ (%) | $s_{y/x}$ | $b_0$ (*p*-value) | $b_1$ (*p*-value) |
| BPF_1 | $-2.1\ 10^{-2}$ | $1.0\ 10^{-3}$ | $1.3\ 10^{-6}$ | 99.9 | 0.023 | 0.929 (0.911) | 0.998 (0.905) |
| BPA_1 | $-1.7\ 10^{-2}$ | $1.5\ 10^{-3}$ | $5.8\ 10^{-7}$ | 99.9 | 0.024 | 0.161 (0.984) | 1.000 (0.978) |
| BFDGE_1 | $4.1\ 10^{-3}$ | $7.5\ 10^{-4}$ | $1.6\ 10^{-6}$ | 99.9 | 0.028 | -2.099 (0.748) | 1.004 (0.772) |
| BADGE_1 | $5.1\ 10^{-2}$ | $1.1\ 10^{-4}$ | $2.4\ 10^{-6}$ | 99.4 | 0.058 | 17.169 (0.387) | 0.974 (0.508) |
| BPF_2 | $-2.8\ 10^{-4}$ | $7.9\ 10^{-4}$ | $1.8\ 10^{-6}$ | 99.8 | 0.032 | −1.239 (0.872) | 1.002 (0.891) |
| BPA_2 | $8.0\ 10^{-3}$ | $1.4\ 10^{-3}$ | $9.8\ 10^{-7}$ | 99.8 | 0.029 | −0.238 (0.974) | 1.001 (0.978) |
| BFDGE_2 | $2.1\ 10^{-2}$ | $8.8\ 10^{-4}$ | $1.7\ 10^{-6}$ | 99.7 | 0.039 | −3.537 (0.738) | 1.007 (0.763) |
| BADGE_2 | $7.8\ 10^{-3}$ | $1.5\ 10^{-3}$ | $7.9\ 10^{-7}$ | 99.7 | 0.038 | 0.090 (0.993) | 0.9996 (0.986) |
| BPF_3 | $5.8\ 10^{-3}$ | $6.0\ 10^{-4}$ | $1.7\ 10^{-6}$ | 99.8 | 0.035 | −2.771 (0.762) | 1.005 (0.794) |
| BPA_3 | $-2.2\ 10^{-3}$ | $1.2\ 10^{-3}$ | $8.4\ 10^{-7}$ | 99.9 | 0.015 | 0.182 (0.972) | 1.000 (0.967) |
| BFDGE_3 | $1.8\ 10^{-2}$ | $5.1\ 10^{-4}$ | $1.8\ 10^{-6}$ | 99.9 | 0.026 | −7.278 (0.476) | 1.013 (0.532) |
| BADGE_3 | $1.1\ 10^{-2}$ | $5.4\ 10^{-4}$ | $1.7\ 10^{-6}$ | 99.0 | 0.071 | −0.854 (0.968) | 1.000 (0.999) |
| BPF_4 | $-7.8\ 10^{-3}$ | $8.5\ 10^{-4}$ | $1.7\ 10^{-6}$ | 99.8 | 0.031 | −0.102 (0.990) | 1.000 (0.999) |
| BPA_4 | $-6.5\ 10^{-3}$ | $1.6\ 10^{-3}$ | $5.0\ 10^{-7}$ | 99.9 | 0.018 | 0.045 (0.993) | 1.000 (0.990) |
| BFDGE_4 | $1.3\ 10^{-2}$ | $9.6\ 10^{-4}$ | $1.4\ 10^{-6}$ | 99.9 | 0.013 | −0.975 (0.823) | 1.002 (0.833) |
| BADGE_4 | $-1.2\ 10^{-2}$ | $1.6\ 10^{-3}$ | $5.7\ 10^{-7}$ | 99.5 | 0.049 | 0.202 (0.988) | 0.999 (0.980) |

From these regression models, five figures of merit were determined: (1) Precision (i.e. the residual standard deviation of the polynomial curves, $s_{yx\_load}$); (2) Residual standard deviation of the accuracy line ($s_{yx\_accur}$); (3) Decision limit, $CC\alpha$, for a nominal value $x_0$ equal to 600 μg $L^{-1}$ with the probability of false non-compliance, $\alpha$, equal to 0.05; (4) Mean (*ERRORm*) and (5) Standard deviation (*ERRORstd*) of the absolute values of relative errors. In this way, for every studied temperature program, each analyte was identified through the PARAFAC decomposition previously detailed and the five figures of merit shown in Table 4.2.2.

#### *4.2.5.6. Optimization. Principal Component Analysis*

The procedure followed to analyse the figures of merit for the four compounds altogether was based on a PCA of the data in Table 4.2.2 previously autoscaled. The cross-validation step was carried out by means of the "*leave-one-out*" technique. The lowest value of the root mean square error of cross-validation (RMSECV) was reached with the one-component model, which explained 93.13 % of the variability of the data.

Figure 4.2.4.a shows the loadings of the PCA model performed, all of them positive. Therefore, the more similar and lower values of the four scores achieved with a temperature program, the better the analytical quality of the procedure, since better figures of merit would be then obtained. As the scores for BADGE in the Experiments 1 and 3 (see Figure 4.2.4.b) were rather high and positive, this meant that the quality of the determinations of this compound in those chromatographic conditions was poor. So they were rejected because suitable analytical quality would be desirable simultaneously for all the four analytes.

Between the other two oven temperature programs, none of them was especially appropriate for all the analytes in a simultaneous way. For the determination of BADGE, it would be clearly better to use the conditions of Experiment 2, just like for that of BPF, but the temperature program of Experiment 4 would be the best choice for the analysis of the other two analytes (Figure 4.2.4.b). The criterion used to decide which temperature program was the most suitable was based on the legislation currently in force. Thus, the oven temperature program finally selected was that of Experiment 4 ($\phi_3$ = 30 ºC $min^{-1}$ and $\phi_4$ = 30 ºC $min^{-1}$), because it made it possible to obtain the best figures of merit for BPA and BFDGE, both banned in certain cases. Furthermore, it meant the shortest total run time, which satisfied the third optimization criterion established in Section 4.2.5.3.

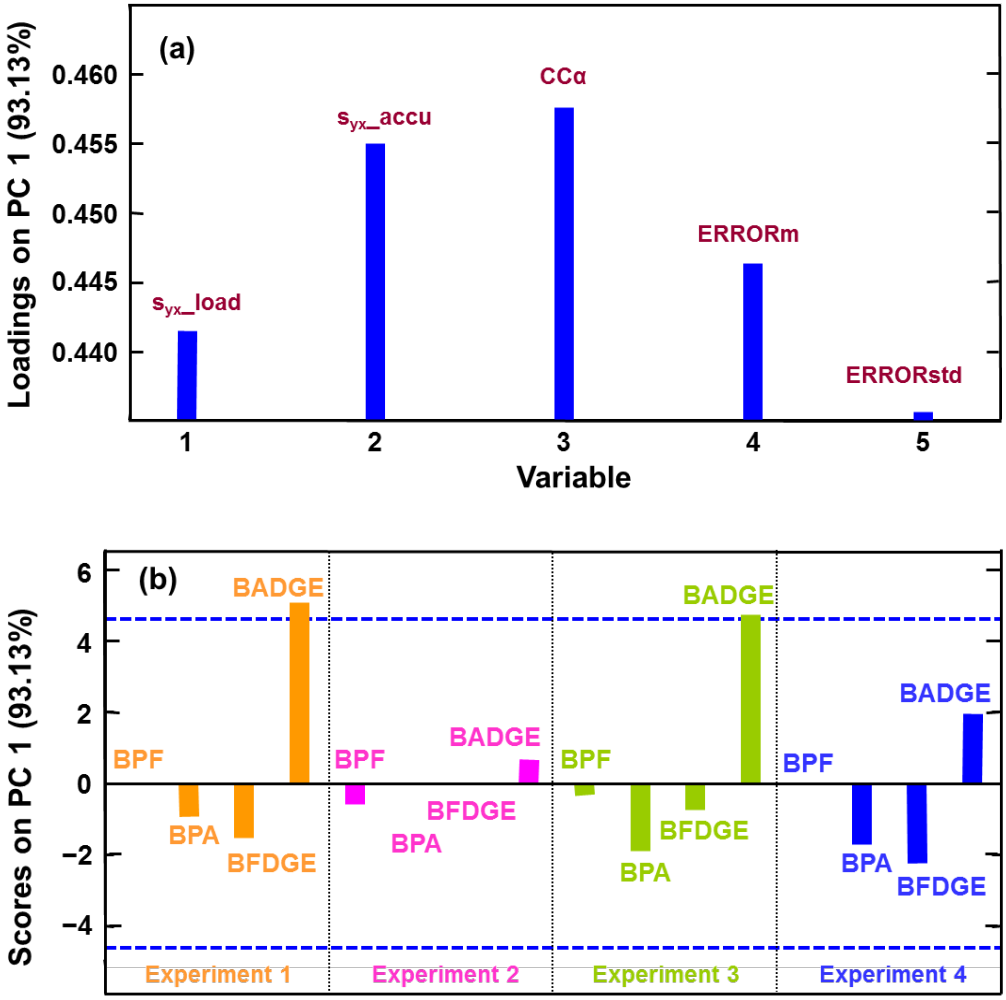


**Figure 4.2.4.** **(a)** Loadings and **(b)** scores of the one-component PCA model. *syx_load*: residual standard deviation of the polynomial curves; *syx_accur*: residual standard deviation of the accuracy line; *CCα*: decision limit for a nominal value of 600 μg $L^{-1}$ and a probability of false non-compliance equal to 0.05; *ERRORm*: mean of absolute values of relative errors; *ERRORstd*: standard deviation of absolute values of relative errors.

### *4.2.5.7. Validation of the analytical procedure*

Given that second-degree polynomial regressions were needed for fitting the calibration curves in Section 4.2.5.5, a reduction of the calibration range was performed. In this case, the concentration of the analytes ranged from 0 to 60 μg $L^{-1}$, whereas that of the internal standard remained constant at 8 μg $L^{-1}$. Sixteen calibration samples were analyzed at the chromatographic conditions of Experiment 4. The MSD was operated in SIM mode, as detailed in Section 4.2.4.4. Thus, four new data tensors were obtained, whose dimensions were 16×5×22 for BPF, 16×9×21 for BPA-$d_{16}$ and BPA (as previously, an only data tensor

was considered for both analytes simultaneously), 16×12×20 for BFDGE, and 16×12×17 for BADGE.

From these data tensors, PARAFAC decompositions were carried out. Next, a linear least squares (LS) regression model "*Standardized loading versus True concentration*" was performed for each analyte, and the corresponding accuracy line from it. Firstly, outlier data were detected using a robust regression by least median of squares (LMS) and then removed out of the calibration set if their absolute values of standardized residual and/or diagnostic resistance were higher than 2.5. The final PARAFAC models had one factor for BPF and BADGE, whereas for BPA-$d_{16}$, BPA and BFDGE two-factor models were needed, which achieved CORCONDIA values of 100%. In all cases, it was checked if the relative ratios of the loadings of the spectral profiles were within the established tolerance intervals, and it was concluded that all the analytes were unequivocally identified in this way.

The parameters of the accuracy line for each peak, along with the figures of merit related to it, are shown in Table 4.2.4; the number of outliers is shown in brackets. In all cases, the property of trueness was fulfilled. Values of capability of detection, $CC\beta$, between 2.65 and 4.71 μg $L^{-1}$ were achieved for probabilities of false positive ($\alpha$) and false negative ($\beta$) equal to 0.05. The mean of the absolute values of relative errors ranged from 5.0 to 10.9%.

**Table 4.2.4.** Parameters of the accuracy line and figures of merit determined for the four analytes. Linear calibration range: 0 - 60 μg $L^{-1}$ for BPF, BPA, BFDGE and 0 - 8 μg $L^{-1}$ for BADGE.

| | **BPF** | **BPA** | **BFDGE** | **BADGE** |
|---|---|---|---|---|
| Number of standards (outliers) | 16 (3) | 16 (2) | 16 (5) | 5 (0) |
| $b_0$ | 3.01 $10^{-9}$ | 1.42 $10^{-8}$ | 2.68 $10^{-4}$ | 0.00 |
| $b_1$ | 0.99 | 0.99 | 1.00 | 1.00 |
| $s_{y/x}$ | 0.88 | 0.70 | 1.19 | 0.62 |
| $R^2$ (%) | 99.84 | 99.89 | 99.37 | 97.20 |
| $CC\alpha$ (μg $L^{-1}$) | 1.90 | 1.50 | 2.42 | 2.48 |
| $CC\beta$ (μg $L^{-1}$) | 3.34 | 2.65 | 4.71 | 3.50 |
| $MAE^a$ (%) | 5.53 | 5.04 | 10.85 | 8.16 |

(a) Mean of the absolute values of relative errors

### *4.2.5.8. Application to food simulant samples: Migration from PC tableware*

At this stage of the work, a change of the injection solvent from ACN to acetone was decided so as to extend the useful life of the GC-MS system. On the other hand, the

capillary column was also replaced by a new one of the same type as the former. It was verified firstly that the new solvent was evaporated properly in the PTV injector during the solvent vent step and that the GC-MS method optimized using acetone as solvent was as suitable as it was for ACN. The only change in the GC-MS method involved the four SIM acquisition windows detailed in Section 4.2.4.4, which had to be modified as the absolute retention times of the analytes had changed as a result of having installed a new column.

24 simulant food samples were analysed, from which 13 samples, both spiked and non-spiked, came from the migration testing performed on some PC tableware (2 cups and 2 glasses). The 11 remaining samples were prepared from food simulant that had not been in contact with any PC containers; 5 out of these 11 samples were aimed at evaluating the repeatability of the analytical procedure, so each of them was fortified at a 200 ng $L^{-1}$ concentration level of BPF, BPA, BFDGE and BADGE. This fortification level would mean a final concentration of 25 μg $L^{-1}$ of every compound in the vial if a 100% recovery was reached. As a result of the final reconstitution step, all these 24 simulant samples contained 10 μg $L^{-1}$ of the internal standard. Table 4.2.5 collects these samples together with their origin and fortification level. 10 standards ranging from 0 to 40 μg $L^{-1}$ of the four analytes (with 10 μg $L^{-1}$ of the internal standard) were also prepared in acetone for assessing the recovery of the method.

**Table 4.2.5.** Food simulant samples analysed, indicating their fortification level and concentration of every analyte predicted from the linear regression "*Standardized sample loading versus True concentration*". The symbol "—" denotes an outlier sample.

| Container | Sample | Sample number[a] | Fortification level of BPF, BPA, BFDGE and BADGE (ng $L^{-1}$) | $c_{PRED}$ (BPF) (μg $L^{-1}$) | $c_{PRED}$ (BPA) (μg $L^{-1}$) | $c_{PRED}$ (BFDGE) (μg $L^{-1}$) | $c_{PRED}$ (BADGE) (μg $L^{-1}$) |
|---|---|---|---|---|---|---|---|
| Glass flask 1 | Repeatability 1 | 11 | 200[b] | 17.27 ± 1.70 | 39.11 ± 2.77 | 0 | 0 |
| Glass flask 2 | Repeatability 2 | 12 | 200[b] | 14.85 ± 1.71 | 39.29 ± 1.34 | 4.31 ± 2.53 | 5.90 ± 3.4 |
| Glass flask 3 | Repeatability 3 | 13 | 200[b] | 12.94 ± 1.70 | 29.36 ± 1.24 | 12.70 ± 2.42 | 8.29 ± 3.37 |
| Glass flask 4 | Repeatability 4 | 14 | 200[b] | 10.49 ± 1.72 | 27.45 ± 1.23 | 14.61 ± 2.41 | 15.90 ± 3.36 |
| Glass flask 5 | Repeatability 5 | 15 | 200[b] | 13.14 ± 1.71 | 31.89 ± 1.26 | 0 | 0 |
| Glass flask 6 | Spiked 1 – 1st r. | 16 | 200[b] | 13.23 ± 1.71 | 38.50 ± 2.78 | 4.58 ± 2.53 | 6.41 ± 3.39 |
| Glass flask 7 | Spiked 1 – 2nd r. | 17 | 200[b] | 12.93 ± 1.71 | — | 7.53 ± 2.48 | — |
| Glass flask 8 | Spiked 2 – 1st r. | 18 | 400[c] | 25.04 ± 1.74 | 56.14 ± 2.66 | 12.99 ± 2.42 | 5.67 ± 3.4 |
| Glass flask 9 | Spiked 2 – 2nd r. | 19 | 400[c] | 26.07 ± 1.74 | 62.76 ± 2.68 | 13.95 ± 2.42 | 7.96 ± 3.37 |
| Glass flask 10 | Non-spiked – 1st r. | 20 | 0 | — | — | 0 | 0 |
| Glass flask 11 | Non-spiked – 2nd r. | 21 | 0 | 0 | — | 0 | 0 |
| PC cup 1 | Spiked 1 | 22 | 200[b] | 17.53 ± 1.71 | — | 7.56 ± 2.48 | 0 |
| | Spiked 2 | 23 | 400[c] | 41.05 ± 2.99 | 248.45 ± 33.75 | 19.71 ± 2.41 | 14.00 ± 3.35 |
| | Non-spiked – 1st r. | 24 | 0 | 0 | 227.93 ± 37.96 | 0 | 0 |
| | Non-spiked – 2nd r. | 25 | 0 | 0 | 225.73 ± 37.98 | 0 | 0 |

**Table 4.2.5. (cont.)**

| Container | Sample | Sample number[a] | Fortification level of BPF, BPA, BFDGE and BADGE (ng $L^{-1}$) | $c_{PRED}$ (BPF) (μg $L^{-1}$) | $c_{PRED}$ (BPA) (μg $L^{-1}$) | $c_{PRED}$ (BFDGE) (μg $L^{-1}$) | $c_{PRED}$ (BADGE) (μg $L^{-1}$) |
|---|---|---|---|---|---|---|---|
| PC cup 2 | Spiked 2 | 26 | 400[c] | 38.70 ± 3.02 | 276.32 ± 37.61 | — | 30.37 ± 3.67 |
| | Non-spiked – 1st r. | 27 | 0 | 0 | 263.94 ± 37.65 | 0 | 0 |
| | Non-spiked – 2nd r | 28 | 0 | 0 | 160.75 ± 39.16 | 0 | 0 |
| PC glass 1 | Spiked 2 | 29 | 400[c] | 36.32 ± 3.05 | 228.10 ± 37.96 | 21.31 ± 2.41 | 11.02 ± 3.36 |
| | Non-spiked – 1st r. | 30 | 0 | 0 | 151.63 ± 39.38 | 0 | 0 |
| | Non-spiked – 2nd r. | 31 | 0 | 0 | 127.50 ± 40.04 | 0 | 0 |
| PC glass 2 | Spiked 2 | 32 | 400[c] | 37.02 ± 3.04 | 199.93 ± 38.35 | 23.56 ± 2.43 | 14.20 ± 3.35 |
| | Non-spiked – 1st r. | 33 | 0 | 0 | 168.59 ± 38.97 | 0 | 0 |
| | Non-spiked – 2nd r. | 34 | 0 | 0 | 121.76 ± 40.21 | 0 | 0 |

[a] In the tensor containing the calibration standards within the range 0 – 40 μg $L^{-1}$ (sample numbers 01 to 10).

[b] Equivalent to 25 μg $L^{-1}$ of every compound in vial if a 100% of recovery is reached.

[c] Equivalent to 50 μg $L^{-1}$ of every compound in vial if a 100% of recovery is reached.

4.2.5.8.1. Tolerance intervals for the unequivocal identification of the analytes

Before quantifying, it is necessary to establish the tolerance intervals for the unequivocal identification of every analyte, as it is laid down in [55]. Duplicated standards with concentrations of 6, 20 and 35 μg $L^{-1}$ of BPF, BPA, BFDGE and BADGE in acetone (with 10 μg $L^{-1}$ of the internal standard) were first analysed and used as references at those three different concentration levels for the confirmation of the presence of every analyte according to its mass spectrum and relative retention time. In each case and taking the permitted maximum tolerances into account, the intervals for the diagnostic ions were estimated in the same way as in Section 4.2.5.2.

A change in the most intense ion of BPF (*m/z* 107 or *m/z* 200) when either the 6-μg $L^{-1}$ standard or the 20-μg $L^{-1}$ and/or the 35-μg $L^{-1}$ standard was considered was observed (see Table 4.2.6 – second column). This incongruous situation, which may be caused by the different contribution of the background noise to the mass spectrum at different concentration levels of BPF, led to the performance of the PARAFAC decomposition of a data tensor where the six matrices (reference standards) containing the data of the absolute intensities of the five characteristic *m/z* ratios for BPF at its elution range were included (tensor dimension 6×5×11). The non-negativity constraint on both chromatographic and spectral ways was imposed, and a PARAFAC model with one factor was consequently obtained. This methodology enabled the extraction of the unique spectral profile of BPF as it was expected to be from the experimentation already conducted (the base peak was *m/z* 200), so the tolerance intervals for the relative ion abundances could then be estimated. The comparison between the tolerance intervals for BPF obtained for each standard analysed and those estimated from the PARAFAC decomposition performed are shown in Table 4.2.6. As long as the GC-MS data are trilinear, which guarantees the second-order advantage, if the common mass spectrum extracted from the PARAFAC decomposition of the data tensor of the standards at different concentrations is considered, a subjective choice of the reference mass spectrum will be avoided. If this strategy had not been followed, using as reference the mass spectrum obtained from one standard or another would have led to a high number of false negatives.

**Table 4.2.6.** Tolerance intervals for the five diagnostic ions for BPF estimated from the mass spectrum recorded at the retention time of every standard and from the PARAFAC decomposition of the tensor including the six standards at three different concentrations. The base peak in each case is in bold.

| | ***m/z*** **ratio** | **Abundance** | **Relative abundance (%)** | **Tolerance interval (%)** |
|---|---|---|---|---|
| 6-μg $L^{-1}$ standard (1st replicate) | 94 | 796 | 14.34 | (11.47 – 17.21) |
| | 107 | 4820 | 86.85 | (78.16 – 95.53) |
| | 152 | 973 | 17.53 | (14.03 – 21.04) |
| | 183 | 1553 | 27.98 | (23.79 – 32.18) |
| | **200** | 5550 | **100** | — |
| 6-μg $L^{-1}$ standard (2nd replicate) | 94 | 772 | 17.11 | (13.69 – 20.53) |
| | **107** | 4512 | **100** | — |
| | 152 | 707 | 15.67 | (12.54 – 18.80) |
| | 183 | 1185 | 26.26 | (22.32 – 30.20) |
| | 200 | 4357 | 96.56 | (86.91 – 106.22) |
| 20-μg $L^{-1}$ standard (1st replicate) | 94 | 5848 | 17.03 | (13.62 – 20.43) |
| | **107** | 34344 | **100** | — |
| | 152 | 5560 | 16.19 | (12.95 – 19.43) |
| | 183 | 7089 | 20.64 | (17.55 – 23.74) |
| | 200 | 30072 | 87.56 | (78.81 – 96.32) |
| 20-μg $L^{-1}$ standard (2nd replicate) | 94 | 4423 | 12.63 | (10.10 – 15.15) |
| | 107 | 27080 | 77.32 | (69.59 – 85.05) |
| | 152 | 4679 | 13.36 | (10.69 – 16.03) |
| | 183 | 7202 | 20.56 | (17.48 – 23.65) |
| | **200** | 35024 | **100** | — |
| 35-μg $L^{-1}$ standard (1st replicate) | 94 | 10158 | 17.51 | (14.01 – 21.01) |
| | **107** | 58016 | **100** | — |
| | 152 | 9418 | 16.23 | (12.99 – 19.48) |
| | 183 | 11474 | 19.78 | (15.82 – 23.73) |
| | 200 | 49304 | 84.98 | (76.49 – 93.48) |
| 35-μg $L^{-1}$ standard (2nd replicate) | 94 | 14991 | 16.79 | (13.43 – 20.15) |
| | 107 | 74368 | 83.29 | (74.96 – 91.62) |
| | 152 | 13622 | 15.26 | (12.21 – 18.31) |
| | 183 | 17856 | 20.00 | (16.00 – 24.00) |
| | **200** | 89288 | **100** | — |

**Table 4.2.6. (cont.)**

| | ***m/z* ratio** | **Abundance** | **Relative abundance (%)** | **Tolerance interval (%)** |
|---|---|---|---|---|
| PARAFAC decomposition (6×5×11 tensor)[a] | 94 | 0.111 | 15.12 | (12.09 – 18.14) |
| | 107 | 0.643 | 87.75 | (78.97 – 96.52) |
| | 152 | 0.111 | 15.12 | (12.10 – 18.15) |
| | 183 | 0.157 | 21.36 | (18.16 – 24.57) |
| | **200** | 0.733 | **100** | — |

[a] For the determination of the tolerance intervals, the spectral loadings were used.

The same PARAFAC strategy was followed in the case of the other four compounds. In consequence, a two-factor model was fitted from the decomposition of the common tensor for BPA-$d_{16}$ and BPA (dimension 6×9×15), with a CORCONDIA value of 100%. For BFDGE, the best PARAFAC model estimated from the corresponding tensor, with dimension 6×5×17, turned out to be the three-factor one, where the chromatographic, spectral and sample profiles showed all complete coherence, being the CORCONDIA index equal to 90%; a PARAFAC model with one factor was determined from the BADGE tensor (dimension 6×5×17). The resultant tolerance intervals for the diagnostic ions of BPA-$d_{16}$, BPA, BFDGE and BADGE from the corresponding PARAFAC decomposition are gathered in Table 4.2.7. On the other hand, the tolerance intervals for the relative retention time of every analyte were also estimated in order to ensure its unequivocal identification as it is established in the European legislation [39,55] (data not shown in this work).

**Table 4.2.7.** Tolerance intervals for the diagnostic ions of BPA-$d_{16}$, BPA, BFDGE and BADGE estimated from the PARAFAC decomposition of the corresponding tensor including the six standards at three different concentrations. The base peak in each case is in bold.

| **Analyte** | ***m/z* ratio** | **Abundance** | **Relative abundance (%)** | **Tolerance interval (%)** |
|---|---|---|---|---|
| BPA-$d_{16}$ (6×9×15 tensor) | 125 | $2.03 \cdot 10^{-1}$ | 21.04 | (17.89 – 24.20) |
| | **224** | $9.63 \cdot 10^{-1}$ | **100** | — |
| | 242 | $1.75 \cdot 10^{-1}$ | 18.15 | (14.52 – 21.78) |
| | 244 | $2.71 \cdot 10^{-3}$ | 0.28 | (0.14 – 0.42) |
| BPA (6×9×15 tensor) | 91 | $8.27 \cdot 10^{-2}$ | 8.56 | (4.28 – 12.84) |
| | 107 | $3.90 \cdot 10^{-2}$ | 4.04 | (2.02 – 6.06) |
| | 119 | $1.48 \cdot 10^{-1}$ | 15.31 | (12.25 – 18.37) |
| | **213** | $9.66 \cdot 10^{-1}$ | **100** | — |
| | 228 | $1.90 \cdot 10^{-1}$ | 19.69 | (15.75 – 23.62) |

**Table 4.2.7. (cont.)**

| Analyte | *m/z* ratio | Abundance | Relative abundance (%) | Tolerance interval (%) |
|---|---|---|---|---|
| BFDGE (6×5×17 tensor) | 107 | $2.76 \cdot 10^{-1}$ | 36.59 | (31.10 – 42.08) |
| | 181 | $4.54 \cdot 10^{-1}$ | 60.21 | (54.19 – 66.23) |
| | **197** | $7.54 \cdot 10^{-1}$ | **100** | — |
| | 279 | $1.60 \cdot 10^{-2}$ | 2.12 | (1.06 – 3.18) |
| | 312 | $3.87 \cdot 10^{-1}$ | 51.40 | (46.26 – 56.54) |
| BADGE (6×5×17 tensor) | 91 | $3.51 \cdot 10^{-2}$ | 3.56 | (1.78 – 5.34) |
| | 119 | $5.07 \cdot 10^{-2}$ | 5.14 | (2.57 – 7.70) |
| | 169 | $2.35 \cdot 10^{-2}$ | 2.38 | (1.19 – 3.57) |
| | **325** | $9.86 \cdot 10^{-1}$ | **100** | — |
| | 340 | $1.51 \cdot 10^{-1}$ | 15.25 | (12.20 – 18.30) |

4.2.5.8.2. Quantification and identification in the food simulant samples

Next, PARAFAC decompositions of the four data tensors that comprised in each case the data matrices from both the 10 calibration standards and the 24 simulant food samples were performed with the chromatographic and spectral ways non-negativity-constrained. The dimensions of these four tensors were, respectively, (34×5×18) for BPF; (34×9×23) for the common tensor for BPA-$d_{16}$ and BPA; (34×5×30) for BFDGE; and (34×5×25) for BADGE.

For BPF, a three-factor PARAFAC model (CORCONDIA value of 98%) was achieved after the removal of Sample 20, which was considered as an outlier since its $Q$ and $T^2$ indices exceeded their threshold values at the 99% confidence level. The loadings of the chromatographic, spectral and sample profiles of Factor 2 (in green dots in Figure 4.2.5) were coherent with BPF; this enabled the unequivocal identification of this analyte. The other two factors were linked with unidentified non-target compounds derived from the simulant samples, since, as it is depicted in Figure 4.2.5, they were only present in some of them.

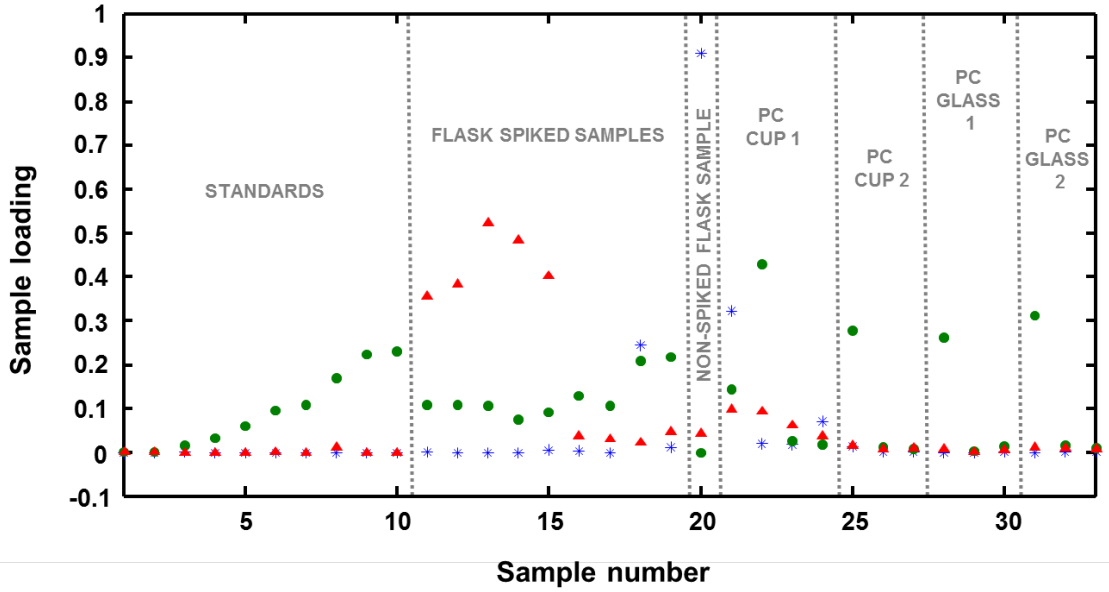


**Figure 4.2.5.** Sample-mode loadings for the three factors resultant from the PARAFAC decomposition of the BPF tensor including standards (from 0 to 40 μg $L^{-1}$) and the samples in Table 4.2.5. Factor 1, in blue asterisks; Factor 2 (BPF), in green dots; Factor 3, in red triangles.

In the case of the PARAFAC decomposition of the common tensor for BPA-$d_{16}$ and BPA, the factor corresponding to the internal standard was not at first extracted, quite the opposite of what happened with BPA. This setback was put down to a numeric effect because of the significantly larger chromatographic and sample loadings of the factor that matched BPA. As the characteristic $m/z$ ratios for the detection of BPA-$d_{16}$ were singular and different from those chosen for BPA, this problem could be solved by dividing the original tensor including BPA-$d_{16}$ and BPA data into two new tensors: the BPA-$d_{16}$ one, which contained the first 11 elution times of the former tensor (dimension 34×9×11), and the other for BPA, with dimension 34×9×12.

On the one hand, the PARAFAC decomposition of the first new tensor produced a 4-factor model (CORCONDIA value of 87%) without having detected any outliers. The loadings of the chromatographic, spectral and sample profiles of Factor 2 matched BPA-$d_{16}$, while those of Factor 1 did BPA, since this analyte coeluted in the range of times used for constructing the BPA-$d_{16}$ tensor. The other two factors could not be clearly identified and, as in the previous case, were regarded as interferents only encountered in some simulant samples. From that PARAFAC model estimated, the presence of the internal standard was unequivocally confirmed. It must be noticed that this unequivocal identification together with the quantification of this compound could never have been achieved from the TIC of most simulant samples analysed, because in that case those two

non-target compounds, especially the one associated with Factor 3, concealed the presence of BPA-$d_{16}$ to such an extent that the mass spectrum extracted from the TIC at the retention time of this compound did not match what was expected for it at all. This is illustrated in Figure 4.2.6, where Figure 4.2.6.a collects the chromatographic profile resultant from the PARAFAC decomposition of the BPA-$d_{16}$ tensor into 4 factors, and Figure 4.2.6.b represents the spectral profile of Factor 2, identified as BPA-$d_{16}$; these two graphics contrast with Figures 4.2.6.c and 4.2.6.d, being the former one the TIC obtained for one of the samples analysed in the same range of time as Figure 4.2.6.a, while the latter reflects the mass spectrum registered at 10.508 min, the retention time of BPA-$d_{16}$. The mass spectrum in Figure 4.2.6.d is clearly different from that of BPA-$d_{16}$, which appears in Figure 4.2.6.b and is detailed in Table 4.2.7.

On the other hand, BPA was both unequivocally identified and quantified from a 3-factor model resulting from the PARAFAC decomposition of the second new tensor. The CORCONDIA index equalled 100% and no outliers were detected. The first factor (in blue dots in Figures 4.2.7.a and 4.2.7.b) was coherent with BPA, while the other two remained unknown again.

When the PARAFAC decomposition of the BFDGE tensor was performed, a numeric effect was again detected, and the factor for this analyte could not be extracted. After the division of the original tensor to obtain a new one (with dimension 34×5×18) where the greatest factor (an unidentified interferent) in the former decomposition was only partly included, a 3-factor model was estimated from the PARAFAC decomposition of this new BFDGE tensor (CORCONDIA value of 99%). One outlier object, Sample 26, was removed. It was verified that the loadings of the three profiles of Factor 3 matched *p-p'* BFDGE, so this analyte was unequivocally identified.

In the case of BADGE, the PARAFAC decomposition of its corresponding tensor yielded a 2-factor model (CORCONDIA index of 91%). Sample 17 was rejected as an outlier. Factor 1 was coherent with the chromatographic, spectral and sample profiles expected for BADGE, and its correct identification was ensured from it.

After the decomposition of every data tensor into the appropriate number of factors and the identification of the factor related to each compound, the sample-mode loadings of every analyte were standardized with the sample-mode loadings of the internal standard. In each case, a linear LS regression between the standardized sample-mode loadings and the concentration of the corresponding analyte in the calibration standards was performed for the quantification stage. Previously, a LMS regression was carried out to detect possible outlier data. A new LS regression "*Standardized sample-mode loading versus True concentration*" was then estimated and validated.

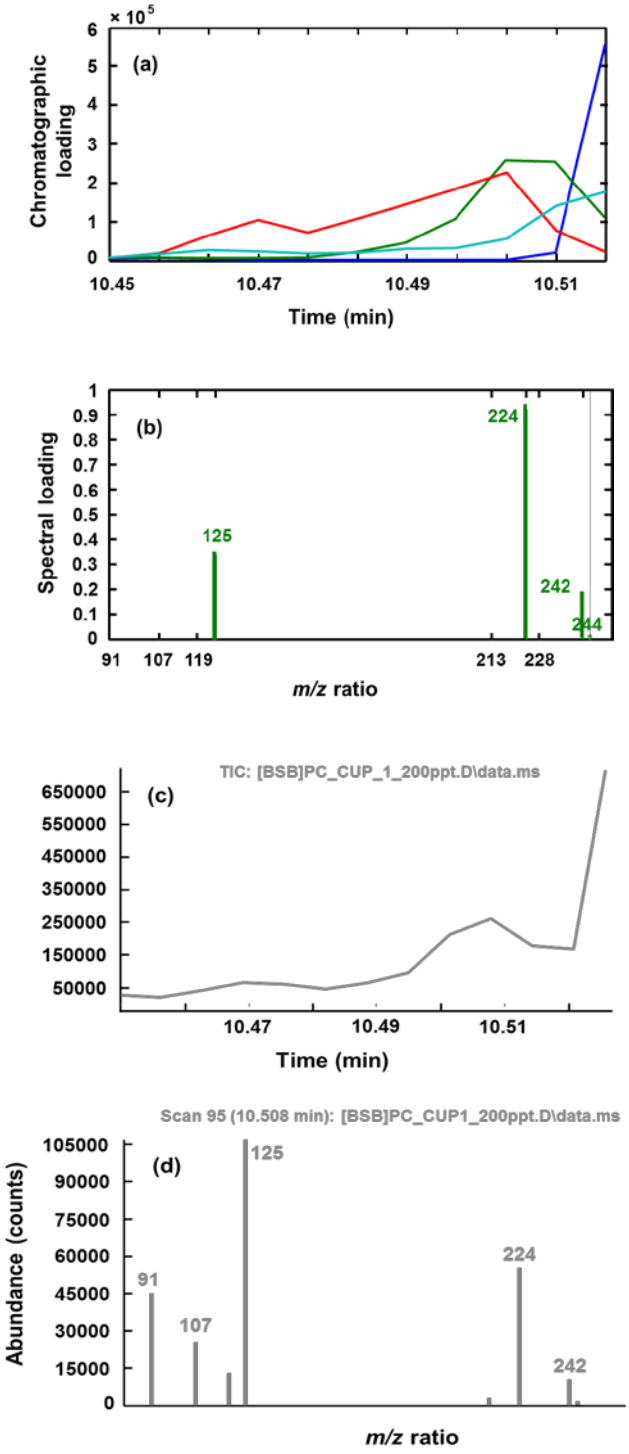


**Figure 4.2.6.** Chromatographic and spectral data derived from the PARAFAC decomposition of the BPA-$d_{16}$ tensor and those from the original TIC registered for one of the samples. **(a)** Loadings of the chromatographic profiles of the 4-factor PARAFAC model; Factor 1 (in blue) matched BPA, while Factor 2 (in green) did BPA-$d_{16}$. **(b)** Loadings of the spectral profile for Factor 2. **(c)** Enlarged TIC of the simulant sample 22 (see Table 4.2.5) within the elution zone of BPA-$d_{16}$. **(d)** Mass spectrum registered at the maximum of the peak considered as BPA-$d_{16}$.

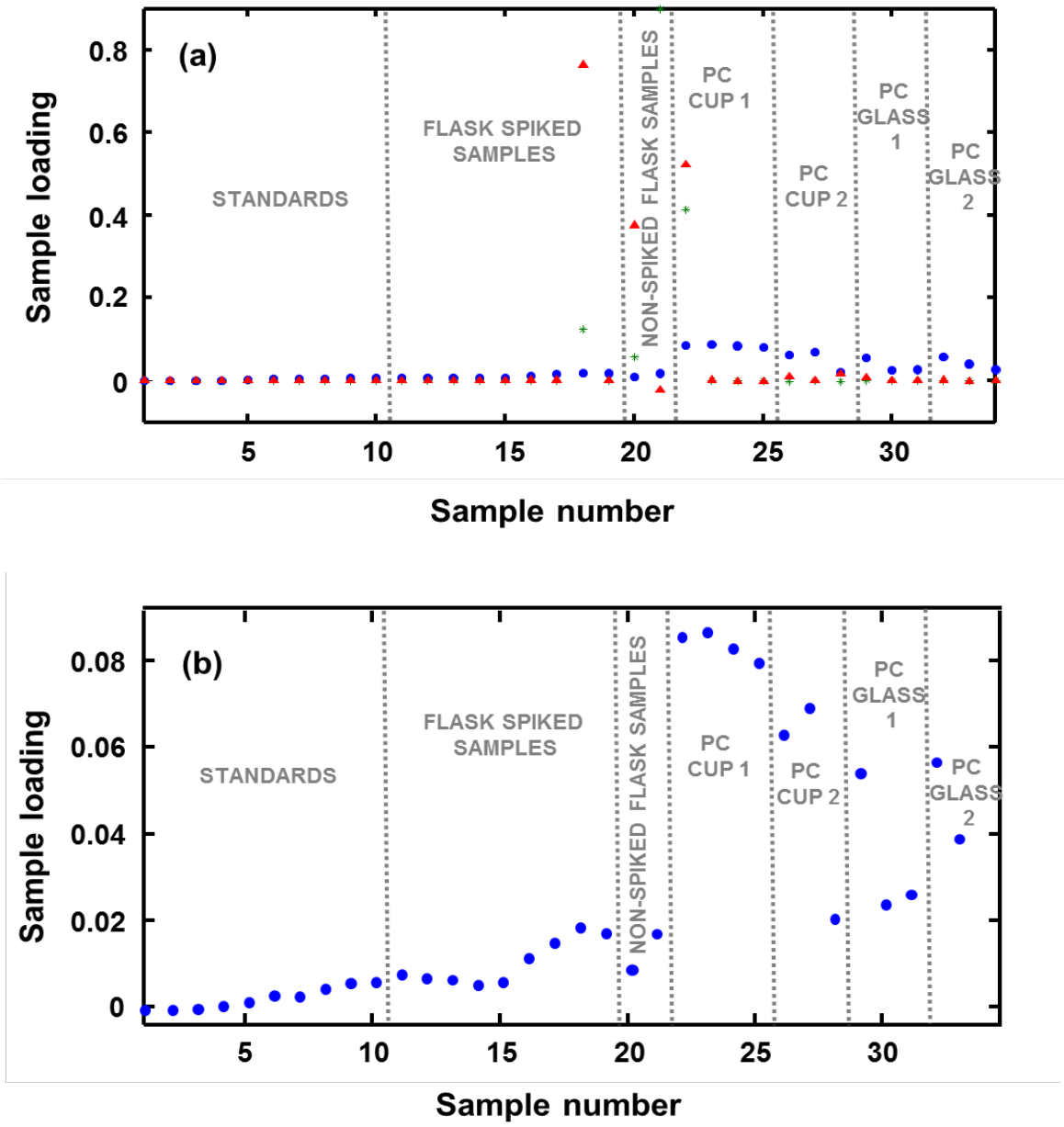


**Figure 4.2.7.** Sample-mode loadings from the PARAFAC decomposition of the BPA tensor including the calibration standards (from 0 to 40 μg $L^{-1}$) and the samples in Table 4.2.5.

**(a)** The three factors: Factor 1 (BPA), in blue dots; Factor 2, in green asterisks; Factor 3, in red triangles.

**(b)** Factor 1 (BPA).

The fact that the number of factors in the PARAFAC models when both standards and samples were included in the data tensors was greater than that when only standards were is remarkable. That increase may be due to the coelution of interferents that were present in some samples. The injection volume of 10 μL may have also contributed to this situation, since the more sample volume injected, the higher the risk of coelution. However, what could have meant a severe problem in the identification and quantification of the analytes did not, because in all cases PARAFAC succeeded in extracting the factor associated to every analyte correctly, even in the presence of coeluents. This allowed the unequivocal identification of all the compounds to be verified.

The standardized sample-mode loadings of some samples for BPF and/or BPA were greater than those of the 40 μg $L^{-1}$ standards, so two new sets of calibration standards (from 30 to 90 μg $L^{-1}$ of BPF and BPA and, additionally, from 100 to 650 μg $L^{-1}$ of BPA) were prepared and analysed in order to avoid extrapolation when determining the concentration of those samples. The second new calibration range exclusively designed for BPA was aimed at quantifying that analyte in all the PC samples. Therefore, three new tensors were built including in each case the new standards and only the samples that were out of the original calibration range: one for BPF, with dimension 13×5×18 (9 standards from 30 to 90 μg $L^{-1}$, together with 4 samples); and the other two for BPA, with dimension (i) 15×9×12 (9 standards from 30 to 90 μg $L^{-1}$, together with 6 samples) and (ii) 22×9×12 (9 standards from 100 to 650 μg $L^{-1}$, together with the 13 PC samples). In these three new data tensors, two standards were aimed at estimating the relative error in prediction. A PARAFAC decomposition of each of these three tensors was then performed, and again, from the relative ratios of the loadings of the spectral profiles and from the relative retention time, it was verified that both BPF and BPA were unequivocally identified. New LMS and LS regressions were estimated from the corresponding PARAFAC model. The results of all the LMS and LS regressions built and the *p*-values for the hypothesis tests posed to validate every model are listed in Table 4.2.8. The mean of the absolute value of the relative errors, both in calibration (from 0.9% to 8.8%) and in prediction (from 2.1% to 10.6%), also figures on this table.

Two hypothesis tests enabled to verify the trueness of every calibration model by checking if, at a 95% significance level, there were no statistically significant differences between the values obtained, respectively, for the slope and 1, and the intercept and 0, of a linear LS regression "*Estimated concentration versus True concentration*". The equations of all the accuracy lines and the *p*-values for those two hypothesis tests are collected in Table 4.2.9.

The detection capability for every analyte was determined for probabilities of false positive ($\alpha$) and false negative ($\beta$) equal to 0.05, so values between 1.97 and 5.53 μg $L^{-1}$ were

achieved (see Table 4.2.9, last column). The detection capability regarding the specific migration limit of 0.6 mg $kg^{-1}$ established for BPA in [18] was also calculated. As a result, $CC\alpha$ equalled 630.3 μg $L^{-1}$ and $CC\beta$, 657.2 μg $L^{-1}$, which meant that the GC-MS method was able to distinguish 657.2 from 600 μg $L^{-1}$ with probabilities of false non-compliance and false compliance equal to 0.05.

**Table 4.2.8.** Results of the LMS and LS regressions performed "*Standardized sample loading versus True concentration*".

| Analyte | Calibration range (μg $L^{-1}$) | LMS regression | | LS regression | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Model | Outlier ratio | Model ($R^2$; $s_{yx}$) | *n* | Hypothesis test | *p*-value | MAE in calibr.[f] | MAE in pred.[g] |
| BPF | 0 – 40 | $y = 9.00 \cdot 10^{-3} + 4.28 \cdot 10^{-2} x$ | 0/10 | $y = -7.68 \cdot 10^{-3} + 4.27 \cdot 10^{-2} x$ (99.78%; $3.01 \cdot 10^{-2}$) | 10 | F-test[a] | 0.00 | 5.54% | Not estimated |
| | | | | | | $\chi^2$ test[b] | 0.33 | | |
| | | | | | | S-W test[c] | 0.05 | | |
| | | | | | | K-S test[d] | 0.87 | | |
| | | | | | | D-W test[e] | 0.50 | | |
| | 30 – 90 | $y = -0.23 + 3.05 \cdot 10^{-2} x$ | 2/7 | $y = -0.23 + 3.06 \cdot 10^{-2} x$ (99.90%; $2.57 \cdot 10^{-2}$) | 5 | F-test | 0.00 | 0.94% | 2.14% |
| | | | | | | $\chi^2$ test | N.C.[h] | | |
| | | | | | | S-W test | 0.58 | | |
| | | | | | | K-S test | 0.23 | | |
| | | | | | | D-W test | 0.63 | | |
| BPA | 0 – 40 | $y = 3.00 \cdot 10^{-3} + 1.13 \cdot 10^{-3} x$ | 2/10 | $y = 2.50 \cdot 10^{-3} + 1.15 \cdot 10^{-3} x$ (99.94%; $4.99 \cdot 10^{-4}$) | 8 | F-test | 0.00 | 6.39% | Not estimated |
| | | | | | | $\chi^2$ test | 0.35 | | |
| | | | | | | S-W test | 0.14 | | |
| | | | | | | K-S test | 0.62 | | |
| | | | | | | D-W test | 0.56 | | |
| | 30 – 90 | $y = -9.97 \cdot 10^{-2} + 2.37 \cdot 10^{-2} x$ | 1/7 | $y = -0.14 + 2.43 \cdot 10^{-2} x$ (99.88%; $2.15 \cdot 10^{-2}$) | 6 | F-test | 0.00 | 1.34% | 10.59% |
| | | | | | | $\chi^2$ test | N.C. | | |
| | | | | | | S-W test | 0.73 | | |
| | | | | | | K-S test | 0.94 | | |
| | | | | | | D-W test | 0.14 | | |

**Table 4.2.8. (cont.)**

| Analyte | Calibration range ($\mu g\ L^{-1}$) | LMS regression | | LS regression | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Model | Outlier ratio | Model ($R^2$; $s_{yx}$) | $n$ | Hypothesis test | $p$-value | MAE in calibr.[f] | MAE in pred.[g] |
| BPA (cont.) | 100 – 650 | $y = -8.75\cdot10^{-2} + 3.90\cdot10^{-3}\,x$ | 3/7 | $y = -8.49\cdot10^{-2} + 3.89\cdot10^{-3}\,x$ (99.92%; $3.04\cdot10^{-2}$) | 4 | F-test<br>$\chi^2$ test<br>S-W test<br>K-S test<br>D-W test | 0.00<br>0.17<br>0.85<br>N.C.<br>0.09 | 2.72% | 7.02% |
| BFDGE | 0 – 40 | $y = 1.05\cdot10^{-1} + 6.35\cdot10^{-2}\,x$ | 1/10 | $y = 1.25\cdot10^{-1} + 6.40\cdot10^{-2}\,x$ (99.56%; $6.18\cdot10^{-2}$) | 9 | F-test<br>$\chi^2$ test<br>S-W test<br>K-S test<br>D-W test | 0.00<br>0.42<br>0.25<br>0.77<br>0.72 | 5.76% | Not estimated |
| BADGE | 0 – 40 | $y = 4.35\cdot10^{-2} + 6.21\cdot10^{-2}\,x$ | 2/10 | $y = 6.98\cdot10^{-2} + 6.30\cdot10^{-2}\,x$ (99.31%; $8.13\cdot10^{-2}$) | 8 | F-test<br>$\chi^2$ test<br>S-W test<br>K-S test<br>D-W test | 0.00<br>0.12<br>0.02<br>0.39<br>0.51 | 8.77% | Not estimated |

([a]) Test for the significance of the regression, where $H_0$: the regression model is not statistically significant.

([b]) Chi-square test, where $H_0$: residuals are compatible with a normal distribution.

([c]) Shapiro-Wilk test, where $H_0$: residuals are compatible with a normal distribution.

([d]) Kolmogorov-Smirnov test, where $H_0$: residuals are compatible with a normal distribution.

([e]) Durbin-Watson test, where $H_0$: residuals are not autocorrelated.

([f]) Mean of the absolute value of the relative errors in calibration.

([g]) Mean of the absolute value of the relative errors in prediction.

([h]) Not computed.

**Table 4.2.9.** Accuracy analysis and values of the decision and detection limits.

| Analyte | Calibration range (μg $L^{-1}$) | LS regression "*Predicted concentration versus True concentration*" | | | | *CCα*; *CCβ* (μg $L^{-1}$) |
|---|---|---|---|---|---|---|
| | | *n* | Model ($R^2$; $s_{yx}$) | *p*-value for the test $b_0 = 0$ | *p*-value for the test $b_1 = 1$ | |
| BPF | 0 – 40 | 10 | $y = 0.0 + 1.0\,x$ (99.78%; $7.05 \cdot 10^{-1}$) | 1.0 | 1.0 | 2.85; 2.90[a] |
| | 30 – 90 | 5 | $y = 4.72 \cdot 10^{-5} + 1.0\,x$ (99.90%; $8.40 \cdot 10^{-1}$) | 0.99 | 0.99 | — |
| BPA | 0 – 40 | 8 | $y = 4.82 \cdot 10^{-5} + 9.99 \cdot 10^{-1}\,x$ (99.94%; $4.35 \cdot 10^{-1}$) | 0.99 | 0.99 | 1.968; 1.969[a] |
| | 30 – 90 | 6 | $y = 0.0 + 1.0\,x$ (99.88%; $8.86 \cdot 10^{-1}$) | 1.0 | 1.0 | — |
| | 100 – 650 | 4 | $y = 0.0 + 1.0\,x$ (99.92%; 7.82) | 1.0 | 1.0 | 630.3; 657.2[b] |
| BFDGE | 0 – 40 | 9 | $y = 2.32 \cdot 10^{-4} + 9.99 \cdot 10^{-1}\,x$ (99.56%; $9.65 \cdot 10^{-1}$) | 0.99 | 0.99 | 4.07; 4.17[a] |
| BADGE | 0 – 40 | 8 | $y = 1.11 \cdot 10^{-4} + 9.99 \cdot 10^{-1}\,x$ (99.31%; 1.29) | 0.99 | 0.99 | 5.40; 5.53[a] |

([a]) For a nominal value $x_0 = 0$, $\alpha = \beta = 0.05$.

([b]) For a nominal value $x_0 = 600$ μg $L^{-1}$, $\alpha = \beta = 0.05$.

The LS calibration models (see Table 4.2.8) enabled to determine the quantity of every analyte in every food simulant sample. The values of the concentration of every analyte in the 24 final simulant samples are shown in the last four columns of Table 4.2.5. On the other hand, the repeatability and the recovery of the method regarding every analyte were also estimated. The first performance characteristic was calculated in all cases as the standard deviation of the concentration of every analyte in the five repeatability simulant samples; in this way, the value for the repeatability standard deviation was 2.80 for BPF, 5.53 for BPA, 6.95 for BFDGE and 6.62 for BADGE. Recovery was expressed as the percentage of the amount initially added at every spiked sample that was recovered at last. The average recovery rates appear in Table 4.2.10. It must be noticed that, in general, better recoveries were obtained when the food simulant had been previously in contact with a PC container. On the other hand, the sample preparation method clearly performed better for BPF and BPA than for BFDGE and BADGE, as recoveries were higher for the first two compounds. This is the reason why that step will be optimized throughout Section 4.3.

**Table 4.2.10.** Recovery rates of every analyte at the two fortification levels considered.

| Analyte | Recovery rate (%) | | | |
|---|---|---|---|---|
| | Fortification level of 200 ng $L^{-1}$ | | Fortification level of 400 ng $L^{-1}$ | |
| | Samples not in contact with PC tableware | Samples in contact with PC tableware | Samples not in contact with PC tableware | Samples in contact with PC tableware |
| BPF | 55.17 | 70.13 | 51.11 | 76.55 |
| BPA | 149.57 | — | 125.53 | 113.95 |
| BFDGE | 24.99 | 30.23 | 26.94 | 43.06 |
| BADGE | 24.34 | 0 | 13.63 | 34.80 |

It was confirmed that the migration of BPA from the new PC containers into the food simulant had occurred, since the concentration of BPA in every PC non-spiked sample was different from 0. To be exact, the average BPA concentration in the simulant volume analysed was 181.46 μg $L^{-1}$ for Cup 1, 148.64 μg $L^{-1}$ for Cup 2, 104.67 μg $L^{-1}$ for Glass 1, and 108.88 μg $L^{-1}$ for Glass 2. Bearing in mind the concentration factor and the recovery rate (113.95%), the average BPA concentration migrated from every PC container was 1.27 μg $L^{-1}$, 1.04 μg $L^{-1}$, 0.73 μg $L^{-1}$, and 0.76 μg $L^{-1}$, respectively. None of these values exceeded the specific migration limit laid down for BPA (0.6 mg $kg^{-1}$) in [18] on plastic materials and articles intended to come into contact with food.

### 4.2.6. Acknowledgements

Authors thank the financial support via projects Ministerio de Economía y Competitividad CTQ2011-26022 and JCyL BU108A11-2. M.L. Oca is particularly grateful to Universidad de Burgos for her FPI grant.

### 4.2.7. References

[1] D. Arroyo, M.C. Ortiz, L.A. Sarabia, *Optimization of the derivatization reaction and the solid-phase microextraction conditions using a D-optimal design and three-way calibration in the determination of non-steroidal anti-inflammatory drugs in bovine milk by gas chromatography-mass spectrometr*y, Journal of Chromatography A 1218 (2011) 4487-4497.

[2] R. Bro, *Review on Multiway Analysis in Chemistry – 2000-2005*, Critical Reviews in Analytical Chemistry 36 (2006) 279-293.

[3] M.C. Ortiz, L. Sarabia, *Quantitative determination in chromatographic analysis based on n-way calibration strategies*, Journal of Chromatography A 1158 (2007) 94-110.

[4] V. Gómez, M.P. Callao, *Analytical applications of second-order calibration methods*, Analytica Chimica Acta 627 (2008) 169-183.

[5] R. Bro, N. Viereck, M. Toft, H. Toft, P.I. Hansen, S.B. Engelsen, *Mathematical chromatography solves the cocktail party effect in mixtures using 2D spectra and PARAFAC*, Trends in Analytical Chemistry 29 (2010) 281-284.

[6] L. Rubio, L.A. Sarabia, A. Herrero, M.C. Ortiz, *Advantages of a programmed temperature vaporizer inlet and parallel factor analysis in the determination of triazines in the presence of non-intentionally added substances by gas chromatography*, Analytical and Bioanalytical Chemistry 403 (2012) 1131-1143.

[7] R. Morales, M.C. Ortiz, L. A. Sarabia, *Usefulness of a PARAFAC decomposition in the fiber selection procedure to determine chlorophenols by means SPME-GC-MS*, Analytical and Bioanalytical Chemistry 403 (2012) 1095-1107.

[8] D. Arroyo, M.C. Ortiz, L.A. Sarabia, *Multiresponse optimization and parallel factor analysis, useful tools in the determination of estrogens by gas chromatography–mass spectrometry*, Journal of Chromatography A 1157 (2007) 358-368.

[9] A. Guart, F. Bono-Blay, A. Borrell, S. Lacorte, *Migration of plasticizers phthalates, bisphenol A and alkylphenols from plastic containers and evaluation of risk*, Food Additives and Contaminants: Part A 28 (2011) 676-685.

[10] W.F. Smith, *Foundations of Materials Science and Engineering*, McGraw-Hill, New York, fourth ed., 2005.

[11] A. Ballesteros-Gómez, F.J. Ruiz, S. Rubio, D. Pérez-Bendito, *Determination of bisphenols A and F and their diglycidyl ethers in wastewater and river water by coacervative extraction and liquid chromatography–fluorimetry*, Analytica Chimica Acta 603 (2007) 51-59.

[12] J. Simal-Gandara, P. López Mahía, P. Paseiro Losada, J. Simal Lozano, S. Paz Abuín, *Overall migration and specific migration of bisphenol A diglycidyl ether monomer and m-xylylenediamine hardener from an optimized epoxy-amine formulation into water-based food simulants*, Food Additives and Contaminants 10 (1993) 555-565.

[13] T. Yamamoto, A. Yasuhara, *Quantities of bisphenol a leached from plastic waste samples*, Chemosphere 38 (1999) 2569-2576.

[14] R. Gachter, H. Muller, *Plastic Additives Handbook*, Hanser, New York, 1990.

[15] K.A. Mountfort, J. Kelly, S.M. Jickells, L. Castle, *Investigations into the potential degradation of polycarbonate baby bottles during sterilization with consequent release of bisphenol A*, Food Additives and Contaminants 14 (1997) 737-740.

[16] K.A. Ehlert, C.W.E. Beumer, M.C.E. Groot, *Migration of bisphenol A into water from polycarbonate baby bottles during microwave heating*, Food Additives and Contaminants: Part A 25 (2008) 904-910.

[17] A.G. Cabado, S. Aldea, C. Porro, G. Ojea, J. Lago, C. Sobrado, J.M. Vieites, *Migration of BADGE (bisphenol A diglycidyl-ether) and BFDGE (bisphenol F diglycidyl-ether) in canned seafood*, Food and Chemical Toxicology 46 (2008) 1674-1680.

[18] Commission Regulation (EU) No 10/2011 of 14 January 2011 on plastic materials and articles intended to come into contact with food, Official Journal of the European Union, L12, 1-89.

[19] A. Ballesteros-Gómez, S. Rubio, D. Pérez-Bendito, *Analytical methods for the determination of bisphenol A in food*, Journal of Chromatography A 1216 (2009) 449-469.

[20] E. Ferrer, E. Santoni, S. Vittori, G. Font, J. Mañes, G. Sagratini, *Simultaneous determination of bisphenol A, octylphenol, and nonylphenol by pressurised liquid extraction and liquid chromatography–tandem mass spectrometry in powdered milk and infant formulas*, Food Chemistry 126 (2011) 360-367.

[21] B. Shao, H. Han, D. Li, Y. Ma, X. Tu, Y. Wu, *Analysis of alkylphenol and bisphenol A in meat by accelerated solvent extraction and liquid chromatography with tandem mass spectrometry*, Food Chemistry 105 (2007) 1236-1241.

[22] J.E. Biles, T.P. McNeal, T.H. Begley, H.C. Hollifield, *Determination of Bisphenol-A in Reusable Polycarbonate Food-Contact Plastics and Migration to Food-Simulating Liquids*, Journal of Agricultural and Food Chemistry 45 (1997) 3541-3544.

[23] J. López-Cervantes, P. Paseiro-Losada, *Determination of bisphenol A in, and its migration from, PVC stretch film used for food packaging*, Food Additives and Contaminants 20 (2003) 596-606.

[24] M. Del Olmo, A. González-Casado, N.A. Navas, J.L. Vílchez, *Determination of bisphenol A (BPA) in water by gas chromatography-mass spectrometry*, Analytica Chimica Acta 346 (1997) 87-92.

[25] D.A. Markham, D.A. Mcnett, J.H. Birk, G.M. Klecka, M.J. Bartels, C.A. Staples, *Quantitative determination of bisphenol-A in river water by cool-on-column injection gas chromatography mass spectrometry*, International Journal of Environmental and Analytical Chemistry 69 (1998) 83-98.

[26] A.K. Malik, C. Blasco, Y. Picó, *Liquid chromatography–mass spectrometry in food safety*, Journal of Chromatography A 1217 (2010) 4018-4040.

[27] N. De Coensel, F. David, P. Sandra, *Study on the migration of bisphenol-A from baby bottles by stir bar sorptive extraction-thermal desorption-capillary GC-MS*, Journal of Separation Science 32 (2009) 3829-3836.

[28] P. Viñas, N. Campillo, N. Martínez-Castillo, M. Hernández-Córdoba, *Comparison of two derivatization-based methods for solid-phase microextraction–gas chromatography–mass spectrometric determination of bisphenol A, bisphenol S and biphenol migrated from food cans*, Analytical and Bioanalytical Chemistry 397 (2010) 115-125.

[29] J.D. Stuart, C.P. Capulong, K.D. Launer, X. Pan, *Analyses of phenolic endocrine disrupting chemicals in marine samples by both gas and liquid chromatography–mass spectrometry*, Journal of Chromatography A 1079 (2005) 136-145.

[30] N. Rastkari, R. Ahmadkhaniha, M. Yunesian, L.J. Baleh, A. Mesdaghinia, *Sensitive determination of bisphenol A and bisphenol F in canned food using a solid-phase microextraction fibre coated with single-walled carbon nanotubes before GC/MS*, Food Additives & Contaminants: Part A 27 (2010) 1460-1468.

[31] E. Hoh, K. Mastovska, *Large volume injection techniques in capillary gas chromatography*, Journal of Chromatography A 1186 (2008) 2-15.

[32] A.C. Heiden, B. Kolahgar, E. Pfannkoch. Gerstel AppNote 7/2001, Gerstel GmbH & Co. KG, Mülheim an der Ruhr. Available at: http://www.gerstel.com/pdf/p-gc-an-2001-07.pdf. Accessed 29 July 2011

[33] C.A. Andersson, R. Bro, *The N-way Toolbox for MATLAB*, Chemometrics and Intelligent Laboratory Systems 52 (2000) 1-4.

[34] R. Bro, *Explorative study of sugar production using fluorescence spectroscopy and PARAFAC analysis*, Chemometrics and Intelligent Laboratory Systems 46 (1999) 133-147.

[35] R.B. Cattell, *Parallel proportional profiles and other principles for determining the choice of factors by rotation*, Psychometrika 9 (1944) 267-283.

[36] R. Bro, PhD Thesis, Royal Veterinary and Agricultural University, Denmark, 1998.

[37] R. Bro, H.A.L. Kiers, A new efficient method for determining the number of components in PARAFAC models, Journal of Chemometrics 17 (2003) 274-286.

[38] T. Pihlström, *Method validation and quality control procedures for pesticide residues analysis in food and feed (SANCO/10684/2009)*. EU, Brussels, 2010.

[39] Commission Decision (EC) No 2002/657/EC of 12 August 2002, Brussels, implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Official Journal of the European Communities, L 221, 8- 36.

[40] International Standard ISO 11843, *Capability of detection*, Geneva, Switzerland, *Part 1, Terms and definitions*, 1997 and *Part 2, Methodology in the linear calibration case*, 2000.

[41] J. Inczédy, T. Lengyel, A.M. Ure, A. Gelencsér, A. Hulanicki, *Compendium of Analytical Nomenclature IUPAC*, third ed., Pot City Press Inc., Baltimore, 2000.

[42] Working Group 2, *Joint Committee for Guides in Metrology International vocabulary of metrology. Basic and general concepts and associated terms*, third ed., approved and adopted by each of the eight JCGM member organizations: BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML is available since 2008-06, Hard copies or downloads are on sale from ISO there, Free copies can be downloaded from the BIPM website in Paris-Sèvres. www.bipm.org.

[43] IUPAC, *Nomenclature, symbols, units and their usage in spectrochemical analysis – II. Data interpretation, analytical chemistry division*, Spectrochimica Acta Part B 33 (1978) 241-245.

[44] L.A. Currie, *Detection and quantification limits: origins and historical overview*, Analytica Chimica Acta 391 (1999) 127-134.

[45] C.A. Clayton, J.W. Hines, P.D. Elkins, *Detection limits with specified assurance probabilities*, Analytical Chemistry 59 (1987) 2506-2514.

[46] L. Sarabia, M.C. Ortiz, *DETARCHI: A program for detection limits with specified assurance probabilities and characteristic curves of detection*, Trends in Analytical Chemistry 13 (1994) 1-6.

[47] M.C. Ortiz, M.S. Sánchez, L.A. Sarabia, *Quality of Analytical Measurements: Univariate Regression,* in: S. Brown, R. Tauler, B. Walczak (Eds.), *Comprehensive Chemometrics, vol. 1*, Elsevier, Oxford, 2009.

[48] M.C. Ortiz, L.A. Sarabia, I. García, D. Giménez, E. Meléndez, *Capability of detection and three-way data*, Analytica Chimica Acta 559 (2006) 124-136.

[49] M.C. Ortiz, L. Sarabia, M.S. Sánchez, *Tutorial on evaluation of type I and type II errors in chemical analyses: From the analytical detection to authentication of products and process control*, Analytica Chimica Acta 674 (2010) 123-142.

[50] NIST Mass Spectral Search Program for the NIST/EPA/NIH Mass Spectral Library Version 2.0 a. build July 1 2002.

[51] B.M. Wise, N.B. Gallagher, R. Bro, J.M. Shaver, W. Windig, R.S. Koch, *PLS Toolbox 5.0*, Eigenvector Research, Inc. 3905 West Eaglerock Drive, Wenatchee, 2006.

[52] *STATGRAPHICS Centurion XVI (Version 16.1.11)*, StatPoint Technologies, Inc., Herndon, VA, 2010.

[53] P.J. Rousseeuw, A.M. Leroy, *Robust regression and outliers detection*, John Wiley and Sons, New Jersey, 2001.

[54] R. Sendón García, C. Pérez Lamela, P. Paseiro Losada, *Determination of bisphenol F diglycidyl ether and related compounds by high-performance liquid chromatography/mass spectrometry*, Rapid Communications in Mass Spectrometry 19 (2005) 1569-1574.

[55] S. Bratinova, B. Raffael, C. Simoneau, *Guidelines for performance criteria and validation procedures of analytical methods used in controls of food contact materials - EUR 24105 EN (JRC53034)*, Publications Office of the European Union, Luxembourg, 1st ed., 2009.

## 4.3. Optimum pH for the determination of bisphenols and their corresponding diglycidyl ethers by gas chromatography-mass spectrometry. Migration kinetics of bisphenol A from polycarbonate glasses[2]

### 4.3.1. Abstract

This work presents, on the one hand, the study of the influence of the pH of the medium on the determination of bisphenol F (BPF), bisphenol A (BPA) and their corresponding diglycidyl ethers (BFDGE and BADGE, respectively) by GC-MS after a solid-phase extraction step, using BPA-$d_{16}$ as internal standard and Parallel Factor Analysis (PARAFAC) decompositions as a multi-way tool for the unequivocal identification and quantification of the four analytes. As the structure of both BFDGE and BADGE has two 2,3-epoxypropoxy groups that can undergo an acid- or base-catalyzed ring-opening via nucleophilic substitution reactions, several samples spiked with the four analytes were set to different pH values between 2 and 12. The best results were obtained in the pH region 8–10, being 9 the most suitable value. Coelution of interferents was overcome using the PARAFAC decomposition; otherwise, the presence of some analytes could not have been ensured according to the regulations currently in force.

Secondly, the release of BPA from polycarbonate glasses into food simulant *D1* (ethanol 50 % (v/v)) over time was studied through seven migration tests and the differences found in this migration process with the incubation temperature (50 and 70 ºC) were evaluated. A nonlinear regression was used to fit the experimental data following an exponential relation between the concentration of BPA transferred from every glass and the respective migration test. None of the quantities of BPA released exceeded the specific migration limit of 0.6 mg $kg^{-1}$ laid down for this compound in the Commission Regulation (EU) No 10/2011, so the compliance of the glasses evaluated was ensured. The average recovery percentages of the four analytes at a fortification level of 800 ng $L^{-1}$ ranged from 50.14% to 92.75%. The detection capability (*CCβ*) of the method for BPA was 2.60 μg $L^{-1}$ for $n = 2$ replicates, with probabilities of false positive and false negative fixed at 0.05.

### 4.3.2. Introduction

Polycarbonates (PC) are linear polyesters produced by the reaction of phosgene with bisphenol A (BPA). A wide variety of food contact materials are derived from PCs. BPA is also used in the production of thermal paper for cash receipts, which may contain relatively high levels of this chemical [1] that can be transferred readily to the skin in small amounts

[2] This section is published as *M.L. Oca et al. / J. Chromatogr. A 1360 (2014) 23–38* (18 citations up to August 30, 2021 according to *Scopus*).

[2]. The controversy over the safety of BPA, considered an endocrine disruptor that can migrate from the packaging material into food, has led to its reduction from food containers and from surface treatments for containers that are thermally processed [3] and to the prohibition of the use of BPA for the manufacture, placing on the market and importation into the European Union of PC feeding bottles for infants since mid-2011 [4]; this regulation amends the Commission Regulation (EU) No 10/2011 [5], where a specific migration limit for BPA of 0.6 mg $kg^{-1}$ is set out. However, the resolutions on BPA are being revised by the European Union after food agencies of several member states have informed lately about the need to remove this compound from any kind of materials intended to come into contact with food.

On the other hand, the lacquers used as inner coatings of food cans and containers to reduce food spoilage are also under consideration. Some of these lacquers are epoxy phenolic resins based on polymerization products of bisphenol F diglycidyl ether (BFDGE) or bisphenol A diglycidyl ether (BADGE) that can release these compounds, as well as oligomers and derivatives, by degradation. BFDGE and BADGE are generated by the reaction of bisphenol F (BPF) and BPA, respectively, with epichlorohydrin, so both hydroxyl compounds are also likely to migrate as non-reacted monomers from the epoxy coating into foods and beverages. With regard to their legal framework, the use and/or presence of BFDGE in the manufacture of materials and articles intended to come into contact with food is prohibited, while specific migration limits have been established for the sum of BADGE and its hydrolyzed derivatives (9 mg $kg^{-1}$ or 9 mg/6 $dm^2$) and for the sum of BADGE·HCl, BADGE·2HCl and BADGE·HCl·$H_2O$ (1 mg $kg^{-1}$ or 1 mg/6 $dm^2$) [6].

For materials and articles not yet in contact with food, verification of compliance with specific migration limits shall be carried out either in food or in the food simulants set out in Annex III of [5], which are test media imitating food. In particular, food simulant *D1* (ethanol 50% (v/v)) is assigned for juices, nectars and soft drinks containing fruit pulp, foods that have a lipophilic character, alcoholic foods with an alcohol content of above 20% and for oil in water emulsions, all of them able to extract lipophilic substances [5]. This is the context which the present work is framed in. Here the simultaneous determination of BPF, BPA, BFDGE and BADGE (BPA-$d_{16}$ as internal standard) in food simulant *D1* using a programmed-temperature vaporizer–gas chromatography-mass spectrometry (PTV-GC-MS) system is pursued, together with the kinetics study for the migration of BPA from PC glasses not yet in contact with food into simulant *D1*. All these compounds can be considered lipophilic.

The development of this multiresidue method must bear in mind the distinct nature of the substituent groups on the aromatic ring of bisphenols and their diglycidyl ethers, which suggests that their chemical behaviour should differ. Both BFDGE and BADGE contain an epoxide functional group as part of their 2,3-epoxypropoxy substituents, while BPF and BPA are phenol derivatives. Epoxides show a great reactivity toward nucleophilic reagents because of their angle strain, so reactions that open the ring with cleavage of the carbon-oxygen bond relieve this strain [7]. These reactions are carried out in two different ways depending on pH (neutral or basic media versus acidic media). As both BFDGE and BADGE are epoxy derivatives, they are expected to undergo nucleophilic substitution reactions that open their epoxide rings, yielding compounds different from the target ones. This may indeed occur when they come into contact with aqueous or acidic food during storage, thus generating their corresponding chlorinated and/or hydrolyzed derivatives, such as, for example, BFDGE·2$H_2O$, BFDGE·2HCl, BADGE·$H_2O$·HCl, BADGE·$H_2O$ or BADGE·HCl [8]. The kinetics of the hydrolysis of BADGE and BFDGE in water-based food simulants have been studied in [9] and [10], respectively. Several analytical methods have been developed for the determination of BADGE, BFDGE and their derivatives in different foodstuffs [8,11,12,13,14,15,16].

pH plays a major role in the stability of epoxy derivatives such as BFDGE and BADGE in aqueous media, as well as it does in that of BPF and BPA because of the presence of hydroxyl groups in their structures. Since the estimated $pK_a$ at 25 ºC of these bisphenols is 7.55-10.80 for BPF [17] and 9.59-11.30 for BPA [18], the most appropriate pH for the simultaneous determination of both compounds at that temperature is expected to be less than these values, the lower the better. Otherwise, there would be a risk of losing them as the salt of the corresponding bisphenolate anion, so extremely basic pH conditions are thought to be unsuitable. However, very acidic media, while the best option for BPF and BPA, will cause the loss of BFDGE and BADGE by a ring-opening reaction of the epoxide group.

In this context, this work deals first with the study of the influence of pH on the simultaneous determination of these four analytes in food simulant *D1*. Secondly, the migration kinetics of BPA from PC glasses into food simulant *D1* has been investigated as a function of the heating temperature during the migration testing. The possibility of posing a general mathematical model with analytical sense capable of making theoretical predictions about the migration of BPA from any container of this kind is then evaluated. If achieved, the resultant model would involve a more economical methodology that would reduce the time-consuming and laborious steps of sample treatment and analysis as much as possible by only knowing the number of migration tests performed and the concentration of BPA in the food simulant after the first migration test. Different

mathematical functions have been proposed to describe the behaviour of other migrants from plastic materials [19,20,21,22,23].

Parallel Factor Analysis (PARAFAC) decomposition has been used for both the identification and quantification steps of the analytes in all the stages of this work. It has proved to be decisive for the achievement of these objectives according to the guidelines in [24], where it is established that the presence of a migrant in a sample analysed by GC-MS will be confirmed if both its relative retention time and the relative intensities of its diagnostic ions agree with those of a standard at comparable concentrations, measured under the same conditions, within the corresponding maximum permitted tolerances. No reasonable conclusions on an analyte can be drawn without first having ensured its unequivocal presence in the sample. This identification, based on mass spectra and retention times, guarantees the specificity of the analysis, even though shared ions are present in the same retention time.

## 4.3.3. Theory

### *4.3.3.1. PARAFAC*

PARAFAC is a decomposition method for multi-way data. An example of three-way data is that provided by GC-MS, where, for each of the $K$ samples analysed, the abundance of $J$ characteristic ions is recorded at $I$ times around the retention time of every compound, producing a mass spectrum at every scan; the resultant experimental data obtained for every peak can thus be arranged in a cube or tensor $\underline{\mathbf{X}}$ with dimension $I \times J \times K$.

PARAFAC performs a decomposition of the data tensor $\underline{\mathbf{X}}$ into trilinear factors [25], where each one of them consists of three loading vectors $\mathbf{a_f}$, $\mathbf{b_f}$ and $\mathbf{c_f}$. So, a PARAFAC model of $\underline{\mathbf{X}}$ is given by three loading matrices, **A**, **B**, and **C,** whose $f$-th column is $\mathbf{a_f}$, $\mathbf{b_f}$ and $\mathbf{c_f}$, respectively, with elements $a_{if}$, $b_{jf}$, and $c_{kf}$. The model is found by mimimizing the sum of squares of the residuals, that is, the unmodelled experimental error, $e_{ijk}$, in

$$x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk}, \quad i = 1,2,...,I; \quad j = 1,2,...,J; \quad k = 1,2,...,K \tag{4.3.1}$$

where $x_{ijk}$ is an element of $\underline{\mathbf{X}}$ that represents the abundance of the $j$-th ion at the $i$-th time recorded for the $k$-th sample and $F$ is the number of factors in the model. In a GC-MS calibration [26], the vectors $\mathbf{a_f}$, $\mathbf{b_f}$ and $\mathbf{c_f}$ are the chromatographic, spectral and sample profiles of the $f$-th analyte, respectively.

If the GC-MS data are trilinear, i.e., they fulfil Eq. (4.3.1), the PARAFAC least squares solution is unique, and the achievement of the true underlying mass spectrum and chromatogram for every analyte is ensured.

The core consistency diagnostic (CORCONDIA) has been suggested for determining the proper number of factors for multi-way models by measuring the trilinearity degree of the experimental data [27].

The uniqueness property means that each *f*-th factor (analyte) is distinguished by its own chromatographic, spectral and sample modes in a GC-MS determination. This makes it possible to identify compounds unequivocally by their spectral and chromatographic profiles as laid down in several official regulations and guidelines [24,28,29], even if an interferent with shared ions coelutes.

Once the proper number of factors has been decided, the resultant model is validated. Unless otherwise stated, unconstrained models have been built in this work. The indices $Q$ and Hotelling's $T^2$ have been used to identify outlier samples: if both indices of a sample exceed the threshold value at a 99 % confidence level, that sample will be rejected and the PARAFAC model will be estimated again.

#### *4.3.3.2. Comparison of regression curves*

Suppose that two calibration sets of similar data on a response **y** and a predictor **x** are available, and that a quadratic regression model is considered for each set.

In order to check if the same least squares (LS) model can be used for both sets and, if so, what the fitted coefficients should be, a multivariate regression involving a binary variable, $z$, is posed [30]. $z$ is 0 for the values $(x_i, y_i)$ of the first calibration set, *C1*, and 1 for the data of the second calibration set, *C2*, so the following model is fitted to both sets of data at one time

$$y = \beta_0 + \beta_1 x + \beta_{11} x^2 + a_0 z + a_1 z x + a_{11} z x^2 + \varepsilon \qquad (4.3.2)$$

where $\varepsilon$ denotes the independent, equal and normally distributed residual error. Eq. (4.3.2) provides separate functional models for *C1* and *C2* by letting $z$ attain the value 0 or 1, respectively. The parameters $a_0$, $a_1$ and $a_{11}$ represent the changes needed to get from the *C1* model function to the *C2* one. So, if $a_0 = a_1 = a_{11} = 0$, it can be concluded that both models are the same.

The decision on the equality of two calibration curves is important in chemical analysis to decide if there is a matrix effect and, if so, what type. For this study, calibration set *C1*

would be, for example, made up of matrix-matched standards and calibration set *C2* would consist of standards of similar concentrations in the solvent [31].

#### *4.3.3.3. Migration model*

The functional relation between the concentration of BPA released ($c_{BPA,Test\ t}$) and the number of the migration test performed ($Test\ t$) is established for every container following an exponential decay model that takes the initial values of both variables into consideration

$$c_{BPA,\ Test\ t} = a + \left(c_{BPA,\ Test\ 1} - a\right) e^{-b\left(Test\ t - Test\ 1\right)} \tag{4.3.3}$$

where $Test\ 1 = 1$ and $c_{BPA,Test\,1}$ is the experimental value of the average amount of BPA transferred from the glass to the food simulant after the first migration test.

Eq. (4.3.3) is the solution of the following differential equation

$$\frac{d\left(c_{BPA,\ Test\ t} - a\right)}{d\left(Test\ t\right)} = \frac{d\left(c_{BPA,\ Test\ t}\right)}{d\left(Test\ t\right)} = -b\left(c_{BPA,\ Test\ t} - a\right) \tag{4.3.4}$$

being *a* and *b* constants. Eq. (4.3.3) corresponds to a migration kinetics in which, as can be seen in Eq. (4.3.4), the rate of release of BPA (as a function of $Test\ t$) from the PC container into the food simulant is linearly related to the concentration of BPA in that container.

In Eqs. (4.3.3) and (4.3.4), *a* represents the concentration of BPA in the whole volume of food simulant after infinite migration tests. On the other hand, *b* is the relative rate of release of BPA from a container. Basically, Eq. (4.3.3) shows that the amount of BPA migrated from a certain PC glass decays exponentially at a rate that depends on the parameter *b*, so its value is characteristic of both the polymer (bearing in mind its composition and structure) and the experimental conditions for the migration study (the heating temperature in this work).

### 4.3.4. Material and experimental

#### *4.3.4.1. Chemicals*

BPF, BPA-$d_{16}$ (internal standard), BPA, BFDGE and BADGE, all of 97% or higher purity, were obtained from Sigma–Aldrich (Steinheim, Germany). Both gradient-grade-for-liquid-chromatography acetone and methanol were purchased from Merck (Darmstadt, Germany), while ethanol absolute GPR RECTAPUR® was from VWR International, LLC

(Radnor, Pennsylvania, USA). Hydrochloric acid 37% AnalaR NORMAPUR (VWR International) and sodium hydroxide (Panreac Química SLU, Barcelona, Spain) were used for adjusting the pH of the samples. Deionised water was obtained by using the Milli-Q gradient A10 water purification system from Millipore (Bedford, MA, USA).

#### *4.3.4.2. Standard solutions*

Stock solutions of BPF, BPA-$d_{16}$, BPA, BFDGE and BADGE at 1000 mg $L^{-1}$ were prepared in acetone. Diluted solutions of 5 mg $L^{-1}$ were obtained from the former ones by dilution in the same solvent. Working solutions of the five compounds were also prepared daily in acetone at the concentrations needed in each experimental stage. All the solutions were kept in crimp vials and stored at 4 ºC. The set of standards and/or samples that were evaluated in every study, as well as the dimensions of the resulting data tensors, are detailed in Table 4.3.1.

#### *4.3.4.3. Migration testing*

Two kinetic studies on the migration of BPA from new PC glasses were carried out at 70 and 50 ºC, respectively. Three different glasses were evaluated throughout each of the two studies. Every migration test consisted of filling every glass with 250 mL of food simulant *D1* at room temperature, covering it with a lid and placing the three glasses into a thermostatic water bath at a given temperature (70 ºC for the first kinetic study and 50 ºC for the second one) for 10 h, and then at 25 ºC for 12 h. Seven migration tests were performed on each of the six PC glasses analysed in order to estimate the migration models. All the containers were rinsed with deionised water several times before evaluation.

#### *4.3.4.4. Treatment of migration samples*

In the two kinetic studies, after each migration test, two 50-mL portions of every resultant simulant sample were taken to evaluate the amount of BPA migrated. For tests number 4, 5, 6 and 7 in the 70 ºC kinetics, another two 50-mL aliquots were pipetted out of every glass and spiked with BPF, BPA, BFDGE and BADGE at a concentration of 800 ng $L^{-1}$ to assess the recovery of the method for every analyte. The pH of all these samples was adjusted to 9 using sodium hydroxide before they were pretreated by SPE.

SPE cartridges were conditioned without vacuum by adding 15 mL of methanol and, once 1 min had elapsed, equilibrated with 6 mL of Milli-Q water at pH 9. After 2 min, the whole volume of every sample passed through the SPE cartridge at a flow rate of approximately 3 mL $min^{-1}$ by adjusting vacuum. After every sample had been loaded, cartridges were rinsed with 3 mL of a methanol-water (10:90 v/v) mixture and dried under vacuum for 5 min.

Elution of the analytes was performed using 6 mL of methanol, and then SPE cartridges were left to dry under vacuum for 30 min. Each eluate was collected in a 10-mL glass tube and evaporated to dryness at 40 ºC using the miVac evaporator. Samples were redissolved in 400 μL of a solution containing 200 μg $L^{-1}$ of BPA-$d_{16}$, vortex-mixed for 35 s and transferred into an insert contained in a vial for GC-MS analysis.

#### *4.3.4.5. GC-MS analysis*

Analyses were carried out on an Agilent 7890A gas chromatograph coupled to an Agilent 5975C mass selective detector (Agilent Technologies, Waldbronn, Germany). The gas chromatograph was equipped with an Agilent HP-5MS Ultra Inert (30 m × 0.25 mm i.d., 0.25 μm film thickness) column. Helium was used as the carrier gas and maintained at a constant flow rate of 1.1 mL $min^{-1}$; the initial pressure was set at 56.07 kPa. The GC oven temperature was 40°C for 2 min after injection, and then programmed at 60 ºC $min^{-1}$ to 150 ºC, which was maintained for 2 min, and next ramped again at 30 ºC $min^{-1}$ to the final temperature of 280 ºC, which was held for 7 min. The total run time was 17.17 min.

Injections were carried out using the MultiPurpose Sampler MPS 2XL from GERSTEL GmbH & Co. KG (Germany) with a 10-μL syringe. A volume of 10 μL was injected at a controlled speed of 3.5 μL $s^{-1}$; after every injection, two washings of the syringe with acetone were performed. The injection system consisted of a septumless head and a PTV inlet (cooled injection system CIS 6 from GERSTEL GmbH & Co. KG) equipped with an empty multi-baffled deactivated glass liner. The PTV inlet was operated in the solvent-vent mode. During the injection and for 0.5 min after it (vent time), the inlet temperature was held at 50 ºC, whereas the column head pressure was reduced to 7 kPa and the flow rate through the split vent was set at 20 mL/min to purge out most of the solvent. Once the vent time had elapsed, the split valve was closed, and after 0.8 min after injection, the temperature of the inlet was ramped from the initial value of 50 ºC up to 280 ºC at a heating rate of 10.8 ºC $s^{-1}$ in order to transfer the analytes retained in the liner to the GC column head. The split valve was reopened 2 min after injection to purge the inlet at a vent flow of 47.4 mL $min^{-1}$.

Analyses were performed in the electron ionization mode at an electron energy of 70 eV. Data were acquired in selected ion monitoring (SIM) mode after a solvent delay of 10 min. Four acquisition windows for SIM data were designed: *i)* for BPF peak (ion dwell time: 75 ms), with diagnostic ions 94, 107, 152, 183 and 200 (molecular ion and base peak); *ii)* for BPA-$d_{16}$ and BPA peaks (ion dwell time: 40 ms), where the diagnostic ions for BPA-$d_{16}$ were 125, 224 (base peak), 242 and 244 (molecular ion), and the diagnostic ions for BPA were 91, 107, 119, 213 (base peak) and 228 (molecular ion); *iii)* for BFDGE peak (ion dwell

time: 75 ms), with diagnostic ions 107, 181, 197 (base peak), 279 and 312 (molecular ion); and *iv)* for BADGE peak (ion dwell time: 75 ms), with diagnostic ions 91, 119, 169, 325 (base peak) and 340 (molecular ion). The electron multiplier operated in relative mode (relative voltage of 300 V). The ion source temperature was fixed at 230 °C, whereas those of the quadrupole and the GC-MS interface were, respectively, 150 and 300 °C.

#### *4.3.4.6. Instrumental*

The temperature during the migration testing was controlled by a water bath equipped with an immersion thermostat Digiterm 200 (JP Selecta S.A., Barcelona, Spain). The CRISON micropH 2002 pH-meter (CRISON Instruments, S.A., Barcelona, Spain) was used for controlling the pH of the samples. The vacuum manifold used for the SPE step was purchased from Waters Corporation (Milford, MA, USA). The negative pressure required for the SPE was achieved by coupling a laboratory vacuum pump (Sartorius AG, Goettingen, Germany) to the manifold. SPE cartridges Oasis® HLB 200 mg/6 cc (Waters Corporation, Milford, MA, USA) were employed. The evaporation of the solvent in the eluates from the SPE cartridges was performed in a miVac Modular Concentrator (GeneVac Limited, Ipswich, UK), which consisted of a miVac Duo concentrator, a SpeedTrap™ and a Quattro pump. A ZX3 vortex mixer (VELP Scientifica, Usmate (MB), Italy) was used for homogenizing samples.

#### *4.3.4.7. Software*

MSD ChemStation E.02.01.1177 (Agilent Technologies, Inc.) and GERSTEL MAESTRO 1 (version 1.3.20.41/3.5) were used for acquiring data and subtracting the background signal. PARAFAC decompositions were performed with the PLS_Toolbox 5.0 [32] for use with MATLAB (version 7.9.0.529 (R2009b)). The one-way ANOVA, the LS calibration curves and the migration models were built and validated with STATGRAPHICS Centurion XVI [33]. Decision limit, $CC\alpha$, and capability of detection, $CC\beta$, were determined using the DETARCHI program [34].

**Table 4.3.1.** Summary of the standards and/or samples analysed at every experimental stage and of the three-way tensors built for the identification and quantification of every compound. All solvent standards were prepared in acetone

| Analytical stage | | Standards and/or samples analysed | Concentration range | Dimensions of the resulting data tensor (Scans×Ions×Samples) | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | **BPF** | **BPA-$d_{16}$** | **BPA** | **BFDGE** | **BADGE** |
| pH evaluation | Tolerance intervals | 2 system blanks<br>2 solvent blanks<br>12 solvent standards at 5 concentration levels | 100 μg $L^{-1}$ (in triplicate)<br>200 μg $L^{-1}$ (in triplicate)<br>300 μg $L^{-1}$ (in duplicate)<br>400 μg $L^{-1}$ (in duplicate)<br>500 μg $L^{-1}$ (in duplicate) | 15×5×16 | 8×9×16 | 13×9×16 | 12×5×16 | 15×5×16 |
| | Study | 2 system blanks<br>2 solvent blanks<br>14 samples at 7 different pH values (in duplicate)<br>2 solvent standards | Samples spiked with 24 μg $L^{-1}$ of every analyte (300 μg $L^{-1}$ after reconstitution)<br>Standards with 300 μg $L^{-1}$ of every compound | 12×5×20 | 9×9×20 | 13×9×19 (1 outlier sample rejected) | 15×5×20 | 16×5×20 |
| One-way ANOVA | Tolerance intervals | Same as in "pH evaluation" | | | | | | |
| | Study | 5 system blanks<br>4 solvent blanks<br>18 samples at 3 different pH values (6 replicates each)<br>4 solvent standards | Samples spiked with 24 μg $L^{-1}$ of every analyte (300 μg $L^{-1}$ after reconstitution)<br>Standards with 300 μg $L^{-1}$ of every compound | 15×5×31 | 9×9×31 | 16×9×31 | 12×5×31 | 16×5×31 |

**Table 4.3.1. (cont. (I))**

| Analytical stage | | Standards and/or samples analysed | Concentration range | Dimensions of the resulting data tensor (Scans×Ions×Samples) | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | **BPF** | **BPA-$d_{16}$** | **BPA** | **BFDGE** | **BADGE** |
| Solvent calibration vs. Matrix-matched calibration | Tolerance intervals | Same as in "pH evaluation" | | | | | | |
| | Solvent calibration | 4 system blanks<br>2 solvent blanks<br>11 solvent standards at 6 concentration levels | Calibration range 0–500 μg $L^{-1}$ (IS[a] at 200 μg $L^{-1}$)<br>(Blank and 100 μg $L^{-1}$ in triplicate; 300 μg $L^{-1}$ in duplicate) | 15×5×17 | 9×9×17 | 16×9×17 | 12×5×17 | 16×5×17 |
| | Matrix-matched calibration | 3 system blanks<br>1 solvent blank<br>10 matrix-matched standards at 6 concentration levels | Calibration range 0–500 μg $L^{-1}$ (IS[a] at 200 μg $L^{-1}$)<br>(4 blank replicates and 2 replicates at 300 μg $L^{-1}$) | 15×5×14 | 9×9×14 | 16×9×14 | 12×5×14 | 16×5×14 |
| BPA migration from PC glasses (70 ºC) | Tolerance intervals | Same as in "pH evaluation" | | | | | | |
| | Study (migration tests 1–7) | 14 system blanks<br>7 solvent blanks<br>14 solvent standards at 11 concentration levels of every analyte<br>42 non-spiked migration samples from tests 1–7 (2 samples per PC glass)<br>22 spiked migration samples from tests 4–7 (2 samples per PC glass, except for glass number 1 at test 6) | Calibration range 0–250 μg $L^{-1}$ (IS[a] at 200 μg $L^{-1}$)<br>(2 blank replicates and 3 replicates at 100 μg $L^{-1}$)<br>Spiked samples containing 800 ng $L^{-1}$ of every analyte (100 μg $L^{-1}$ after reconstitution) | 21×5×99 | 13×9×99 | 14×9×99 | 12×5×99 | 18×5×99 |

**Table 4.3.1. (cont. (II))**

| Analytical stage | | Standards and/or samples analysed | Concentration range | Dimensions of the resulting data tensor (Scans×Ions×Samples) | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | **BPF** | **BPA-$d_{16}$** | **BPA** | **BFDGE** | **BADGE** |
| BPA migration from PC glasses (50 ºC) | Tolerance intervals | 2 system blanks<br>8 solvent standards at 4 concentration levels (in duplicate) | 50, 100, 150, 200 μg $L^{-1}$ | — | 8×9×10 | 14×9×10 | — | — |
| | Study (migration tests 1–3) | 14 system blanks<br>5 solvent blanks<br>6 solvent standards at 6 concentration levels of BPA<br>18 migration samples from tests 1–3 (2 samples per PC glass) | Calibration range 0–40 μg $L^{-1}$ (IS[a] at 200 μg $L^{-1}$) | — | 13×9×43 | 14×9×43 | — | — |
| | Study (migration tests 4–7) | 19 system blanks<br>9 solvent blanks<br>6 solvent standards at 6 concentration levels of BPA<br>24 migration samples from tests 4–7 (2 samples per PC glass) | Calibration range 0–40 μg $L^{-1}$ (IS[a] at 200 μg $L^{-1}$) | — | 13×9×58 | 14×9×58 | — | — |

([a]) Internal standard (BPA-$d_{16}$)

## 4.3.5. Results and discussion

### *4.3.5.1. Optimization of the experimental method*

#### 4.3.5.1.1. Testing the effect of pH on BPF, BPA, BFDGE and BADGE

Guaranteeing the unequivocal identification of every compound is mandatory before whatever quantification in accordance with the confirmatory criteria laid down in [24]. So, standards with concentrations of 100, 200, 300, 400 and 500 μg $L^{-1}$ of BPF, BPA-$d_{16}$, BPA, BFDGE and BADGE in acetone were first prepared and analysed, as shown in Table 4.3.1 (first row, related to the tolerance intervals in the pH evaluation stage). Taking advantage of the trilinearity of GC-MS data, the PARAFAC decompositions from the five resulting tensors enabled the extraction of the unique chromatographic and spectral profiles of every compound. This led to the estimation of the tolerance intervals for the relative ion abundances and for the relative retention time of every substance from the loadings of its spectral and chromatographic profiles, respectively. Those intervals, which are listed in Table 4.3.2, were used as reference to confirm the presence of the corresponding compound in every sample.

Different pH values between 2 and 12 were considered for the evaluation of the effect of this variable on the recovery of BPF, BPA, BFDGE and BADGE after the pH adjustment step of the method, so samples were not subjected to SPE afterwards. The aim of these tests was to check if the structures of the four analytes stayed unaltered after the corresponding change in the pH of the sample. To that end, 5 mL of simulant *D1* were fortified with 24 μg $L^{-1}$ of BPF, BPA, BFDGE and BADGE; then the pH of every sample was adjusted by adding the smallest possible volume of a solution of either hydrochloric acid or sodium hydroxide. The pH values tested in duplicate were specifically 2, 4, 5.6 (pH of freshly prepared simulant *D1*), 6, 8, 10 and 12. Next, the 14 samples were reconstituted after evaporation to dryness in 400 μL of 300 μg $L^{-1}$ of BPA-$d_{16}$ in acetone prior to analysis. Two standards at 300 μg $L^{-1}$ of the five compounds in acetone were also prepared and analysed. In addition, both system blanks (no liquid) and solvent blanks (only acetone) were injected to test the performance of the GC-MS system. Five three-way tensors containing the data from the analyses of all the samples, the blanks and the standards were thus built. The dimensions of these tensors are specified in Table 4.3.1 (second row, related to the study carried out in the pH evaluation stage, columns 5 to 9).

**Table 4.3.2.** Tolerance intervals for the relative retention time and for the diagnostic ions of every compound estimated from the PARAFAC decomposition of the corresponding tensor (see Table 4.3.1, second row). The base peak in each case is in bold.

| Compound | $t_{R,rel}$ | Tolerance interval | $m/z$ ratio | Spectral loading $L_i$ | Relative abundance ($L_i$ / $L_{BP}$) (%) | Tolerance interval (%) |
|---|---|---|---|---|---|---|
| BPF (15×5×16 tensor) (1-factor unconstrained model, explained variance of 94.26%) | 0.966 | (0.961 – 0.971) | 94 | $1.21 \cdot 10^{-1}$ | 17.45 | (13.96 – 20.94) |
| | | | 107 | $6.88 \cdot 10^{-1}$ | 99.40 | (89.46 – 109.34) |
| | | | 152 | $1.09 \cdot 10^{-1}$ | 15.72 | (12.58 – 18.87) |
| | | | 183 | $1.46 \cdot 10^{-1}$ | 21.11 | (17.94 – 24.27) |
| | | | **200** | $6.92 \cdot 10^{-1}$ | **100** | — |
| BPA-$d_{16}$ (8×9×16 tensor) (1-factor unconstrained model, explained variance of 99.51%) | 1.000 | — | 125 | $1.37 \cdot 10^{-1}$ | 14.06 | (11.25 – 16.88) |
| | | | **224** | $9.74 \cdot 10^{-1}$ | **100** | — |
| | | | 242 | $1.82 \cdot 10^{-1}$ | 18.65 | (14.92 – 22.38) |
| | | | 244 | $3.09 \cdot 10^{-3}$ | 0.32 | (0.16 – 0.48) |
| BPA (13×9×16 tensor) (1-factor unconstrained model, explained variance of 98.78%) | 1.003 | (0.998 – 1.008) | 91 | $1.06 \cdot 10^{-1}$ | 11.03 | (8.83 – 13.24) |
| | | | 107 | $5.42 \cdot 10^{-2}$ | 5.65 | (2.83 – 8.48) |
| | | | 119 | $1.76 \cdot 10^{-1}$ | 18.39 | (14.71 – 22.06) |
| | | | **213** | $9.57 \cdot 10^{-1}$ | **100** | — |
| | | | 228 | $1.85 \cdot 10^{-1}$ | 19.32 | (15.46 – 23.19) |
| BFDGE (12×5×16 tensor) (2-factor unconstrained model, CI[a] 100, explained variance of 99.06%) | 1.224 | (1.218 – 1.230) | 107 | $2.45 \cdot 10^{-1}$ | 33.29 | (28.30 – 38.29) |
| | | | 181 | $5.04 \cdot 10^{-1}$ | 68.55 | (61.69 – 75.40) |
| | | | **197** | $7.35 \cdot 10^{-1}$ | **100** | — |
| | | | 279 | $2.07 \cdot 10^{-2}$ | 2.82 | (1.41 – 4.23) |
| | | | 312 | $3.83 \cdot 10^{-1}$ | 52.14 | (46.93 – 57.35) |

**Table 4.3.2. (cont.)**

| Compound | $t_{R,rel}$ | Tolerance interval | $m/z$ ratio | Spectral loading $L_i$ | Relative abundance ($L_i / L_{BP}$) (%) | Tolerance interval (%) |
|---|---|---|---|---|---|---|
| BADGE (15×5×16 tensor) (1-factor unconstrained model, explained variance of 99.77%) | 1.397 | (1.390 – 1.404) | 91 | $4.68 \cdot 10^{-2}$ | 4.75 | (2.37 – 7.12) |
| | | | 119 | $6.36 \cdot 10^{-2}$ | 6.45 | (3.22 – 9.67) |
| | | | 169 | $2.81 \cdot 10^{-2}$ | 2.85 | (1.43 – 4.28) |
| | | | **325** | $9.86 \cdot 10^{-1}$ | **100** | — |
| | | | 340 | $1.43 \cdot 10^{-1}$ | 14.52 | (11.61 – 17.42) |

(a) CI stands for CORCONDIA index

The PARAFAC decomposition of the tensor constructed for BPF yielded a two-factor model (CORCONDIA of 100, explained variance of 95.56%, no outlier samples detected), where the chromatographic and spectral ways had been non-negativity-constrained. The first factor in this model (in dark blue in Figures 4.3.1.a, 4.3.1.b and 4.3.1.c) was unequivocally associated to BPF, whereas the second one (in light green in Figures 4.3.1.a, 4.3.1.b and 4.3.1.c) was to an unidentified coeluent that only appeared in the samples at pH 2, especially in the first replicate. As for the tensor for BPA-$d_{16}$, a two-factor model (CORCONDIA of 100, explained variance of 98.90%, no outliers found) resulted after no constraints had been imposed. As Figures 4.3.1.d and 4.3.1.e show, the first factor (in dark blue) matched BPA-$d_{16}$ unequivocally; the second factor was attributed to an interferent eluting near the beginning of the BPA-$d_{16}$ peak. The unconstrained PARAFAC model estimated from the BPA tensor indicated the existence of one single factor corresponding to BPA (explained variance of 99.09%, 1 outlier rejected). On the other hand, the PARAFAC decomposition of the BFDGE tensor revealed, through a two-factor unconstrained model (CORCONDIA of 100, explained variance of 99.05%, no outliers detected), the coelution of an interferent with BFDGE, mainly present in the samples whose pH ranged between 6 and 12; however, that did not hinder the unequivocal identification of this analyte (see Figures 4.3.1.f, 4.3.1.g and 4.3.1.h). Lastly, in the case of BADGE, a one-factor unconstrained model was obtained (explained variance of 99.40%, no outliers found), being this factor coherent with this compound.

Checking the presence of a compound in some of those 14 samples by only considering the mass spectrum recorded at its retention time would lead to a false negative when comparing the relative abundances of its diagnostic ions with the tolerance intervals. An example of this is presented in Table 4.3.3 for BPF and BFDGE. The identification from the PARAFAC decompositions above ensured the presence of both BPF and BFDGE in all the samples and the standards. Instead, if each of these two analytes had been evaluated from the spectral data obtained at its retention time in the sample specified in Table 4.3.3, it would have been wrongly concluded that BPF and BFDGE were missing in that sample, since the relative abundances of several diagnostic ions would not lie within their tolerance intervals (see columns number 8 and 9 in Table 4.3.3): there would not have been enough well-characterized ions to guarantee the unequivocal identification of the analyte in the sample as it is set out in [24].

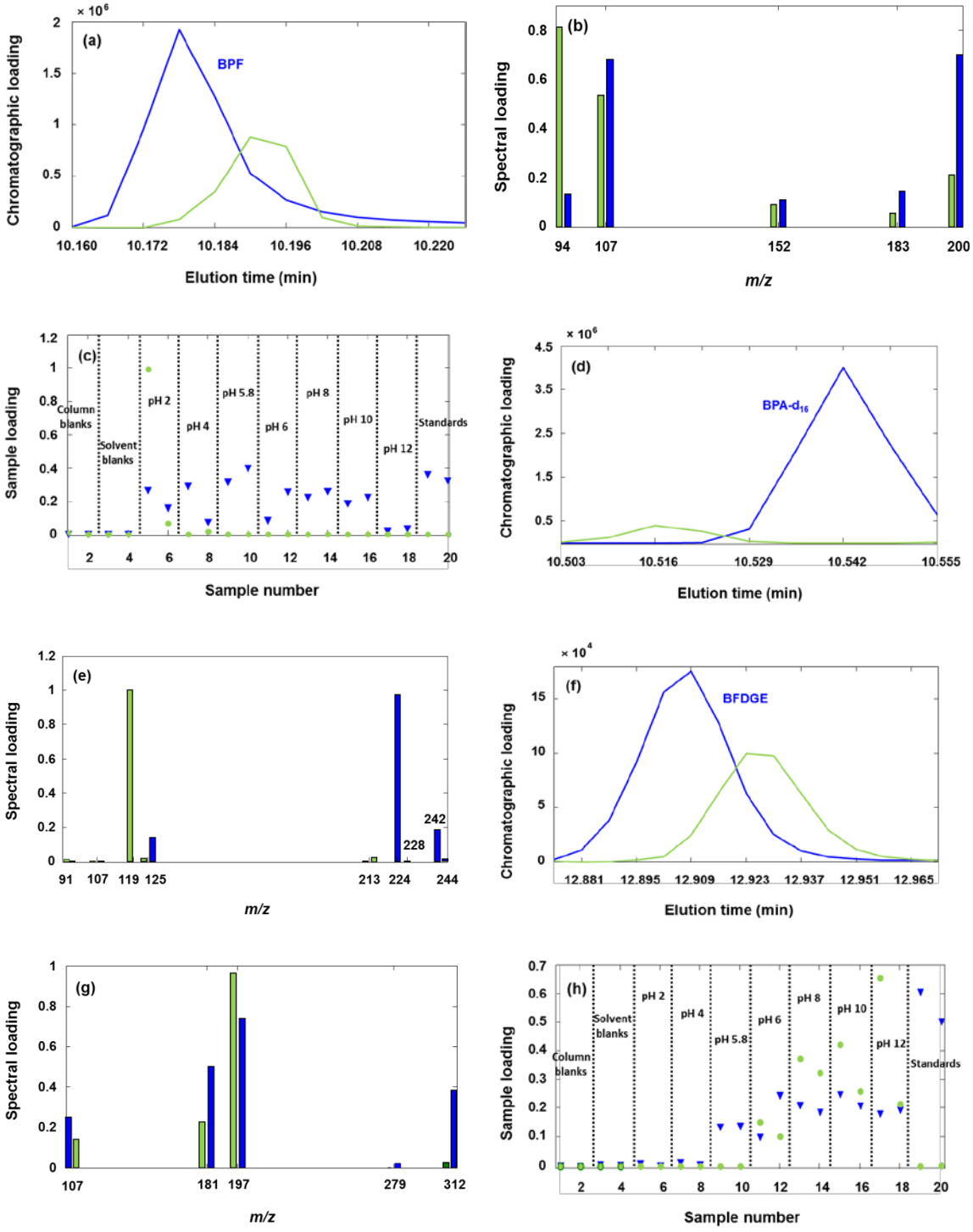


**Figure 4.3.1.** PARAFAC decompositions of the tensors that revealed the presence of coeluents obtained at the pH evaluation stage. Loadings of the **(a)** chromatographic, **(b)** spectral and **(c)** sample profiles of the PARAFAC model for BPF; Factor 1 (BPF): dark blue; Factor 2: light green. Loadings of the **(d)** chromatographic and **(e)** spectral profiles of the PARAFAC model for BPA-$d_{16}$; Factor 1 (BPA-$d_{16}$): dark blue; Factor 2: light green. Loadings of the **(f)** chromatographic, **(g)** spectral and **(h)** sample profiles of the PARAFAC model for BFDGE; Factor 1 (BFDGE): dark blue; Factor 2: light green.

**Table 4.3.3.** Comparison of the results obtained for the spectral identification of BPF and BFDGE in the pH evaluation study depending on whether the PARAFAC methodology or just the mass spectrum at the retention time is used for that purpose. The base peak in both cases is in bold.

| Analyte | *m/z* ratio | Tolerance interval (%) | Identification based on the PARAFAC decomposition of the data tensor including all samples | | | Identification based on the mass spectrum at the retention time | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Relative abundance ($L_i / L_{BP}$) (%) | Compliance verified | Result | Sample | Relative abundance (%) | Compliance verified | Result |
| BPF | 94 | (13.96 – 20.94) | 18.77 | Yes | Correct identification: analyte present in the 14 samples and the 2 standards | First replicate at pH 2 | 36.31 | No | Wrong identification: analyte missing in the sample |
| | 107 | (89.46 – 109.34) | 97.20 | Yes | | | 167.80 | No | |
| | 152 | (12.58 – 18.87) | 15.27 | Yes | | | 22.81 | No | |
| | 183 | (17.94 – 24.27) | 20.51 | Yes | | | 25.48 | No | |
| | **200** | — | **100** | — | | | **100** | — | |
| BFDGE | 107 | (28.30 – 38.29) | 33.77 | Yes | Correct identification: analyte present in the 14 samples and the 2 standards | First replicate at pH 8 | 28.35 | Yes | Wrong identification: analyte missing in the sample |
| | 181 | (61.69 – 75.40) | 67.74 | Yes | | | 54.40 | No | |
| | **197** | — | **100** | — | | | **100** | — | |
| | 279 | (1.41 – 4.23) | 2.91 | Yes | | | 2.15 | Yes | |
| | 312 | (46.93 – 57.35) | 51.85 | Yes | | | 36.41 | No | |

Once the identification step had been completed, the standardized sample loadings for each of the four analytes at the seven pH levels tested were determined; these values are gathered in Table 4.3.4. For every analyte, the higher the standardized sample loading, the greater the amount of that compound in that sample. The low values for BPF and BPA at pH 12 suggested a possible loss of both substances in the form of the corresponding bisphenolate salt. In the same way, the poor results obtained for BFDGE and BADGE at pH 2 and 4 were consistent with possible acid-catalyzed ring-opening reactions of the epoxide groups to yield glycols. On the other hand, the datum for BADGE at pH 10 was considered an outlier.

**Table 4.3.4.** Average standardized sample loadings ($n$ = 2 replicates) obtained for every analyte at the seven pH values evaluated.

| pH value | Average standardized sample loading | | | |
|---|---|---|---|---|
| | **BPF** | **BPA** | **BFDGE** | **BADGE** |
| 2 | 0.80 | 1.02 | 0.02 | 0.03 |
| 4 | 1.23 | 1.44 | 0.07 | 0.05 |
| 5.6 | 1.13 | 1.19 | 0.43 | 0.23 |
| 6 | 1.00 | 1.13 | 1.01 | 0.42 |
| 8 | 1.11 | 1.16 | 0.91 | 0.50 |
| 10 | 1.01 | 1.21 | 1.10 | 2.63 |
| 12 | 0.17 | 0.31 | 0.90 | 0.57 |

The pH range 8-10 showed the best possible standardized sample loadings for a simultaneous determination of the four analytes. This result agreed with the expectations reflected in Section 4.3.2. That does not necessarily mean that the ring-opening of epoxides could not have occurred in this pH range. Nucleophilic substitution reactions are indeed likely to have taken place on the epoxide groups of BFDGE and BADGE at these conditions, but at a reaction rate far slower than that in very acidic media (pH 2–4), where that process would have been practically immediate. This fact would account for the difference between the standardized sample loadings of BFDGE and BADGE obtained at pH 2–6 and those at pH 6–12. The final value for the pH adjustment step in the experimental method was selected through a one-way analysis of variance (ANOVA) for the factor "pH" to three levels (8, 9 and 10), where the response variable was the standardized sample loading of every analyte.

4.3.5.1.2. One-way ANOVA for the selection of the optimum pH

According to the conclusion in the previous section, six replicates at each of the three pH levels were prepared to perform the ANOVA following the same procedure as in Section 4.3.5.1.1. After GC-MS analysis, five data tensors were built from these 18 samples, together with solvent and system blanks and 300 μg $L^{-1}$ standards, as it appears in Table 4.3.1 (fourth row, referred to the evaluation of the one-way ANOVA).

All PARAFAC decompositions resulted in one-factor unconstrained models, with no outlier samples detected at a 99% confidence level. In each case, the percentage of variance explained by the three-way model was 94.20% for BPF, 99.53% for BPA-$d_{16}$, 98.66% for BPA, 98.89% for BFDGE and 99.59% for BADGE. As indicated in the third row of Table 4.3.1, the presence of every compound was unequivocally verified as previously (see Table 4.3.2).

Table 4.3.5 shows the value of the standardized sample loading estimated for each of the 18 samples analysed, together with some results of the ANOVA. The hypotheses of normal distribution of the residuals and of equality of variances were confirmed for the four analytes at a 95% confidence level. On the other hand, testing the equality of the means obtained at the three pH levels led to the conclusion that, for both BPF and BFDGE, there was no effect on the response variable due to the factor "pH" at a 5% significance level, whereas that effect was significant in the case of BPA and BADGE. It must be emphasized that the ANOVA test produced the same results for every bisphenol and its derivative.

Since the null hypothesis in the ANOVA test had to be rejected for BPA and BADGE, both Tukey HSD and Student-Newman-Keuls multiple range tests were posed to look for significant differences among the means of the three pH levels. As can be seen in Table 4.3.5, for both analytes, adjusting the pH to either 8 or 9 enabled to get equal values of the standardized sample loading that were significantly different from those obtained at pH 10 at a 95% confidence level. Identical results were achieved for BPA and its epoxy derivative BADGE again.

Taking the values of the coefficient of variation into account (see Table 4.3.5), pH 9 should be selected for both BPF and BPA, whereas pH 8 should be for BFDGE and BADGE. Although either pH 8 or 9 would have been just as appropriate for the simultaneous determination of the four analytes in accordance with the result of the multiple range tests posed for BPA and BADGE, after comparing the coefficients of variation at those two factor levels, it was realized that the best option for BPF and not so bad for the others was 9, being the highest coefficient of variation at this pH equal to 10.45%.

**Table 4.3.5.** Results of the one-way ANOVA posed to test for differences in the standardized sample loading of every analyte among three pH values (8, 9, 10). Coefficient of variation in % for each pH level. *p*-values for some preliminary tests for homoscedasticity and residual normality (Kolmogoroff-Smirnoff (K-S) and chi-square ($\chi^2$) tests), the ANOVA *F*-test. Conclusions from the Tukey HSD and Student-Newman-Keuls multiple range tests.

| Analyte | Replic. | Standardized sample loading | | | Coefficient of variation (%) | | | *p*-values | | | | | Multiple range tests ($\alpha = 0.05$) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Preliminary tests | | | | One-way ANOVA[c] | | |
| | | Factor level | | | Factor level | | | Normality of residuals | | Homoscedasticity of the response variable | | | | |
| | | pH 8 | pH 9 | pH 10 | pH 8 | pH 9 | pH 10 | K-S test[a] | $\chi^2$ test[a] | Cochran's test[b] | Bartlett's test[b] | | Tukey HSD test | Student-Newman-Keuls test |
| BPF | 1 | 1.01 | 1.18 | 1.19 | 14.28 | 7.88 | 20.22 | 0.83 | 0.12 | 0.26 | 0.24 | 0.45 | — | — |
| | 2 | 1.11 | 0.94 | 0.85 | | | | | | | | | | |
| | 3 | 0.87 | 1.02 | 1.09 | | | | | | | | | | |
| | 4 | 1.16 | 1.02 | 0.91 | | | | | | | | | | |
| | 5 | 0.95 | 0.98 | 0.69 | | | | | | | | | | |
| | 6 | 0.79 | 1.05 | 0.81 | | | | | | | | | | |
| BPA | 1 | 1.01 | 1.11 | 1.19 | 9.44 | 6.48 | 9.84 | 0.67 | 0.15 | 0.31 | 0.43 | $1 \cdot 10^{-4}$ | pH 8 equal to pH 9 and different from pH 10 | pH 8 equal to pH 9 and different from pH 10 |
| | 2 | 0.80 | 0.95 | 1.18 | | | | | | | | | | |
| | 3 | 0.95 | 0.99 | 1.03 | | | | | | | | | | |
| | 4 | 0.87 | 0.94 | 1.20 | | | | | | | | | | |
| | 5 | 0.93 | 0.96 | 1.23 | | | | | | | | | | |
| | 6 | 0.80 | 1.00 | 1.40 | | | | | | | | | | |

**Table 4.3.5. (cont.)**

| Analyte | Replic. | Standardized sample loading | | | Coefficient of variation (%) | | | *p*-values | | | | | Multiple range tests ($\alpha = 0.05$) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Preliminary tests | | | | One-way ANOVA[c] | | |
| | | Factor level | | | Factor level | | | Normality of residuals | | Homoscedasticity of the response variable | | | | |
| | | pH 8 | pH 9 | pH 10 | pH 8 | pH 9 | pH 10 | K-S test[a] | $\chi^2$ test[a] | Cochran's test[b] | Bartlett's test[b] | | Tukey HSD test | Student-Newman-Keuls test |
| BFDGE | 1 | 1.05 | 1.08 | 0.87 | 7.48 | 9.80 | 4.10 | 0.94 | 0.86 | 0.22 | 0.16 | 0.15 | — | — |
| | 2 | 0.90 | 0.87 | 0.83 | | | | | | | | | | |
| | 3 | 0.92 | 0.97 | 0.88 | | | | | | | | | | |
| | 4 | 0.89 | 1.04 | 0.88 | | | | | | | | | | |
| | 5 | 0.99 | 0.84 | 0.86 | | | | | | | | | | |
| | 6 | 0.86 | 0.94 | 0.93 | | | | | | | | | | |
| BADGE | 1 | 1.10 | 1.20 | 1.05 | 7.98 | 10.45 | 22.72 | 0.20 | 0.20 | 0.17 | 0.27 | 0.002 | pH 8 equal to pH 9 and different from pH 10 | pH 8 equal to pH 9 and different from pH 10 |
| | 2 | 1.04 | 0.94 | 0.68 | | | | | | | | | | |
| | 3 | 0.90 | 0.98 | 0.68 | | | | | | | | | | |
| | 4 | 0.94 | 0.89 | 0.67 | | | | | | | | | | |
| | 5 | 0.95 | 1.04 | 0.66 | | | | | | | | | | |
| | 6 | 0.92 | 1.03 | 0.59 | | | | | | | | | | |

(a) Null hypothesis, $H_0$: The residuals are compatible with a normal distribution.

(b) Null hypothesis, $H_0$: The variances at all pH levels are equal.

(c) Null hypothesis, $H_0$: There is no effect due to pH.

#### *4.3.5.2. Solvent calibration versus Matrix-matched calibration*

To test the suitability of either a solvent calibration or a matrix-matched calibration to quantify the four analytes, using BPA-$d_{16}$ as internal standard, both calibration sets were designed as indicated in Table 4.3.1 (rows number 6 and 7, respectively). Every matrix-matched standard was prepared by spiking 5 mL of simulant *D1* with the corresponding amount of the four analytes, so that, after pH adjustment to 9, SPE as explained in Section 4.3.4.4, evaporation to dryness and reconstitution in 400 μL of BPA-$d_{16}$ (200 μg $L^{-1}$), the final concentration of the sample fell within the range from 100 to 500 μg $L^{-1}$ of every analyte.

After data processing, two tensors were built for every compound, one for each calibration methodology. The dimensions of these three-way arrays are given in Table 4.3.1 (rows number 6 and 7, columns number 5 to 9) while the features of the model estimated from the PARAFAC decomposition of every tensor are listed in Table 4.3.6. The results regarding the chromatographic and spectral identification of every compound also appear in this last table (see columns number 3 and 7); from the comparison of those values with the tolerance intervals used as reference (see Table 4.3.2), it could be stated that all compounds were unequivocally identified.

Table 4.3.6 also collects the equations of the two LS curves "*Standardized sample loading versus True concentration*" estimated for every analyte, together with that of the accuracy line derived from each one of them. The outliers detected in some cases had studentized residuals greater than 3 in absolute value, so they were removed and a new LS fitting was performed and validated with the remaining data. All the calibration curves followed a quadratic model, both in solvent and in matrix-matched quantifications, and trueness was verified in all cases at a 95% confidence level from the accuracy lines.

As explained in Section 4.3.3.2, the statistical equality of the two calibration models was checked for every analyte by fitting the model in Eq. (4.3.2) to the experimental data from both sets. The hypothesis test on the significance of the estimated coefficient of every term including the dummy variable $z$ was then posed, where $H_0$: $a_i = 0$ versus $H_a$: $a_i \neq 0$; the resultant *p*-values figure in the last column of Table 4.3.6. For every analyte, at a 99 % confidence level, all the estimates of $a_0$, $a_1$ and $a_{11}$ were significantly null, so both calibration methodologies were considered to produce the same statistical results and no matrix effect was thus observed. Due to economic and time considerations, a solvent calibration would be preferred for the quantification of the analytes in the next stage of the work.

**Table 4.3.6.** Comparison between a solvent calibration and a matrix-matched calibration for the quantification of BPF, BPA, BFDGE and BADGE. The 3rd and 7th columns (**Identification**) show the relative retention time and the relative abundances (in brackets) estimated for the diagnostic ions of every compound from the spectral loadings of the PARAFAC models.

| Comp. | Solvent calibration | | | | Matrix-matched calibration | | | | *p*-values[d] |
|---|---|---|---|---|---|---|---|---|---|
| | **PARAFAC model (Outliers)** | **Identification** | ***SSL*[a] = f($c_{true}$) ($R^2$; $s_{yx}$) (Outliers)** | **$c_{pred} = b_0 + b_1\, c_{true}$ ($R^2$; $s_{yx}$)** | **PARAFAC model (Outliers)** | **Identification** | ***SSL*[a] = f($c_{true}$) ($R^2$; $s_{yx}$) (Outliers)** | **$c_{pred} = b_0 + b_1\, c_{true}$ ($R^2$; $s_{yx}$)** | |
| BPF | 1 factor<br>No constraints<br>95.55% (EV[b])<br>(0/17) | $t_{R,rel}$ = 0.966<br>94 (18.02%)<br>107 (101.92%)<br>152 (16.27%)<br>183 (21.17%)<br>200 (100%) | $y = -1.7\ 10^{-2} - 3.7\ 10^{-5}x + 6.7\ 10^{-6}x^2$<br>(99.30%; 0.05)<br>(0/11) | $y = 22.85 + 0.92x$<br>(96.78%; 30.53) | 1 factor<br>No constraints<br>94.21% (EV)<br>(0/14) | $t_{R,rel}$ = 0.966<br>94 (17.83%)<br>107 (100.85%)<br>152 (15.85%)<br>183 (21.10%)<br>200 (100%) | $y = -1.0\ 10^{-3} + 1.1\ 10^{-3}x + 5.2\ 10^{-6}x^2$<br>(99.90%; 0.02)<br>(0/10) | $y = 1.40 + 0.99x$<br>(99.90%; 7.75) | 0.60 ($a_0$)<br>0.01 ($a_1$)<br>0.06 ($a_{11}$) |
| BPA-$d_{16}$ | 1 factor<br>No constraints<br>98.68% (EV)<br>(0/17) | $t_{R,rel}$ = 1.000<br>125 (15.45%)<br>224 (100%)<br>242 (18.52%)<br>244 (0.31%) | Internal standard | | 3 factors<br>Factor 1 = BPA-$d_{16}$<br>Non-negativity constraint in modes 1 and 2<br>99.19% (EV)<br>CI[c] 88<br>(0/14) | $t_{R,rel}$ = 1.000<br>125 (13.71%)<br>224 (100%)<br>242 (18.88%)<br>244 (0.31%) | Internal standard | | |
| BPA | 1 factor<br>No constraints<br>99.06% (EV)<br>(0/17) | $t_{R,rel}$ = 1.003<br>91 (11.43%)<br>107 (6.09%)<br>119 (18.71%)<br>213 (100%)<br>228 (19.24%) | $y = 4.8\ 10^{-4} + 1.9\ 10^{-3}x + 2.1\ 10^{-6}x^2$<br>(99.79%; 0.03)<br>(1/11) | $y = 0.15 + 0.99x$<br>(99.79%; 8.78) | 1 factor<br>No constraints<br>99.18% (EV)<br>(0/14) | $t_{R,rel}$ = 1.002<br>91 (10.85%)<br>107 (5.48%)<br>119 (18.43%)<br>213 (100%)<br>228 (19.45%) | $y = 9.3\ 10^{-3} + 2.5\ 10^{-3}x + 1.8\ 10^{-6}x^2$<br>(99.50%; 0.05)<br>(0/10) | $y = 0.12 + 0.99x$<br>(99.50%; 13.95) | 0.79 ($a_0$)<br>0.14 ($a_1$)<br>0.66 ($a_{11}$) |

**Table 4.3.6. (cont.)**

| Comp. | Solvent calibration | | | | Matrix-matched calibration | | | | *p*-values[d] |
|---|---|---|---|---|---|---|---|---|---|
| | **PARAFAC model (Outliers)** | **Identification** | ***SSL*[a] = f($c_{true}$) ($R^2$; $s_{yx}$) (Outliers)** | **$c_{pred} = b_0 + b_1\, c_{true}$ ($R^2$; $s_{yx}$)** | **PARAFAC model (Outliers)** | **Identification** | ***SSL*[a] = f($c_{true}$) ($R^2$; $s_{yx}$) (Outliers)** | **$c_{pred} = b_0 + b_1\, c_{true}$ ($R^2$; $s_{yx}$)** | |
| BFDGE | 1 factor<br>No constraints<br>98.51% (EV)<br>(0/17) | $t_{R,rel}$ = 1.224<br>107 (31.41%)<br>181 (68.38%)<br>197 (100%)<br>279 (3.54%)<br>312 (50.83%) | $y = 1.5\ 10^{-2} + 1.8\ 10^{-3}x + 2.2\ 10^{-6}x^2$<br>(99.27%; 0.05)<br>(1/11) | $y = -0.58 + 1.00x$<br>(99.27%; 14.56) | 2 factors<br>Factor 1 = BFDGE<br>No constraints<br>98.94% (EV)<br>CI 100<br>(0/14) | $t_{R,rel}$ = 1.224<br>107 (31.46%)<br>181 (68.47%)<br>197 (100%)<br>279 (3.04%)<br>312 (52.50%) | $y = -5.5\ 10^{-3} + 2.3\ 10^{-3}x + 1.8\ 10^{-6}x^2$<br>(99.56%; 0.04)<br>(2/10) | $y = 0.12 + 1.00x$<br>(99.56%; 13.27) | 0.57 ($a_0$)<br>0.37 ($a_1$)<br>0.67 ($a_{11}$) |
| BADGE | 1 factor<br>No constraints<br>98.89% (EV)<br>(0/17) | $t_{R,rel}$ = 1.394<br>91 (4.87%)<br>119 (6.49%)<br>169 (2.92%)<br>325 (100%)<br>340 (14.19%) | $y = 2.2\ 10^{-2} + 1.9\ 10^{-3}x + 2.0\ 10^{-6}x^2$<br>(99.54%; 0.04)<br>(2/11) | $y = -0.58 + 1.00x$<br>(99.54%; 12.40) | 1 factor<br>No constraints<br>99.63% (EV)<br>(0/14) | $t_{R,rel}$ = 1.397<br>91 (4.74%)<br>119 (6.88%)<br>169 (2.95%)<br>325 (100%)<br>340 (14.67%) | $y = 5.1\ 10^{-3} + 2.6\ 10^{-3}x + 1.2\ 10^{-6}x^2$<br>(99.31%; 0.05)<br>(1/10) | $y = -0.06 + 0.97x$<br>(99.31%; 15.60) | 0.66 ($a_0$)<br>0.14 ($a_1$)<br>0.38 ($a_{11}$) |

(a) Standardized sample loading

(b) Explained variance

(c) CI stands for CORCONDIA index

(d) *p*-values for the test on the significance of the estimates of the coefficients of the dummy terms including the binary variable. See Eq. (4.3.2)

#### *4.3.5.3. Evaluation of the migration of BPA from new PC glasses*

Six PC glasses provided with polypropylene lids were purchased at a local store (Burgos, Spain). The evolution of the release of BPA from these containers to food simulant *D1* in time was examined by means of seven migration tests performed on every glass. As commented on Section 4.3.4.3, three glasses (number 1, 2 and 3) were placed in the thermostatic water bath at 70 ºC, while three more glasses (number 4, 5 and 6) were at 50 ºC. The whole migration test procedure and the sample treatment are detailed in Sections 4.3.4.3 and 4.3.4.4.

All the resultant migration samples were analysed. As always, both system and solvent blanks were included in every analytical sequence. Besides, for the 70 ºC kinetics, a solvent calibration set ranging from 0 to 250 μg $L^{-1}$ of every analyte was also analysed to quantitate both the amount of BPA migrated from every glass and the concentration of BPF, BPA, BFDGE and BADGE recovered in the spiked samples. On the other hand, as no spiked samples were prepared from the 50 ºC kinetics, only calibration standards containing BPA in the range of 0–40 μg $L^{-1}$ were needed.

After GC-MS analyses, a total of 99 and 96 samples resulted from the 70 ºC and 50 ºC studies, respectively. These samples were arranged in data tensors whose dimension and composition are specified in Table 4.3.1 (rows number 9 (70 ºC), 11 and 12 (50 ºC); columns number 5 to 9). It can be seen there that one three-way array was built for each of the five compounds from the 70 ºC kinetics, while only data tensors for BPA-$d_{16}$ (internal standard) and BPA were achieved from the 50 ºC migration. Table 4.3.1 also indicates that two tensors were considered for both BPA and BPA-$d_{16}$ in this last study because the GC-MS analysis of the samples from tests 1–3 was carried out using an inlet liner different from that for analysing those from tests 4–7. Both arrays contained the same solvent calibration set.

Next, the PARAFAC decomposition of each of those data tensors was performed. The most remarkable features of every PARAFAC model are collected in the second column of Table 4.3.7 (migration study at 70 ºC) and in the fourth and seventh columns of Table 4.3.8 (migration study at 50 ºC). New reference tolerance intervals were estimated before the 50 ºC evaluation. The need to cut the GC column after the conclusion of the 70 ºC kinetics as a maintenance task prior to starting with the second migration study made new identification standards necessary, since the retention time of every compound would change. So, since only BPA and the internal standard were to be monitored, two tensors were first built (see Table 4.3.1, tenth row, columns number 6 and 7) and decomposed in order to define the tolerance intervals used as chromatographic and spectral reference throughout the 50 ºC kinetics (see Table 4.3.8, third column). On the other hand, in the

case of the 70 ºC samples, the unequivocal identification of every compound was contrasted with the tolerance intervals that had been utilized above (see Table 4.3.2). The presence of every compound in both migration studies was unequivocally confirmed (see Table 4.3.7, third column, and Table 4.3.8, fifth and eighth columns).

**Table 4.3.7.** PARAFAC models estimated at the evaluation of the release of BPA from PC glasses at 70 ºC. Identification of every compound according to the regulations currently in force and results of the LS regressions performed.

| Comp. | PARAFAC model | Identification | $SSL^a = f(c_{true})$ ($R^2$; $s_{yx}$) (Outliers) | $c_{pred} = b_0 + b_1\ c_{true}$ ($R^2$; $s_{yx}$) |
|---|---|---|---|---|
| BPF | 2 factors<br>Factor 1: BPF<br>No constraints<br>94.76% (EV[b])<br>CI[c] 100 | $t_{R,rel}$ = 0.966<br>94 (18.05%)<br>107 (101.02%)<br>152 (15.34%)<br>183 (20.93%)<br>200 (100%) | $y = -2.5\ 10^{-2} + 7.7\ 10^{-3}x$<br>(95.41%; 0.06)<br>(0/5) | $y = -2.5\ 10^{-2} + 0.92x$<br>(95.41%; 8.2) |
| | | | $y = -6.5\ 10^{-1} + 1.8\ 10^{-2}x$<br>(99.22%; 0.10)<br>(0/9) | $y = 1.1\ 10^{-2} + 1.00x$<br>(99.22%; 5.44) |
| $BPA\text{-}d_{16}$ | 2 factors<br>Factor 1: $BPA\text{-}d_{16}$<br>No constraints<br>97.93% (EV)<br>CI 100 | $t_{R,rel}$ = 1.000<br>125 (23.96%)<br>224 (100%)<br>242 (18.90%)<br>244 (0.31%) | Internal standard | |
| BPA | 3 factors<br>Factor 1: BPA<br>Non-negativity constraint in modes 1 and 2<br>98.96% (EV)<br>CI 86 | $t_{R,rel}$ = 1.002<br>91 (8.93%)<br>107 (4.82%)<br>119 (17.00%)<br>213 (100%)<br>228 (19.87%) | $y = -5.3\ 10^{-2} + 1.2\ 10^{-2}x$<br>(99.74%; 0.05)<br>(2/14) | $y = 3.7\ 10^{-3} + 1.00x$<br>(99.74%; 3.85) |
| BFDGE | 2 factors<br>Factor 1: BFDGE<br>No constraints<br>98.06% (EV)<br>CI 100 | $t_{R,rel}$ = 1.224<br>107 (31.29%)<br>181 (67.92%)<br>197 (100%)<br>279 (2.58%)<br>312 (51.88%) | $y = 2.9{\cdot}10^{-2} + 8.4\ 10^{-3}x + 3.3\ 10^{-5}x^2$<br>(99.81%; 0.06)<br>(1/14) | $y = -3.7{\cdot}10^{-1} + 1.00x$<br>(99.81%; 3.30) |
| BADGE | 1 factor<br>No constraints<br>98.56% (EV) | $t_{R,rel}$ = 1.397<br>91 (4.84%)<br>119 (6.77%)<br>169 (2.87%)<br>325 (100%)<br>340 (14.56%) | $y = 2.6{\cdot}10^{-2} + 9.6\ 10^{-3}x + 3.3\ 10^{-5}x^2$<br>(99.75%; 0.06)<br>(2/14) | $y = -1.1\ 10^{-1} + 1.00x$<br>(99.75%; 3.09) |

(a) Standardized sample loading

(b) Explained variance

(c) CI stands for CORCONDIA index

**Table 4.3.8.** PARAFAC models estimated at the evaluation of the release of BPA from PC glasses at 50 ºC. Identification of every compound according to the regulations currently in force and results of the LS regressions performed.

| Comp. | Tolerance intervals | | | Data tensor for tests 1–3 | | | Data tensor for tests 4–7 | | | Detection capability[b] (μg $L^{-1}$) |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | PARAFAC model | Identification | $SSL^a = f(c_{true})$ ($R^2$; $s_{yx}$) (Outliers) $c_{pred} = f(c_{true})$ ($R^2$; $s_{yx}$) | PARAFAC model | Identification | $SSL^a = f(c_{true})$ ($R^2$; $s_{yx}$) (Outliers) $c_{pred} = f(c_{true})$ ($R^2$; $s_{yx}$) | |
| $BPA-d_{16}$ | $t_{R,rel}$ | 1.000 | — | 2 factors Factor 1: $BPA-d_{16}$ Non-negativity constraint in modes 1 and 2 98.65% ($EV^d$) $CI^e$ 100 | $t_{R,rel}$ = 1.000 | Internal standard | 2 factors Factor 1: $BPA-d_{16}$ Non-negativity constraint in modes 1 and 2 99.01% (EV) CI 98 | $t_{R,rel}$ = 1.000 | Internal standard | Internal standard |
| | $m/z$ | 125 | (10.70 – 16.06)% | | 125 (15.15%) | | | 125 (15.00%) | | |
| | | 224 | BP[c] | | 224 (100 %) | | | 224 (100%) | | |
| | | 242 | (15.61 – 23.41)% | | 242 (19.44%) | | | 242 (19.40%) | | |
| | | 244 | (0.16 – 0.48)% | | 244 (0.31%) | | | 244 (0.32%) | | |
| BPA | $t_{R,rel}$ | 1.003 | (0.998 – 1.008) | 2 factors Factor 1: BPA Non-negativity constraint in modes 1 and 2 98.78% (EV) CI 96 | $t_{R,rel}$ = 1.003 | $SSL = -7.3\ 10^{-3} + 1.8\ 10^{-2} c_{true}$ (99.82%; 0.01) (0/6) $c_{pred} = 7.0\ 10^{-3} + 1.00 c_{true}$ (99.82%; 0.70) | 3 factors Factor 1: BPA Non-negativity constraint in modes 1 and 2 99.31% (EV) CI 97 | $t_{R,rel}$ = 1.003 | $SSL = -7.4\ 10^{-3} + 2.1\ 10^{-2} c_{true}$ (99.83%; 0.01) (0/6) $c_{pred} = 1.3\ 10^{-2} + 1.00 c_{true}$ (99.83%; 0.71) | 3.30 ($n$ = 1) 2.60 ($n$ = 2) 2.30 ($n$ = 3) |
| | $m/z$ | 91 | (4.69 – 14.07)% | | 91 (8.56%) | | | 91 (8.68%) | | |
| | | 107 | (2.36 – 7.07)% | | 107 (4.30%) | | | 107 (4.34%) | | |
| | | 119 | (13.72 – 20.58)% | | 119 (16.76%) | | | 119 (16.83%) | | |
| | | 213 | BP[c] | | 213 (100%) | | | 213 (100%) | | |
| | | 228 | (17.08 – 23.11)% | | 228 (20.01%) | | | 228 (19.99%) | | |

(a) Standardized sample loading

(b) $\alpha = \beta = 0.05$

(c) Base peak

(d) Explained variance

(e) CI stands for CORCONDIA index

After the correct identification of the factor related to every target compound had been guaranteed, a LS regression between the standardized sample loadings of every analyte and the concentration of the solvent standards was built. All these calibration curves are presented in Table 4.3.7 (see fourth and fifth columns) and Table 4.3.8 (see sixth and ninth columns), where it is displayed that a quadratic relationship was still established between those two variables for BFDGE and BADGE in the concentration range 0–250 μg $L^{-1}$. A particular situation arose for BPF in the 70 ºC study. Two sections with quite different linear behaviours were found within that calibration range, and two separate models were thus fitted. Regarding the 50 ºC evaluation, as already mentioned, only a calibration for BPA had been designed, but, since the whole set of standards had been included in each of the two arrays built from these migration analyses, two LS regressions of BPA standardized sample loading *versus* true concentration resulted. Every calibration model was validated and accuracy was contrasted using the corresponding regression of predicted concentration versus true concentration. The values of the capability of detection, $CC\beta$, estimated for BPA when $n$ = 1, 2 or 3 replicates, are also listed in the last column of Table 4.3.8, being the probabilities of false positive ($\alpha$) and false negative ($\beta$) fixed at 0.05; the generalization of the procedure to estimate $CC\beta$ from multi-way calibrations can be found in [35].This performance characteristic has hardly varied over time, as the comparison of these results with those in both Section 4.2 and [36] revealed.

The concentration of each analyte in every migration sample was determined from the corresponding regression model. As for the spiked migration samples prepared from the resultant simulants after tests 4, 5, 6 and 7 at 70 ºC, the value of the average recovery rate of 800 ng $L^{-1}$ of every analyte was calculated, being the results, respectively, 92.75% for BPF, 90.62% for BPA, 89.29% for BFDGE and 50.14% for BADGE. These values were significantly better than those achieved in both Section 4.2 and [36], particularly for both BFDGE and BADGE, which gives clear proof of the improvement of the experimental method after the optimization of the pH of the samples. The comparison between the percentage recoveries obtained for the diglycidyl ethers suggests that the ring-opening reaction of the epoxide group could have taken place faster for BADGE than for BFDGE, since the recovery was lower for the former. The quantity of this analyte that could not be recovered might thus have been lost as its hydrolyzed derivatives.

Since BPA is the monomer in the production of polycarbonate plastics, none of the other three analytes was found in the non-spiked migration samples. The levels of BPA migrated from every PC glass over time at each of the two heating temperatures evaluated are listed in Table 4.3.9. The results of the sixth migration test for the glasses number 1 and 4 had to be rejected as outliers for the next interpretations and estimations. It could be concluded from the data shown in Table 4.3.9 that, at either temperature, the three PC containers

evaluated in each case were somewhat differentiated only in the first two migration tests, where it is clear that the three glasses did not release the same amount of BPA per test. The main differences were observed at test number 1, where the effect of temperature on the migration process became patently obvious: the higher the temperature, the greater the quantity of BPA transferred to the food simulant. On the other hand, the concentration of BPA migrated from every glass after each of the last four tests turned out to be the same, no matter which heating temperature had been set at the incubation step. So, although initially the BPA release differed both within every set of containers and for each temperature, the behaviour of all the six glasses tended to be identical in the end. None of the amounts of BPA migrated after every test exceeded the specific migration limit of 0.6 mg $kg^{-1}$ set out for this compound, so the compliance of the six glasses evaluated was ensured.

**Table 4.3.9.** Results of the seven migration tests performed on the six PC glasses evaluated. The symbol "—" denotes an outlier

| Migration test | Concentration of BPA migrated after every test (ng $L^{-1}$) | | | | | |
|---|---|---|---|---|---|---|
| | Heating temperature of 70 ºC | | | Heating temperature of 50 ºC | | |
| | Glass 1 | Glass 2 | Glass 3 | Glass 4 | Glass 5 | Glass 6 |
| 1 | 8210.97 ± 155.88 | 3363.34 ± 1153.16 | 5102.53 ± 529.85 | 1298.98 ± 175.87 | 999.35 ± 181.14 | 961.09 ± 138.31 |
| 2 | 2222.57 ± 89.31 | 2232.52 ± 7.96 | 1673.25 ± 437.39 | 1113.29 ± 63.28 | 820.93 ± 21.09 | 857.61 ± 222.02 |
| 3 | 1975.83 ± 238.29 | 824.03 ± 356.64 | 769.48 ± 376.81 | 726.68 ± 98.62 | 706.68 ± 30.74 | 716.74 ± 10.66 |
| 4 | 722.96 ± 123.84 | 567.79 ± 32.10 | 746.01 ± 124.43 | 509.48 ± 93.23 | 618.91 ± 53.85 | 703.50 ± 31.93 |
| 5 | 369.36 ± 88.70 | 385.53 ± 7.01 | 372.57 ± 18.65 | 431.72 ± 97.95 | 424.80 ± 38.81 | 498.39 ± 10.14 |
| 6 | — | 607.86 ± 1395.88 | 347.21 ± 75.07 | — | 358.11 ± 52.94 | — |
| 7 | 694.69 ± 219.36 | 654.97 ± 35.66 | 536.52 ± 93.76 | 456.99 ± 14.94 | 376.36 ± 74.82 | 524.64 ± 87.31 |

The fitting of the migration data for every container following the exponential model posed in Section 4.3.3.3 is depicted in Figure 4.3.2 and detailed in Table 4.3.10, where the clear difference in the range of estimated values of the parameter *b* with the heating temperature must be noted. This fact reflected the temperature dependence of the BPA migration, so, the higher the temperature, the higher the value of $\hat{b}$ and the greater the average

degradation of the plastic to release this compound; consequently, BPA was transferred at a faster rate at 70 ºC than at 50 ºC.

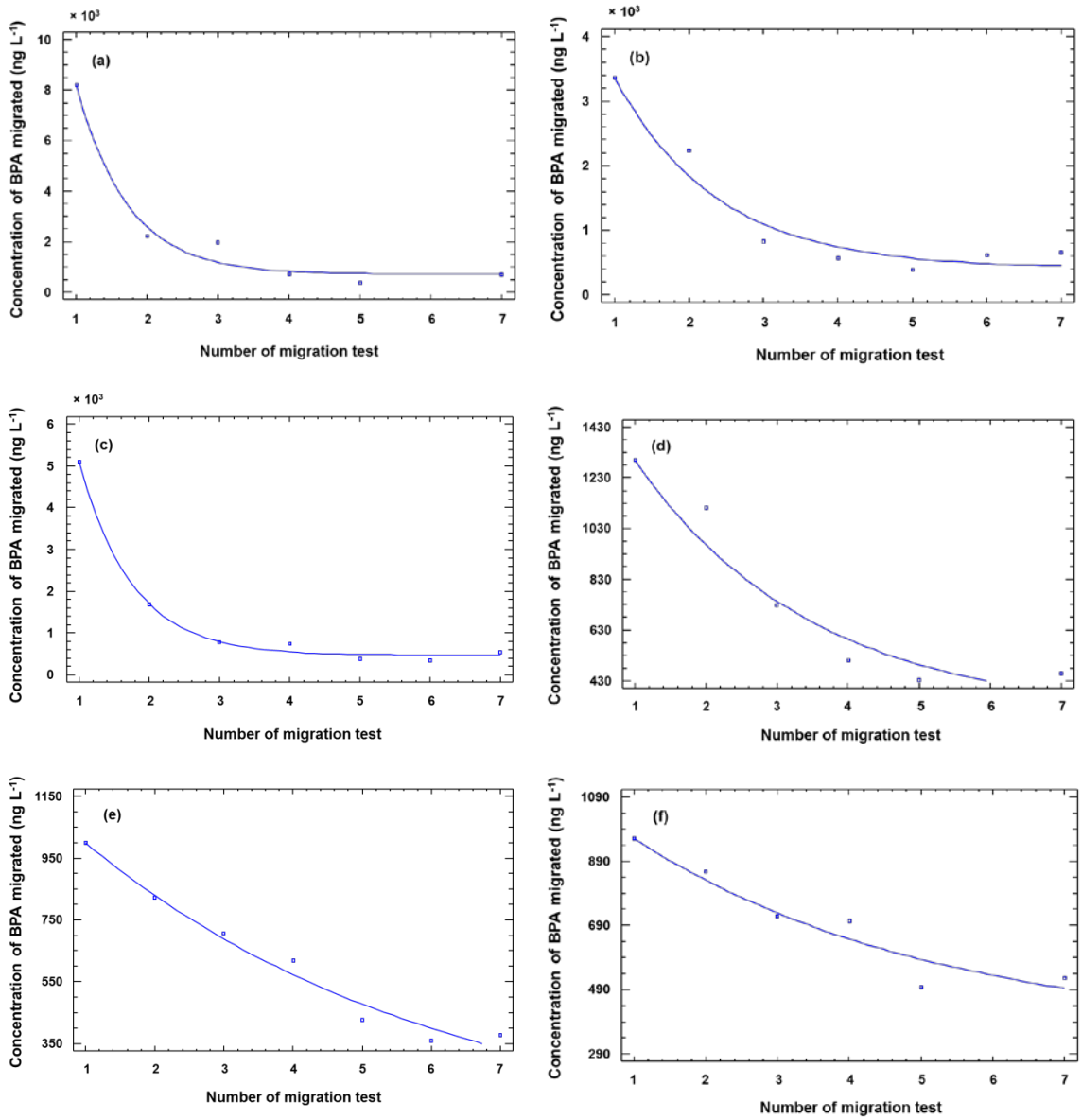


**Figure 4.3.2.** Non-linear fitting of the experimental data on the migration of BPA into food simulant *D1* at 70 ºC (**(a)** Glass 1; **(b)** Glass 2; **(c)** Glass 3) and at 50 ºC (**(d)** Glass 4; **(e)** Glass 5; **(f)** Glass 6); (□) average concentration of BPA after every test; (___) fitted model.

**Table 4.3.10.** Parameters of the models fitted following the equation $c_{BPA,\,Test\ t} = a + \left(c_{BPA,\,Test\ 1} - a\right)e^{-b\left(Test\ t - Test\ 1\right)}$

| PC container | Outlier data | *a* (ng $L^{-1}$) | Standard error of *a* (ng $L^{-1}$) | Asymptotic 95% confidence interval for *a* (ng $L^{-1}$) | *b* ($Test^{-1}$) | Standard error of *b* ($Test^{-1}$) | $R^2$ (%) | Asymptotic 95% confidence interval for *b* ($Test^{-1}$) | Residual standard deviation (ng $L^{-1}$) | Correlation *a*–*b* |
|---|---|---|---|---|---|---|---|---|---|---|
| Glass 1 | 1/7 | 708.33 | 289.67 | (−95.93, 1512.6) | 1.39 | 0.29 | 97.91 | (0.59, 2.18) | 478.21 | 0.62 |
| Glass 2 | 0/7 | 407.04 | 201.97 | (−112.15, 926.23) | 0.73 | 0.18 | 95.33 | (0.26, 1.19) | 266.40 | 0.78 |
| Glass 3 | 0/7 | 462.16 | 63.64 | (298.57, 625.75) | 1.33 | 0.10 | 99.59 | (1.06, 1.59) | 120.23 | 0.58 |
| Glass 4 | 1/7 | 292.11 | 173.01 | (−188.24, 772.46) | 0.40 | 0.15 | 94.32 | (−0.02, 0.83) | 98.15 | 0.93 |
| Glass 5 | 0/7 | 31.15 | 224.10 | (−544.93, 607.22) | 0.19 | 0.07 | 97.56 | (0.01, 0.38) | 41.80 | 0.99 |
| Glass 6 | 1/7 | 350.22 | 214.27 | (−244.69, 945.12) | 0.24 | 0.14 | 92.56 | (−0.16, 0.64) | 55.26 | 0.98 |

On the other hand, none of the confidence intervals for $\hat{b}$ in the 70 ºC kinetics (Glasses 1, 2 and 3) included zero; however, that happened for Glasses 4 and 6, whereas the lower limit of that interval in the case of Glass 5 equalled 0.01. This fact led to the conclusion that the average release of BPA from the PC plastic had not been significant at a 95% confidence level when the migration process had been conducted at 50 ºC; on the contrary, BPA had been significantly transferred into the food simulant from the three glasses heated at 70 ºC. The confidence intervals for *a* showed that the limit value of the concentration of BPA migrated after infinite tests could be considered to be zero with just the exception of Glass 3. Table 4.3.10 also illustrates that the estimates of *a* and *b* were correlated, especially for the migration study at 50 ºC.

The confidence region surrounding the least squares estimator $(\hat{a},\hat{b})$ for each of the six migration models fitted following Eq. (4.3.3) was then determined as explained in chapter 24 of [30]. At each temperature, the representation of these confidence contours is shown in Figures 4.3.3.a (70 ºC) and 4.3.3.b (50 ºC). Since the model in Eq. (4.3.3) was nonlinear, the confidence regions for $\hat{a}$ and $\hat{b}$ were not ellipsoidal but tended to be irregular with a great asymmetry. It can be seen that, at each temperature, the three confidence regions were similar in both orientation and shape, but not in size, since its area is directly related to the residual variance of the model: the lower the value of the residual standard deviation (see Table 4.3.10, tenth column), the greater the quality of the precision on the predictions and the smaller the area within the confidence contour. However, when comparing the two kinetic studies (see Figures 4.3.3.a and 4.3.3.b), a great difference in the shape of the confidence regions calculated at the two heating temperatures was evident, which fully accorded with the differences found between both migration processes. All these facts proved that the analytical behaviour of every glass regarding the migration of BPA could be considered to be the same, but described by different values of the parameters *a* and *b* in each case. So it was not possible to pose a general mathematical equation that served as a migration model valid for all the glasses, not even for those evaluated at the same temperature: the different values of the estimates of the model and especially of the concentration of BPA migrated after the first test could be due to differences in the composition and the structure of the polymer that made up every glass.

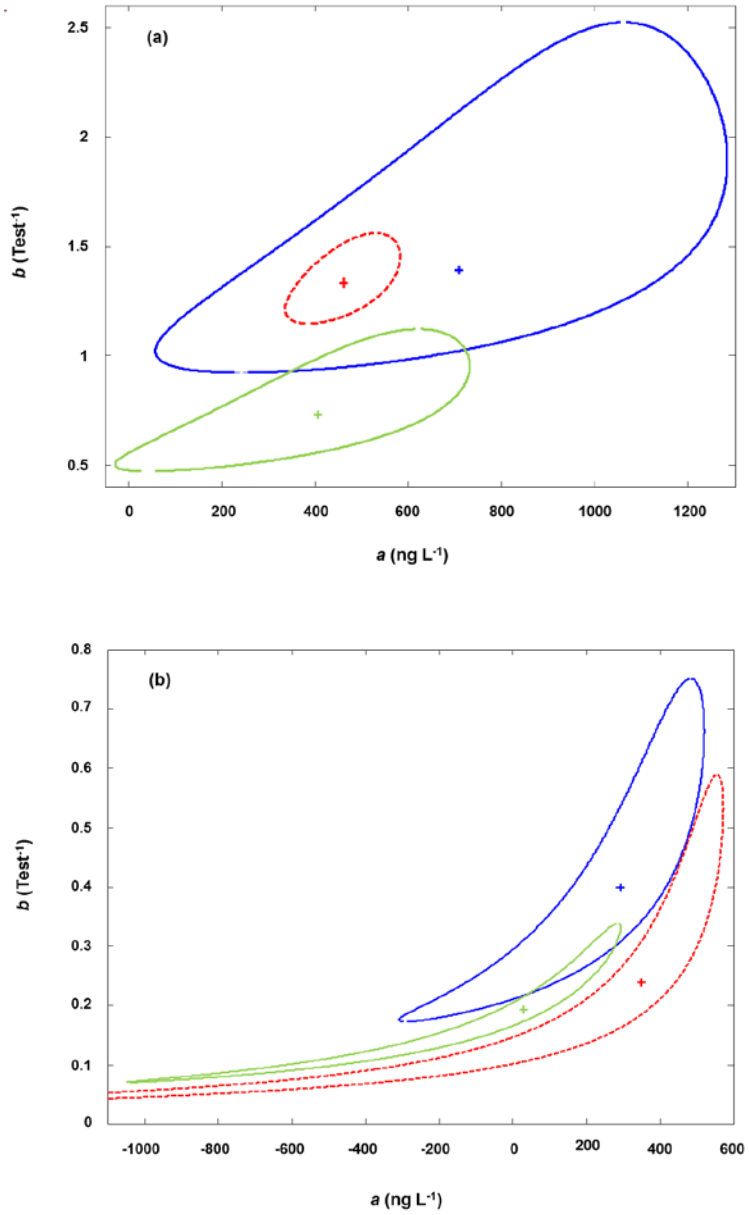


**Figure 4.3.3.** Confidence regions for the estimates $\hat{a}$ and $\hat{b}$ of the migration nonlinear models resulting from **(a)** the 70 ºC study and **(b)** the 50 ºC study. Glasses 1 and 4: dark blue lines; Glasses 2 and 5: light green lines; Glasses 3 and 6: dotted red lines.

### 4.3.6. Conclusions

After a preliminary examination into the pH range 2–12, it has been stated that the best working pH for the simultaneous determination of BPF, BPA, BFDGE and BADGE is 9; this conclusion agrees with the fact that extreme pH conditions will cause the loss of BFDGE and BADGE by epoxide ring-opening and, on the other hand, with the $pK_{a,\ 25\ ºC}$ values of BPF and BPA. The trueness of the analytical method has been also verified at a

95% confidence level, being the capability of detection for BPA equal to 2.60 μg $L^{-1}$ ($n$ = 2 replicates, $\alpha = \beta = 0.05$).

In a second part of the study, the fitting of the experimental data to an exponential migration model for every glass has demonstrated that *i)* the higher the heating temperature, the greater the relative migration rate of BPA, and *ii)* BPA is hardly significantly transferred from the plastic to the food simulant at 50 ºC, whereas it is at 70 ºC. An identical analytical behaviour regarding the release of BPA has been found for all the glasses evaluated at the same temperature. However, this essentially common migration process is not described by the same values of the parameters in the equation for all the containers, which has meant in practice the existence of a different underlying behaviour. This makes it unfeasible to pose a general migration equation.

Carrying out the PARAFAC decomposition of all the GC-MS signals has succeeded in: *i)* identifying every compound unequivocally according to the maximum permitted tolerances for relative ion abundances and relative retention time, which guaranteed the specificity of the analysis, and *ii)* quantifying every analyte, even in the presence of interferents. It must be noticed that, in some cases, these steps of identification and quantification could never have been achieved from the mass spectrum recorded at the retention time because of coeluents that share ions with the analyte of interest, concealing its presence to such an extent that it would have been wrongly concluded that the compound was missing in the sample.

### 4.3.7. Acknowledgements

The authors wish to thank the financial support from Ministerio de Economía y Competitividad and Junta de Castilla y León through projects CTQ2011-26022 and BU108A11-2, respectively. M.L. Oca is particularly grateful to Universidad de Burgos for her FPI grant.

### 4.3.8. References

[1] P. Viñas, I. López-García, N. Campillo, R.E. Rivas, M. Hernández-Córdoba, *Ultrasound-assisted emulsification microextraction coupled with gas chromatography–mass spectrometry using the Taguchi design method for bisphenol migration studies from thermal printer paper, toys and baby utensils*, Analytical and Bioanalytical Chemistry 404 (2012) 671-678.

[2] S. Biedermann, P. Tschudin, K. Grob, *Transfer of bisphenol A from thermal printer paper to the skin*, Analytical and Bioanalytical Chemistry 398 (2010) 571-576.

[3] K. Bhunia, S.S. Sablani, J. Tang, B. Rasco, *Migration of Chemical Compounds from Packaging Polymers during Microwave, Conventional Heat Treatment, and Storage*, Comprehensive Reviews in Food Science and Food Safety 12 (2013) 523-545.

[4] Commission Implementing Regulation (EU) No 321/2011 of 1 April 2011 amending Regulation (EU) No 10/2011 as regards the restriction of use of Bisphenol A in plastic infant feeding bottles, Official Journal of the European Union, L 87, 1-2.

[5] Commission Regulation (EU) No 10/2011 of 14 January 2011 on plastic materials and articles intended to come into contact with food, Official Journal of the European Union, L 12, 1-89.

[6] Commission Regulation (EC) No 1895/2005 of 18 November 2005 on the restriction of use of certain epoxy derivatives in materials and articles intended to come into contact with food, Official Journal of the European Union, L 302, 28-32.

[7] F.A. Carey, R.M. Giuliano, *Organic Chemistry*, McGraw-Hill, New York, 8th ed., 2011.

[8] H. Gallart-Ayala, E. Moyano, M. T. Galceran, *Multiple-stage mass spectrometry analysis of bisphenol A diglycidyl ether, bisphenol F diglycidyl ether and their derivatives*, Rapid Communications in Mass Spectrometry 24 (2010) 3469-3477.

[9] P. Paseiro, J. Simal, S. Paz, P. López, J. Simal, *Kinetics of the hydrolysis of bisphenol A diglycidyl ether (BADGE) in water-based food simulants*, Fresenius' Journal of Analytical Chemistry 345 (1993) 527-532.

[10] P. Paseiro, J. Simal, S. Paz, P. López, J. Simal, *Kinetics of the Hydrolysis of Bisphenol F Diglycidyl Ether in Water-Based Food Simulants. Comparison with Bisphenol A Diglycidyl Ether*, Journal of Agricultural and Food Chemistry 40 (1992) 868-872.

[11] R. Sendón, P. Paseiro, *Determination of bisphenol A diglycidyl ether and its hydrolysis and chlorohydroxy derivatives by liquid chromatography–mass spectrometry*, Journal of Chromatography A 1032 (2004) 37-43.

[12] N. Leepipatpiboon, O. Sae-Khow, S. Jayanta, *Simultaneous determination of bisphenol-A-diglycidyl ether, bisphenol-F-diglycidyl ether, and their derivatives in oil-in-water and aqueous-based canned foods by high-performance liquid chromatography with fluorescence detection*, Journal of Chromatography A 1073 (2005) 331-339.

[13] X. Cao, G. Dufresne, G. Clement, S. Bélisle, A. Robichaud, F. Beraldin, *Levels of Bisphenol A Diglycidyl Ether (BADGE) and Bisphenol F Diglycidyl Ether (BFDGE) in Canned Liquid Infant Formula Products in Canada and Dietary Intake Estimates*, Journal of AOAC International 92 (2009) 1780-1789.

[14] H. Gallart-Ayala, E. Moyano, M.T. Galceran, *Fast liquid chromatography–tandem mass spectrometry for the analysis of bisphenol A-diglycidyl ether, bisphenol F-diglycidyl ether and their derivatives in canned food and beverages*, Journal of Chromatography A 1218 (2011) 1603-1610.

[15] J. Míguez, C. Herrero, I. Quintás, C. Rodríguez, P.G. Gigosos, O.C. Mariz, *A LC-MS/MS method for the determination of BADGE-related and BFDGE-related compounds in canned fish food samples based on the formation of [M + NH4]+ aducts*, Food Chemistry 135 (2012) 1310-1315.

[16] Y. Zou, S. Lin, S. Chen, H. Zhang, *Determination of bisphenol A diglycidyl ether, novolac glycidyl ether and their derivatives migrated from can coatings into foodstuff by UPLC-MS/MS*, European Food Research and Technology 235 (2012) 231-244.

[17] E.P. Serjeant, B. Dempsey, *Ionisation Constants of Organic Acids in Aqueous Solution*, IUPAC chemical data series No. 23, Pergamon Press, New York, 1979.

[18] I.T. Cousins, C.A. Staples, G.M. Klěcka, D. Mackay, *A Multimedia Assessment of the Environmental Fate of Bisphenol A*, Human and Ecological Risk Assessment 8 (2002) 1107-1135.

[19] J. Brandsch, P. Mercea, M. Rüter, V. Tosa, O. Piringer, *Migration modelling as a tool for quality assurance of food packaging*, Food Additives & Contaminants 19 (2002) 29-41.

[20] A. Sanches, J.M. Cruz, R. Sendón, R. Franz, P. Paseiro, *Kinetic migration studies from packaging films into meat products*, Meat Science 77 (2007) 238-245.

[21] M.F. Poças, J.C. Oliveira, J.R. Pereira, R. Brandsch, T. Hogg, *Modelling migration from paper into a food simulant*, Food Control 22 (2011) 303-312.

[22] M.F. Poças, J.C. Oliveira, R. Brandsch, T. Hogg, *Analysis of mathematical models to describe the migration of additives from packaging plastics to food*, Journal of Food Process Engineering 35 (2012) 657-676.

[23] N. Lin, Y. Zou, H. Zhang, *Kinetic migration studies of bisphenol-A-related compounds from can coatings into food simulant and oily foods*, European Food Research and Technology 237 (2013) 1009-1019.

[24] S. Bratinova, B. Raffael, C. Simoneau, *Guidelines for performance criteria and validation procedures of analytical methods used in controls of food contact materials - EUR 24105 EN (JRC53034)*, Publications Office of the European Union, Luxembourg, 1st ed., 2009.

[25] R. Bro, *PARAFAC. Tutorial and applications*, Chemometrics and Intelligent Laboratory Systems 38 (1997) 149-171.

[26] M.C. Ortiz, L. Sarabia, *Quantitative determination in chromatographic analysis based on n-way calibration strategies*, Journal of Chromatography A, 1158 (2007) 94-110.

[27] R. Bro, H.A.L. Kiers, *A new efficient method for determining the number of components in PARAFAC models*, Journal of Chemometrics 17 (2003) 274-286.

[28] T. Pihlström, *Method validation and quality control procedures for pesticide residues analysis in food and feed (SANCO/12571/2013 – Implemented by 01.01.2014)*, EU, Brussels, 2013.

[29] Commission Decision (EC) No 2002/657/EC of 12 August 2002, implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Official Journal of the European Communities, L 221, 8-36.

[30] N.R. Draper, H. Smith, *Applied regression analysis*, John Wiley and Sons, New York, 3rd ed., 1998.

[31] M.C. Ortiz, M.S. Sánchez, L.A. Sarabia, *Quality of Analytical Measurements: Univariate Regression,* in: S. Brown, R. Tauler, B. Walczak (Eds.), *Comprehensive Chemometrics, vol. 1.*, Elsevier, Oxford, 2009.

[32] B.M. Wise, N.B. Gallagher, R. Bro, J.M. Shaver, W. Windig, R.S. Koch, *PLS Toolbox 5.0*, Eigenvector Research, Inc. 3905 West Eaglerock Drive, Wenatchee, 2006.

[33] *STATGRAPHICS Centurion XVI (Version 16.1.11)*, StatPoint Technologies, Inc., Herndon, VA, 2010.

[34] L. Sarabia, M.C. Ortiz, *DETARCHI: A program for detection limits with specified assurance probabilities and characteristic curves of detection*, Trends in Analytical Chemistry 13 (1994) 1-6.

[35] M.C. Ortiz, L.A. Sarabia, I. García, D. Giménez, E. Meléndez, *Capability of detection and three-way data*, Analytica Chimica Acta 559 (2006) 124-136.

[36] M.L. Oca, M.C. Ortiz, A. Herrero, L.A. Sarabia, *Optimization of a GC/MS procedure that uses parallel factor analysis for the determination of bisphenols and their diglycidyl ethers after migration from polycarbonate tableware*, Talanta 106 (2013) 266-280.

# CHAPTER 5

## ENSURING FOOD QUALITY BY FIRST ENSURING ANALYTICAL QUALITY WHEN COMPLEX SIGNALS AND AN UNCONTROLLED SOURCE FOR ONE OF THE TARGET MIGRANTS ARE PRESENT IN THE LABORATORY ENVIRONMENT: DEALING WITH THE UBIQUITY OF PHTHALATES IN THE LABORATORY WHEN DETERMINING PLASTICIZERS BY GAS CHROMATOGRAPHY-MASS SPECTROMETRY AND PARAFAC

## 5.1. Introducción

Hoy en día, los plastificantes, grupo al que pertenece la familia química de los ftalatos, se utilizan de manera muy notable como aditivos en el proceso de fabricación de polímeros plásticos flexibles: desde la Segunda Guerra Mundial, el uso de estos compuestos se ha ido extendiendo progresivamente en todas las industrias manufactureras hasta convertirse en primordial. En este sentido, las industrias química y envasadora de alimentos no han sido ninguna excepción a esta tendencia creciente.

De la misma manera que los plastificantes pueden migrar desde el envase al alimento, tal y como se ha visto en el capítulo anterior, el material y los equipos usados día a día en un laboratorio de análisis también pueden liberar estos mismos compuestos a la muestra que está siendo preparada y analizada en ese entorno en función de las condiciones físico-químicas a las que es sometida la muestra y los componentes plásticos con los que ella interactúa durante la ejecución del método analítico. Así, si ambas fuentes de plastificantes (envase alimentario y entorno de análisis) convergieran en una muestra de laboratorio y los analitos de interés fuesen específicamente dichas sustancias (por ejemplo, en el marco de un ensayo de conformidad de la migración de las mismas desde materiales y objetos plásticos de contacto alimentario), sería necesario trazar en primer lugar la línea base correspondiente a la contribución del entorno de análisis a la cantidad de plastificantes presente en la muestra con el objetivo de minimizar el riesgo de obtener finalmente un falso positivo. En este caso, asegurar la Calidad Alimentaria de los productos sometidos a un ensayo de conformidad de la migración desde sus envases en estas condiciones implica necesariamente ser capaz de asegurar, en primer lugar, la Calidad Analítica de los resultados proporcionados por el método de laboratorio en lo que a la existencia de un sesgo de fondo se refiere.

En este capítulo se recoge la determinación multirresiduo de dos ftalatos (diisobutilftalato (DiBP) y diisononilftalato (DiNP)), junto con otros dos plastificantes (butilhidroxitolueno (BHT) y adipato de bis (2-etilhexilo) (DEHA)) y un estabilizante frente a la radiación ultravioleta (benzofenona (BP)) mediante GC-MS empleando DiBP-$d_4$ como estándar interno y PARAFAC como la estrategia quimiométrica para asegurar la cuantificación y la identificación inequívoca de todas las sustancias anteriores, en conformidad con la legislación europea. Al demostrarse durante el desarrollo del método que DiBP estaba presente en el entorno del laboratorio, muy probablemente proveniente del material o de los equipos analíticos usados, se evaluaron en primera instancia las distribuciones de probabilidad de la concentración de este compuesto en blancos y en disoluciones patrón, preparadas a un determinado nivel de DiBP. De este modo, dadas unas probabilidades de falso no cumplimiento ($\alpha$) y de falso cumplimiento ($\beta$), fue posible la estimación, a través

de PARAFAC, del valor crítico del *loading* muestral estandarizado de DiBP por encima del cual la concentración de este analito es estadísticamente mayor que la existente en los blancos. La determinación de este valor umbral permitía a la señal analítica de DiBP dar cuenta de una manera veraz y precisa de fuentes adicionales de este compuesto más allá de la correspondiente a la contribución del laboratorio.

Además, la estrategia estadística diseñada y detallada en este capítulo hizo posible la superación de un segundo obstáculo analítico: la señal cromatográfica de DiNP tenía una forma de múltiples picos (simulando los dedos de una mano) debido a la presencia de un conjunto de isómeros $C_9$ en la composición del patrón de referencia adquirido para el desarrollo del método. En este contexto, PARAFAC se erigió como una aproximación más fiable para el tratamiento de estas señales analíticas tan complejas, llevando a cabo de manera satisfactoria la determinación de este analito al no emplear para ello áreas de pico, sino *loadings* cromatográficos y espectrales.

Por último, se evaluó y confirmó la existencia de efecto matriz en el análisis de estos compuestos, por lo que su cuantificación tras una extracción exhaustiva desde una tetina infantil hubo de ser realizada mediante la metodología de adición estándar. Así, en el extracto final se verificó la presencia de BHT proveniente de la tetina a una concentración significativa de 37.87 $\mu g\ L^{-1}$.

## 5.2. Dealing with the ubiquity of phthalates in the laboratory when determining plasticizers by gas chromatography-mass spectrometry and PARAFAC[1]

### 5.2.1. Abstract

Determining plasticizers and other additives migrated from plastic materials becomes a hard task when these substances are already present in the laboratory environment. This work dealt with this drawback in the multiresidue determination of four plasticizers (2,6-di-*tert*-butyl-4-methyl-phenol (BHT), diisobutyl phthalate (DiBP), bis(2-ethylhexyl) adipate (DEHA) and diisononyl phthalate (DiNP)) and a UV stabilizer (benzophenone (BP)) by gas chromatography-mass spectrometry (GC-MS) using DiBP-$d_4$ as internal standard. The ubiquity of DiBP by a non-constant leaching process in the laboratory was detected, which could not guarantee the achievement of a trustworthy quantification. To handle this, the assessment of the level of DiBP in solvent blanks having fixed the probabilities of false non-compliance ($\alpha$) and false compliance ($\beta$) at 0.01 was performed. On the other hand, another special case was that of DiNP, in whose chromatogram finger peaks appear because of an array of possible $C_9$ isomers. PARAFAC, used for the identification and quantification of all the substances, is a useful chemometric tool that enabled a more reliable determination of this analyte since no peak areas were considered but chromatographic and spectral loadings.

Since phthalates may migrate from rubber latex items, an evaluation of the existence of matrix effects on the determination of the five analytes was conducted prior to an extraction with hexane from a dummy for infants. As matrix effects were present, the quantification of the compounds under study was performed following the standard addition method using PARAFAC sample loadings as response variable. As a result, the presence of BHT was confirmed, being its concentration equal to 37.87 μg $L^{-1}$. Calibrations based on PARAFAC yielded the following values for the decision limit ($CC\alpha$): 1.16 μg $L^{-1}$ for BHT, 1.34 μg $L^{-1}$ for BP, 1.84 μg $L^{-1}$ for DEHA and 51.42 μg $L^{-1}$ for DINP (for $\alpha = 0.05$ and two replicates).

### 5.2.2. Introduction

Alkyl and aryl esters of 1,2-benzenedicarboxylic acid, better known as phthalates, are a group of synthetic compounds widely used as plasticizers. Mainly added to polyvinyl chloride (PVC), phthalates increase plastic flexibility, resistance and durability. Widely used items such as toys, water pipes, wallpaper, artificial leather, electrical wire insulation, glue,

[1] This section is published as *M.L. Oca et al. / J. Chromatogr. A 1464 (2016) 124–140* (10 citations up to August 30, 2021 according to *Scopus*).

cosmetics, plastic water bottles, paints and printing ink include phthalate additives in their formulation, and as a result, a huge worldwide market has been created around these compounds. However, phthalates have been classified as endocrine disrupting chemicals and potential human carcinogenic agents [1,2,3,4]. They can be easily leached out from plastic materials over time since they are not chemically bound to the polymeric matrix. Due to a direct migration process from packaging films and plastic containers, phthalates can thus be found in a high concentration in foodstuffs and beverages, especially in fatty products because of the hydrophobia of these chemicals [5,6,7,8]. They can even be released into the air [9,10]. In this context, the EU Regulation No 10/2011 [11] only permits a limited use of certain phthalates in food contact materials by establishing specific migration limits (SMLs) on the basis of toxicological evaluations by the European Food Safety Authority.

Aside from health concerns, their ubiquity can cause frustration and big trouble when analysing them, especially if trace analyses of phthalates by chromatographic methods are to be carried out [12,13]. Cross-contamination during the analytical procedure often leads to false positive or overestimated results, while actually only a part of the predicted concentration should have been related to the sample. Phthalates are indeed everywhere in the laboratory environment, including solvents, chemicals, glassware, gloves and consumables, such as rubber tubing, syringes, pipette tips, filters, stir bars, vials, 96-well plates, caps, etc. [12,13]; even both the chromatographic system and laboratory air may be sources of contamination [12]. So a great probability of phthalate interferences exists whenever gas-chromatography and/or liquid-chromatography analyses are performed. Common phthalate-origin lab interferents are diisobutyl phthalate (DiBP), di(2-ethylhexyl) phthalate, di-*n*-butyl phthalate, diisononyl phthalate (DiNP), butyl benzyl phthalate, di-*n*-octyl phthalate and dimethyl phthalate [14]. Among them, the first three compounds are mainly involved in lab environment and blank issues, especially the latter two due to their low molecular weight, easy partition from polymer matrix, solubility in water and high usage in PVC production [13].

Several strategies to reduce the presence of phthalate contaminants have been developed [12,15,16], such as avoiding plastic labware, baking out glassware or rinsing it with redistilled solvents. However, even so, this problem is proving to be very difficult to get rid of, so usually there is no choice but to take its existence into account and face it. Nevertheless, some guarantees should be first established in order to be able to distinguish significantly the amount of phthalates present in a blank from that in a sample. This would be easy to achieve if the blank concentration profile of the contaminant displayed a constant pattern, so that the blank average value could be subtracted to determine the amount of the phthalate under study in the sample. But if this is not the case and a variable

leaching behaviour is detected, difficulties arise, because now the blank average value would not be representative of the true presence of the interferent in the laboratory environment. This drawback could be tackled by using the approach proposed in this work, which has consisted in *i*) estimating the probability distributions of the concentration in blanks and standards at a fixed level of the controversial contaminants, and then *ii*) after setting the probability of making a type I error ($\alpha$) and that of a type II error ($\beta$), computing the critical or rejection concentration value beyond which the amount of interferents in a sample is statistically higher than that at the blank level. A more detailed explanation on the use of this methodology in analytical measurements can be found in [17,18].

The determination of DiBP and DiNP by gas chromatography coupled to mass spectrometry (GC-MS) has been pursued in this work, being the former variably ubiquitous while performing the study. No restrictive regulations are established on DiBP regarding its use in plastic manufacturing, so the generic SML of 60 mg $kg^{-1}$ should apply according to [11]. On the other hand, DiNP (in the form of diesters of the phthalic acid with primary saturated $C_8$–$C_{10}$ branched alcohols, being $C_9$ more than 60%) is only authorised to be used as additive or polymer production aid within the European Union, and a total SML of 9 mg $kg^{-1}$ is set [11]. In this work, these SML values have been taken into account for the performance of the procedure. Due to the composition of the chemical marketed as DiNP, its signal typically appears in the chromatogram as finger peaks instead of a single one because of an array of $C_9$ isomers (with alkyl groups with different degree of branching on both side chains of the phthalic ring) that were impossible to separate before commercializing. In fact, there are two main CAS numbers available for DiNP from several manufacturers: DiNP 1 (CAS no. 68515-48-0), which is a mixture of isomers with alkyl chains from 8-to-10- (mainly 9) carbons long, and DiNP 2 (CAS no. 28553-12-0), composed exclusively of $C_9$-chains [19,20]. Anyway, a quantitation based on the area summation integration of that signal could lead to unreliable results, either greater or lower than the true ones, because of the complex shape of those peaks. However, the statistical methodology Parallel Factor Analysis (PARAFAC) decomposition, used in this work not only to quantify the analytes, but also to identify them unequivocally by taking advantage of the second-order property of the GC-MS data, offers a more trustworthy approach to deal with complex signals such as that of DiNP. Ensuring the unequivocal presence of every analyte in a sample is a key question when target analyses are performed, as official regulations and guidelines [21,22,23] lay down in the form of tolerance interval requirements. This identification, based on the unique mass spectrum and elution time profiles estimated by PARAFAC, guarantees the specificity of the analysis, even though shared ions are present at the same retention time.

As well as these two phthalate derivatives, the multiresidue gas-chromatography method developed in this work also pursues the determination of benzophenone (BP), 2,6-di-*tert*-butyl-4-methyl-phenol (BHT) and bis(2-ethylhexyl) adipate (DEHA). BP is widely used as a flavour ingredient, a fragrance enhancer, in the manufacture of insecticides, agricultural chemicals, hypnotic drugs, antihistamines and other pharmaceuticals; but its main application comes from its ultraviolet- (UV) curing property, which makes it suitable for being added to plastic packaging as a UV blocker. This allows manufacturers to package their products in clear glass or plastic rather than opaque or dark packaging to prevent them from being damaged [24]. DEHA is a common plasticizer primarily used together with phthalates in the PVC industry, while BHT is a manufactured antioxidant commonly used as preservative in plastics, rubber, petroleum products, foods, pharmaceuticals and cosmetics. These three substances are allowed to be used as additive or polymer production aid within the European Union, being the values of their SMLs equal to 0.6 mg $kg^{-1}$ for BP, 3 mg $kg^{-1}$ for BHT and 18 mg $kg^{-1}$ for DEHA [11]. Diisobutyl phthalate-3,4,5,6-$d_4$ (DiBP-$d_4$) has been used as the chromatographic internal standard (IS) for the five analytes of interest.

The presence of these five compounds in the natural rubber latex forming a dummy intended for infants has also been studied, and their leaching into a volume of hexane has thus been tested. But firstly, in order to ensure a proper determination of every substance in these samples, the existence of matrix effects has been evaluated and confirmed. This fact has led to the decision of estimating the concentration of each analyte leached into the volume of hexane by means of the standard addition methodology.

### 5.2.3. Material and methods

#### *5.2.3.1. Chemicals*

Benzophenone (CAS no. 119-61-9; purified by sublimation), 2,6-di-*tert*-butyl-4-methyl-phenol (CAS no. 128-37-0), diisobutyl phthalate (CAS no. 84-69-5), diisobutyl phthalate-3,4,5,6-$d_4$ (CAS no. 358730-88-8; analytical standard), bis(2-ethylhexyl) adipate (CAS no. 103-23-1) and diisononyl phthalate (CAS no. 28553-12-0; ester content ≥ 99%, mixture of $C_9$ isomers), all of 99% or higher purity, were purchased from Sigma-Aldrich (Steinheim, Germany). *n*-Hexane (CAS no. 110-54-3; for liquid chromatography Lichrosolv®) was obtained from Merck KGaA (Darmstadt, Germany).

#### *5.2.3.2. Standard solutions*

Stock solutions were prepared individually in hexane at a concentration of 1000 mg $L^{-1}$ for BHT and BP, of 700 mg $L^{-1}$ for DiBP-$d_4$ and of 2000 mg $L^{-1}$ for DiBP, DEHA and DiNP. Intermediate solutions were prepared from the previous ones by dilution in the same solvent. All the solutions were kept in crimp vials, stored at low temperature (4ºC) and protected from light. Table 5.1 shows the standards and/or samples analysed (concentration ranges and number of standards) together with the dimensions of the data tensors for each experimental stage of this work (Sections 5.2.5.1 to 5.2.5.4). Regarding the step related to the procedure to assess the level of DiBP, as can be seen in Table 5.1, the solvent blanks and standards were analysed in two days' time due to the impossibility of carrying out the whole analysis within the same day. The data from the analysis of 2 out of the 24 solvent blanks containing IS and the 22 non-zero solvent standards performed the first day were also used to estimate the calibration models for BHT, BP, DEHA and DiNP (see Section 5.2.5.1.2).

Only glassware was used throughout all this study, and plastic consumables were avoided as far as possible. A thorough procedure for cleaning the lab glassware was followed to try to minimize cross-contamination from plasticizers.

#### *5.2.3.3. Extraction in hexane conditions*

The surface of a dummy for infants was placed in a beaker with 60 mL of hexane for 1 hour to perform an exhaustive extraction of the potential migrants. The beaker was covered with a watch glass in order to minimize the evaporation of the solvent. This extraction experiment was performed at room temperature. After 1 hour, the dummy was removed and fresh hexane was added till reaching up the original volume again. This extract was stored under refrigeration at 4ºC and used to perform the standard addition method.

#### *5.2.3.4. GC-MS analysis*

Analyses were carried out on an Agilent 7890A gas chromatograph coupled to an Agilent 5975C mass spectrometer detector (Agilent Technologies, Santa Clara, CA, USA). The analytical column used was an Agilent HP-5MS Ultra Inert (30 m × 0.25 mm i.d., 0.25 μm film thickness). Helium was used as the carrier gas at a constant flow of 1.3 mL $min^{-1}$, and the initial pressure was set at 10.121 psi.

**Table 5.1.** Summary of the samples analysed, concentration ranges and dimensions of the data tensors built for the analytes in each experimental stage of this work.

| Experimental stage | Samples analysed | Concentration range | Dimension of the resulting data tensor (Scans × Ions × Samples) | | | | |
|---|---|---|---|---|---|---|---|
| | | | **BHT** | **BP** | **DiBP and DiBP-$d_4$** | **DEHA** | **DiNP** |
| Tolerance intervals | 3 system blanks<br>1 solvent blank without IS<br>6 reference standards | 10–50 μg $L^{-1}$ (for BHT, BP, DiBP and DiBP-$d_4$)<br>25–75 μg $L^{-1}$ (for DEHA)<br>1–3 mg $L^{-1}$ (for DiNP) | 21 × 5 × 10 | 32 × 5 × 10 | 41 × 10 × 10 | 19 × 5 × 10 | 823 × 6 × 10 |
| Calibration | 8 system blanks<br>1 solvent blank without IS<br>22 solvent blanks with IS<br>24 standards at 12 concentration levels (in duplicate) | 0–50 μg $L^{-1}$ (for BHT, BP)<br>25 μg $L^{-1}$ of DiBP and DiBP-$d_4$<br>0–80 μg $L^{-1}$ (for DEHA)<br>0–3 mg $L^{-1}$ (for DiNP) | 21 × 5 × 55 | 32 × 5 × 55 | 41 × 10 × 55 | 19 × 5 × 55 | 823 × 6 × 55 |
| Procedure to assess the level of DiBP | 18 system blanks<br>2 solvent blank without IS<br>50 solvent blanks with IS<br>44 standards at 11 concentration levels (in quadruplicate)<br>16 extracts after contact with latex rubber nipples | 0–50 μg $L^{-1}$ (for BHT, BP)<br>25 μg $L^{-1}$ of DiBP and DiBP-$d_4$<br>0–80 μg $L^{-1}$ (for DEHA)<br>0–3 mg $L^{-1}$ (for DiNP) | 21 × 5 × 130 | 32 × 5 × 130 | 41 × 10 × 119<br>(11 outlier samples rejected) | 19 × 5 × 130 | 823 × 6 × 130 |
| Study of the matrix effect | 8 system blanks<br>1 solvent blank without IS<br>6 solvent blanks and 6 solvent standards measured before matrix-matched standards<br>12 matrix-matched standards<br>6 solvent blanks and 6 solvent standards measured after matrix-matched standards | 0-45 μg $L^{-1}$ (for BHT, BP)<br>0–75 μg $L^{-1}$ (for DiBP)<br>25 μg $L^{-1}$ of DiBP-$d_4$<br>0–56 μg $L^{-1}$ (for DEHA)<br>0–2.5 mg $L^{-1}$ (for DiNP) | 21 × 5 × 45 | 32 × 5 × 45 | 41 × 10 × 45 | 36 × 5 × 45 | 823 × 6 × 45 |

**Table 5.1. (cont.)**

| Experimental stage | Samples analysed | Concentration range | Dimension of the resulting data tensor (Scans × Ions × Samples) | | | | |
|---|---|---|---|---|---|---|---|
| | | | **BHT** | **BP** | **DiBP and DiBP-$d_4$** | **DEHA** | **DiNP** |
| Migration test from a dummy | 15 system blanks<br>1 solvent blank without IS<br>14 matrix-matched standards at 7 concentration levels (in duplicate) | 0–450 μg $L^{-1}$ (for BHT)<br>0–12 μg $L^{-1}$ (for BP)<br>0–40 μg $L^{-1}$ (for DiBP)<br>0–48 μg $L^{-1}$ (for DEHA)<br>0–1.5 mg $L^{-1}$ (for DiNP)<br>DiBP-$d_4$: 25 μg $L^{-1}$ and 225 μg $L^{-1}$ (for the 1st and 2nd standard addition method, respectively) | 21 × 5 × 30 | 32 × 5 × 30 | 41 × 10 × 30 (in both analyses) | 19 × 5 × 30 | 823 × 6 × 30 |

Injections were performed using the MultiPurpose Sampler MPS2XL from GERSTEL GmbH & Co. KG (Mülheim an der Ruhr, Germany) with a 10 μL syringe. The injection system consisted of a PTV inlet with a septumless head (CIS 6 from GERSTEL) equipped with a straight-with-notch quartz glass liner. A volume of 1 μL was injected at a controlled speed of 1 μL $s^{-1}$. Before and after each injection, the syringe was washed twice with acetone and twice with hexane, while ten washings with acetone and hexane were carried out after each injection throughout the measurement of the standard addition samples. The injection penetration was set at 40 mm, whereas the vial penetration was 30 mm. Amber glass vials were used for the analyses. The PTV inlet operated in the cold splitless mode. During the injection and for 0.1 min afterwards, the inlet temperature was 55ºC; this temperature was ramped then at 12ºC $s^{-1}$ from that initial value up to 270ºC, which was held for 15 min. The septum purge flow rate was set at 3 mL $min^{-1}$ while the purge flow rate through the split vent was 30 mL $min^{-1}$ from 0.6 min to 2 min. After 2 min, the flow rate was set at 20 mL $min^{-1}$.

The oven temperature was maintained at 40ºC for 0.6 min. Then the temperature was increased at 20ºC $min^{-1}$ to 250ºC, which was maintained for 1 min and next programmed at 10ºC $min^{-1}$ to 290ºC, which was held for 3 min. The run time was 19.1 min. A post-run step was performed at 300ºC for 4 min.

After a solvent delay of 8 min, the mass spectrometer was operated in the electron impact (EI) ionization mode at 70 eV. The transfer line temperature was set at 300°C, the ion source temperature at 230°C and the quadrupole temperature at 150°C. Data were acquired in single ion monitoring (SIM) mode. Five acquisition windows for SIM data were used: *i)* for BHT (start time: 8 min, ion dwell time: 30 ms), the following $m/z$ ratios were selected: 91, 145, 177, 205 and 220; *ii)* for BP (start time: 8.80 min, ion dwell time: 30 ms), the diagnostic ions were 51, 77, 105, 152 and 182; *iii)* for DiBP and DiBP-$d_4$ (start time: 9.80 min, ion dwell time: 10 ms), where the diagnostic ions for DiBP were 104, 149, 167, 205 and 223, and the diagnostic ions for DiBP-$d_4$ were 80, 153, 171, 209 and 227; *iv)* for DEHA (start time: 12 min, ion dwell time: 30 ms), the $m/z$ ratios recorded were 112, 129, 147, 241 and 259; *v)* for DiNP (start time: 14.60 min, ion dwell time: 25 ms), the diagnostic ions were 57, 127, 149, 167, 275 and 293.

#### *5.2.3.5. Software*

MSD ChemStation E.02.01.1177 (Agilent Technologies, Inc.) with Data Analysis software was used for data acquisition and processing. The NIST mass spectral library [25] was also used. PARAFAC and PARAFAC2 decompositions were performed with the PLS_Toolbox 6.0.1 [26] for use with MATLAB [27]. The building and validation of the regression models

as well as their comparison and the fitting of the probability distributions were carried out using STATGRAPHICS Centurion XVI [28]. Decision limit ($CC\alpha$) and capability of detection ($CC\beta$) were calculated with the DETARCHI program [29].

### 5.2.4. Theory

#### *5.2.4.1. PARAFAC and PARAFAC2 decompositions*

In the case of three-way data, PARAFAC decomposes a data tensor $\underline{\mathbf{X}}$ into triads or trilinear factors [30] and each factor consists of three loading vectors. GC-MS data can be arranged in a three-way array $\underline{\mathbf{X}}$ (of dimension $I \times J \times K$), where for each of the $K$ samples analysed, the abundance measured at $J$ $m/z$ ratios is recorded at $I$ elution times around the retention time of each compound. In this case, the trilinear PARAFAC model is:

$$\underline{\mathbf{X}} = x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk}, \quad i = 1, 2,\ldots, I; \; j = 1, 2,\ldots, J; \; k = 1, 2, \ldots, K \tag{5.1}$$

where $F$ is the number of factors, $\mathbf{a}_f$, $\mathbf{b}_f$ and $\mathbf{c}_f$ are the loading vectors of the chromatographic, spectral and sample profiles, respectively, and $e_{ijk}$ are the residuals of the model. The coordinates of the loading vectors are the columns of matrix **A**, **B** and **C** of size $I \times F$, $J \times F$ and $K \times F$, respectively.

Data are trilinear if the experimental data tensor is compatible with the structure shown in Eq. (5.1). In this case, the least-squares estimation of all its coefficients is unique.

The core consistency diagnostic (CORCONDIA) [31] measures the trilinearity degree of the experimental data tensor. If the data tensor is trilinear, then the maximum CORCONDIA value of 100 is found.

PARAFAC model is highly affected by deviations from the trilinear structure. Shifts in the retention time of analytes from sample to sample often happen in chromatography, which could affect trilinearity. PARAFAC2 [32,33] is a slightly different decomposition technique that overcomes the inequality in the chromatographic mode, thus allowing some deviation:

$$\underline{\mathbf{X}} = \left(x_{ijk}\right) = \left( \sum_{f=1}^{F} a_{if}^{k} b_{jf} c_{kf} + e_{ijk} \right), \quad i = 1, 2,\ldots, I; \; j = 1, 2,\ldots, J; \; k = 1, 2, \ldots, K \tag{5.2}$$

where the superscript $k$ is added to account for the dependence of the chromatographic profile on the $k$th sample. As a consequence, the loading matrices $\mathbf{A}_k$ are not necessarily equal for all $k = 1,\ldots,K$. So, while PARAFAC applies the same profiles (**A**, **B**) to a parallel set of matrices, as shown in Eq. (5.1), PARAFAC2 applies the same profile (**B**) along the

spectral mode instead, allowing the chromatographic mode to vary from one matrix to another. The uniqueness property is maintained whenever the cross-product $\mathbf{A}_k^T \mathbf{A}_k$ is restricted to be the same for $k = 1,\ldots,K$. Details and bibliographical references related to this question can be seen in [34].

The identification of outliers can be done through the indices $Q$ and Hotelling's $T^2$. If both indices of a sample exceeded the threshold values at a certain confidence level, that sample should be rejected and the PARAFAC or PARAFAC2 model should be reestimated.

When PARAFAC and PARAFAC2 have the uniqueness property, the unequivocal identification of compounds is possible by their chromatographic and spectral profiles, as some official regulations and guidelines [21,22,23] state.

#### *5.2.4.2. Comparison of regression models*

The evaluation of possible matrix effects was conducted following the strategy proposed by [35,36]. Suppose that three calibration sets, namely *C1*, *C2* and *C3*, on a response $y$ and a predictor $x$ are available, and that a linear least-squares (LS) fitting is to be carried out on each of them. The comparison of those three lines at once to check if they can be described by the same mathematical model can be performed by posing the following regression equation,

$$y = \beta_0 + \beta_1 x + a_0 z_1 + a_1 z_1 x + a_2 z_2 + a_3 z_2 x + \varepsilon \tag{5.3}$$

being $\varepsilon$ the independent, equal and normally distributed residual error. Two indicator variables, $z_1$ and $z_2$, appear in Eq. (5.3) to define the three separate functional models for *C1*, *C2* and *C3* by letting $z_1 = z_2 = 0$ in the first case, $z_1 = 1$ and $z_2 = 0$ in the second case, and $z_1 = 0$ and $z_2 = 1$ in the third case, respectively; that is,

*C1* model function: $y = \beta_0 + \beta_1 x$ (5.4)

*C2* model function: $y = (\beta_0 + a_0) + (\beta_1 + a_1) x$ (5.5)

*C3* model function: $y = (\beta_0 + a_2) + (\beta_1 + a_3) x$ (5.6)

On the one hand, the coefficients $a_0$ and $a_1$, and $a_2$ and $a_3$ on the other hand, represent the changes needed to get from the *C1* functional model to the *C2* or the *C3* one, respectively. So, if, after hypothesis testing, it is concluded that $a_0 = a_1 = a_2 = a_3 = 0$, it can thus be stated that the three models under comparison are the same.

This strategy to evaluate the equality of several calibration curves is of great application in chemical analysis, especially to check if matrix effects are present [37], which could

adversely affect both identification and quantification of analytes if not properly corrected [38]. In GC determinations, a matrix-induced chromatographic response enhancement effect often occurs: in solvent standards without sample matrix present, only the analytes fill the active sites in the system (injector, column, detector), which reduce the percentage of injected molecules eventually detected. However, in complex injected extracts, the active sites are filled predominantly by matrix components, thereby increasing efficiency of analyte transfer through the GC system to the detector [39]. For the study of the existence of matrix effects in this work (see Section 5.2.5.3), calibration set *C2* was made up of matrix-matched standards, whereas *C1* and *C3* consisted of two arrays of standards of similar concentrations in the solvent. So two solvent calibration sets were selected to check the extent of the possible matrix effects by comparing the regression model estimated before (*C1* set) and after (*C*3 set) analysing the matrix-matched standards (*C2* set).

## 5.2.5. Results and discussion

### *5.2.5.1. Validation of the analytical procedure*

#### 5.2.5.1.1. Unequivocal identification

The requirements laid down in EUR 24105 EN [21] for the unequivocal identification of the analytes were followed in this work. To establish the permitted tolerance intervals, six reference standards were prepared and analysed. Table 5.1 (first row, third column) shows the concentration ranges of these standards. Three of them contained the IS at a fixed concentration and the analytes at three concentration levels, while the rest contained the IS at three concentration levels and the analytes at a fixed concentration. A solvent blank (only hexane) as well as three system blanks (no liquid) injected at the beginning, the middle and the end of these analyses were measured to test the performance of the GC-MS system. After baseline correction, the resulting ten chromatograms were equally fragmented around the retention time of each analyte and the fragments (data matrices) related to the same compound were arranged together into a data tensor **X**. The dimensions of these tensors are specified in Table 5.1 (first row, columns 4–8): it must be noticed that a joint array was built for DiBP and DiBP-$d_4$ peaks.

A PARAFAC decomposition of each tensor was then performed. As can be seen in Table 5.2, a one-factor unconstrained model was needed for BHT, BP and DEHA. The PARAFAC model for DiBP and the internal standard required three factors, being DiBP and DiBP-$d_4$ the first and second factors, respectively, while the third factor was an interferent that eluted before DiBP-$d_4$. The PARAFAC model for DiNP also yielded three factors.

**Table 5.2.** Characteristics of the PARAFAC decomposition performed with the tensor built with the reference samples for each analyte (see Table 5.1, first row) and tolerance intervals for the relative retention time and the diagnostic ions in each case. The base peak is in bold.

| Analyte | PARAFAC model | Retention time | | | Diagnostic ions | | | |
|---|---|---|---|---|---|---|---|---|
| | | $t_R$ (min) | Relative $t_R$ | Tolerance interval | $m/z$ ratio | Spectral loading | Relative abundance (%) | Tolerance interval (%) |
| BHT | 1 factor Unconstrained model Explained variance: 99.85% | 8.401 | 0.808 | (0.804–0.813) | 91 | $6.85\ 10^{-2}$ | 7.13 | (3.57–10.70) |
| | | | | | 145 | $1.14\ 10^{-1}$ | 11.85 | (9.48–14.22) |
| | | | | | 177 | $7.95\ 10^{-2}$ | 8.27 | (4.14–12.41) |
| | | | | | **205** | $9.61\ 10^{-1}$ | 100.00 | — |
| | | | | | 220 | $2.28\ 10^{-1}$ | 23.73 | (20.17–27.29) |
| BP | 1 factor Unconstrained model Explained variance: 92.72% | 9.177 | 0.883 | (0.879–0.888) | 51 | $1.62\ 10^{-1}$ | 21.03 | (17.88–24.18) |
| | | | | | 77 | $4.74\ 10^{-1}$ | 61.70 | (55.53–67.87) |
| | | | | | **105** | $7.68\ 10^{-1}$ | 100.00 | — |
| | | | | | 152 | $3.05\ 10^{-2}$ | 3.97 | (1.99–5.96) |
| | | | | | 182 | $3.97\ 10^{-1}$ | 51.68 | (46.51–56.85) |
| DiBP-$d_4$ | 3 factors (Factor 2: DiBP-$d_4$) Non-negativity constraint in modes 1 and 2 Explained variance: 99.36% CORCONDIA: 98% | 10.391 | 1.000 | — | 80 | $6.49\ 10^{-2}$ | 6.52 | (3.26–9.78) |
| | | | | | **153** | $9.96\ 10^{-1}$ | 100.00 | — |
| | | | | | 171 | $2.43\ 10^{-2}$ | 2.44 | (1.22–3.66) |
| | | | | | 209 | $1.67\ 10^{-2}$ | 1.68 | (0.84–2.52) |
| | | | | | 227 | $5.25\ 10^{-2}$ | 5.27 | (2.64–7.91) |
| DiBP | 3 factors (Factor 1: DiBP) Non-negativity constraint in modes 1 and 2 Explained variance: 99.36% CORCONDIA: 98% | 10.397 | 1.001 | (0.996–1.006) | 104 | $7.86\ 10^{-2}$ | 7.90 | (3.95–11.85) |
| | | | | | **149** | $9.95\ 10^{-1}$ | 100.00 | — |
| | | | | | 167 | $2.85\ 10^{-2}$ | 2.86 | (1.43–4.29) |
| | | | | | 205 | $1.41\ 10^{-2}$ | 1.42 | (0.71–2.13) |
| | | | | | 223 | $5.50\ 10^{-2}$ | 5.53 | (2.77–8.30) |

**Table 5.2. (cont.)**

| Analyte | PARAFAC model | Retention time | | | Diagnostic ions | | | |
|---|---|---|---|---|---|---|---|---|
| | | $t_R$ (min) | Relative $t_R$ | Tolerance interval | $m/z$ ratio | Spectral loading | Relative abundance (%) | Tolerance interval (%) |
| DEHA | 1 factor<br>Unconstrained model<br>Explained variance: 98.47% | 13.319 | 1.282 | (1.275–1.288) | 112 | $3.57\ 10^{-1}$ | 38.82 | (33.00–44.64) |
| | | | | | **129** | $9.21\ 10^{-1}$ | 100.00 | — |
| | | | | | 147 | $1.51\ 10^{-1}$ | 16.36 | (13.09–19.63) |
| | | | | | 241 | $4.11\ 10^{-2}$ | 4.46 | (2.23–6.69) |
| | | | | | 259 | $1.78\ 10^{-2}$ | 1.93 | (0.97–2.90) |
| DiNP | 3 factors (Factor 1: DiNP)<br>Non-negativity constraint in the three modes<br>Explained variance: 99.62%<br>CORCONDIA: 95% | —[a] | —[a] | —[a] | 57 | $3.64\ 10^{-1}$ | 39.71 | (33.75–45.67) |
| | | | | | 127 | $9.54\ 10^{-2}$ | 10.42 | (8.34–12.50) |
| | | | | | **149** | $9.16\ 10^{-1}$ | 100.00 | — |
| | | | | | 167 | $8.48\ 10^{-2}$ | 9.26 | (4.63–13.89) |
| | | | | | 275 | $5.70\ 10^{-3}$ | 0.62 | (0.31–0.93) |
| | | | | | 293 | $1.15\ 10^{-1}$ | 12.61 | (10.09–15.13) |

[a] It is not possible to establish a retention time for DiNP.

These PARAFAC decompositions provided the unique chromatographic and spectral profiles of every compound that are common to all the samples related to it. Therefore, the retention time for each analyte obtained through the chromatographic profile enabled to calculate the tolerances for its relative retention time (the ratio of the chromatographic retention time of the analyte to that of the internal standard), which appears in the fourth column of Table 5.2. The tolerance intervals for the relative retention time of all the analytes are collected in the fifth column of this table. In the case of DiNP, it was not possible to establish its retention time since finger peaks appeared in its chromatogram. In addition, the spectral loadings were used to calculate the relative abundances of each $m/z$ ion with regard to the base peak and thus determine the tolerance intervals for the relative ion abundances (see Table 5.2, columns 6–9). Both kinds of intervals would be now used as reference to confirm the presence of the corresponding compound in every sample.

5.2.5.1.2. Calibration

Twelve calibration standards were prepared within the concentration ranges detailed in Table 5.1 (second row, third column) and analysed in duplicate. The ubiquity of DiBP in the previously analysed solvent blanks made its quantification impossible because this analyte appeared in a different quantity in each solvent blank injected. So, the concentration of DiBP was fixed at 25 $\mu g\ L^{-1}$ in these calibration standards, which is below the generic SML. To carry out the calibration based on PARAFAC, the data tensors were built with these standards together with 8 system blanks, a solvent blank without IS and 22 solvent blanks with IS. The dimensions of these three-way tensors are given in Table 5.1 (second row) for each analyte, while the features of the model estimated from the PARAFAC decomposition of each tensor are included in the second column of Table 5.3.

For all the analytes, it was checked if their relative retention times obtained through the chromatographic profile and the relative abundances calculated with the loadings of the spectral profile for each diagnostic ion (see Table 5.3, third column) were within the corresponding tolerance intervals established previously (in Table 5.2). Only the relative abundance of the $m/z$ ratio 57 for DiNP lay outside its corresponding tolerance interval. However, 5 $m/z$ ratios met the identification conditions for this analyte. So, it could be stated that the presence of all compounds was unequivocally confirmed.

Once the sample loadings for each analyte had been standardized by dividing each of them by that of the internal standard, calibration lines "*Standardized sample loading versus True concentration*" were performed with the 24 standards. Table 5.3 (fourth column) shows the parameters of the LS regression models estimated for each analyte. A second-degree polynomial model was considered for BP and DiNP since a lack of fit was concluded in the

linear model at 95% confidence level. The regression models were significant in all cases and no outlier data were detected. The highest mean of the absolute value of the relative errors in calibration was 6.71% ($n$ = 22) for DEHA, while the lowest value was obtained for DiNP: 2.33% ($n$ = 22). In addition, the intercept of the corresponding accuracy line ("*Estimated concentration versus True concentration*", see the equations in the sixth column of Table 5.3) was equal to 0 and its slope was equal to 1 at a significance level of 5% in all cases. Therefore, trueness was verified for all the analytes at a 95% confidence level. The decision limit ($CC\alpha$) and the detection capability ($CC\beta$) values are collected in the last two columns of Table 5.3. $CC\beta$ values were: 2.3 μg $L^{-1}$ for BHT, 2.7 μg $L^{-1}$ for BP, 3.6 μg $L^{-1}$ for DEHA and 101.7 μg $L^{-1}$ for DiNP (for $\alpha = \beta = 0.05$ and two replicates). The details of the procedure to obtain $CC\beta$ from three-way data can be found in [40].

**Table 5.3.** Characteristics of the PARAFAC models obtained in the calibration step, identification of every analyte according to the regulation and parameters of the calibration line "*Standardized sample loading vs True concentration*" and of the accuracy line. Decision limit ($CC\alpha$) and capability of detection ($CC\beta$) at $x_0 = 0$ ($\alpha = \beta = 0.05$). The third column shows the relative retention time ($t_{R,rel}$) and the relative abundances (in brackets) estimated from the spectral loadings for each diagnostic ion. In bold, non-compliant $m/z$ ratios.

| Analyte | PARAFAC model | Identification ($t_{R,rel}$ and $m/z$ relative abundance) | Calibration line | | Accuracy line | $CC\alpha$ (µg L$^{-1}$) | $CC\beta$[c] (µg L$^{-1}$) |
|---|---|---|---|---|---|---|---|
| | | | Model ($R^2$, $s_{yx}$) | Error (%)[b] | Model ($R^2$, $s_{yx}$) | | |
| BHT | 1 factor<br>Unconstrained model<br>Explained variance: 99.70% | $t_{R,rel}$ = 0.808<br>91 (7.12%)<br>145 (11.79%)<br>177 (8.31%)<br>205 (100%)<br>220 (23.82%) | $y = 9.59\ 10^{-2} + 4.68\ 10^{-2}\,x$<br>(99.75%, 0.04) | 5.37<br>($n$ = 22) | $y = -8.89\ 10^{-5} + 1.00\,x$<br>(99.75%, 0.86) | 1.16 | 2.30 |
| BP | 1 factor<br>Unconstrained model<br>Explained variance: 96.96% | $t_{R,rel}$ = 0.884<br>51 (19.60%)<br>77 (59.12%)<br>105 (100%)<br>152 (4.21%)<br>182 (55.54%) | $y = -7.27\ 10^{-2} + 4\ 10^{-2}\,x$<br>$+ 2.98\ 10^{-4}\,x^2$<br>(99.74%, 0.049) | 5.72<br>($n$ = 22) | $y = 4.82\ 10^{-3} + 1.00\,x$<br>(99.67%, 0.99) | 1.34 | 2.66 |
| DiBP-$d_4$ | 3 factors (Factor 2: DiBP-$d_4$)<br>Non-negativity constraint in modes 1 and 2<br>Explained variance: 99.39%<br>CORCONDIA: 98% | $t_{R,rel}$ = 1.000<br>80 (6.12%)<br>153 (100%)<br>171 (2.62%)<br>209 (1.57%)<br>227 (5.16%) | Internal standard (IS) | | | | |
| DiBP | 3 factors (Factor 1: DiBP)<br>Non-negativity constraint in modes 1 and 2<br>Explained variance: 99.39%<br>CORCONDIA: 98% | $t_{R,rel}$ = 1.001<br>104 (7.83%)<br>149 (100%)<br>167 (2.84%)<br>205 (1.47%)<br>223 (5.33%) | Not quantified | | | | |

**Table 5.3. (cont.)**

| Analyte | PARAFAC model | Identification ($t_{R,rel}$ and $m/z$ relative abundance) | Calibration line | | Accuracy line | $CC\alpha$ (μg $L^{-1}$) | $CC\beta$[c] (μg $L^{-1}$) |
|---|---|---|---|---|---|---|---|
| | | | Model ($R^2$, $s_{yx}$) | Error (%)[b] | Model ($R^2$, $s_{yx}$) | | |
| DEHA | 1 factor<br>Unconstrained model<br>Explained variance: 99.45% | $t_{R,rel}$ = 1.282<br>112 (35.92%)<br>129 (100%)<br>147 (16.04%)<br>241 (4.48%)<br>259 (1.89%) | $y = 1.12\ 10^{-2} + 3.09\ 10^{-2}\,x$<br>(99.75%, 0.04) | 6.71<br>($n$ = 22) | $y = -7.54\ 10^{-4} + 1.00\,x$<br>(99.75%, 1.36) | 1.84 | 3.65 |
| DiNP | 2 factors (Factor 1: DiNP)<br>Non-negativity constraint in modes 1 and 2<br>Explained variance: 99.75%<br>CORCONDIA: 100% | $t_{R,rel}$ = —[a]<br>**57** (32.52%)<br>127 (10.34%)<br>149 (100%)<br>167 (9.48%)<br>275 (0.63%)<br>293 (12.98%) | $y = -1.96\ 10^{-2} + 5.81\ 10^{-4}x$<br>$+ 8.15\ 10^{-8}\,x^2$<br>(99.83%, 3.34 $10^{-2}$) | 2.33<br>($n$ = 22) | $y = 3.91\ 10^{-2} + 1.00\,x$<br>(99.85%, 36.46) | 51.42 | 101.70 |

[a] It is not possible to establish a retention time for DiNP.

[b] Mean of the absolute value of the relative error in calibration.

[c] The first standard analysed was at a concentration of 2.5 μg $L^{-1}$ for BHT, 4 μg $L^{-1}$ for DEHA and 500 μg $L^{-1}$ for DiNP which corresponds to a $\beta$ value of 0.028, 0.025 and lower than $10^{-15}$, respectively.

By way of example, Figure 5.1 shows the total ion chromatogram (TIC) of the calibration standard injected at the highest concentration. As can be seen in this figure, finger peaks appeared in the chromatogram for DiNP because of an array of possible $C_9$ isomers, as commented in Section 5.2.2. Therefore, it was not possible to establish a retention time for this analyte, as can be seen in Tables 5.2 and 5.3. In addition, this analyte appeared as a broad peak, which took about 2 min to elute (see Figure 5.1). This is the reason why a big number of scans (823) were considered for this analyte in this work (see Table 5.1, column 8).

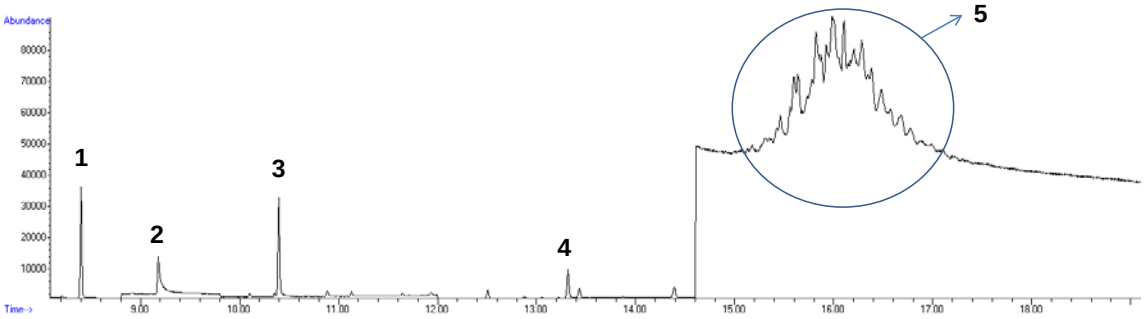


**Figure 5.1.** Total ion chromatogram (TIC) obtained from the injection of a calibration standard containing 50 μg $L^{-1}$ of BHT and BP, 25 μg $L^{-1}$ of DiBP and DiBP-$d_4$, 80 μg $L^{-1}$ of DEHA and 3000 μg $L^{-1}$ of DiNP. Peak labels: **1**, BHT; **2**, BP, **3**, DiBP and DiBP-$d_4$; **4**, DEHA; **5**, DiNP.

The quantification of DiNP is usually performed from the sum of peak areas corresponding to the different isomers [41]. But, as the signal for this analyte has such a complex shape, the results could vary significantly depending on the skill and experience of the chromatographist. However, a more accurate choice, namely, a calibration based on PARAFAC for this special case was used in this work, as explained previously. Figure 5.2 shows the loadings of the PARAFAC decomposition of the tensor **X** (823 × 6 × 55) built for DiNP in the calibration step (see its features in Table 5.3). The loadings of the chromatographic profile (Figure 5.2.a) for the factor identified as DiNP showed its typical finger peaks, whereas the ones for the second factor (the baseline) decreased with the elution time. So, there was still baseline noise together with this analyte although a previous correction had been made for those chromatograms. In Figure 5.2.b, it can be observed that $m/z$ ratio 57 was characteristic of the baseline and shared with DiNP; on the other hand, $m/z$ ratio 149 was the base peak for DiNP, like in most EI spectra of phthalates. The loadings of the sample profile (Figure 5.2.c) for DiNP increased with the concentration of the calibration samples, whereas they were zero for the system and solvent blanks (blue circles near zero). However, the sample loadings of the baseline factor remained nearly constant for the calibration and blank samples and these values were higher than the ones

for the system blank samples (sample numbers 1, 2, 16, 23, 32, 41, 50 and 55, red circles near zero). So, this PARAFAC model was coherent with the experimental knowledge and the calibration results for this analyte were satisfactory. Therefore, it can be concluded that PARAFAC succeeded in the determination of DiNP since no peak areas but loadings were considered. As far as the authors are aware, this is the first time PARAFAC has been used to deal with finger-peak chromatographic signals.

#### *5.2.5.2. Ubiquity of DiBP in the laboratory: strategy to assess if the amount of DiBP present in a sample is higher than the blank level*

As previously mentioned, a changeable presence of DiBP was observed from the first stages of this work, even when solvent blanks were analysed. This fact did not allow the quantification of this compound to be carried out by means of a regression line, because its blank concentration varied significantly over time, even throughout a routine analytical sequence. Therefore, a proper calibration model could not be estimated without having assessed that blank level properly first.

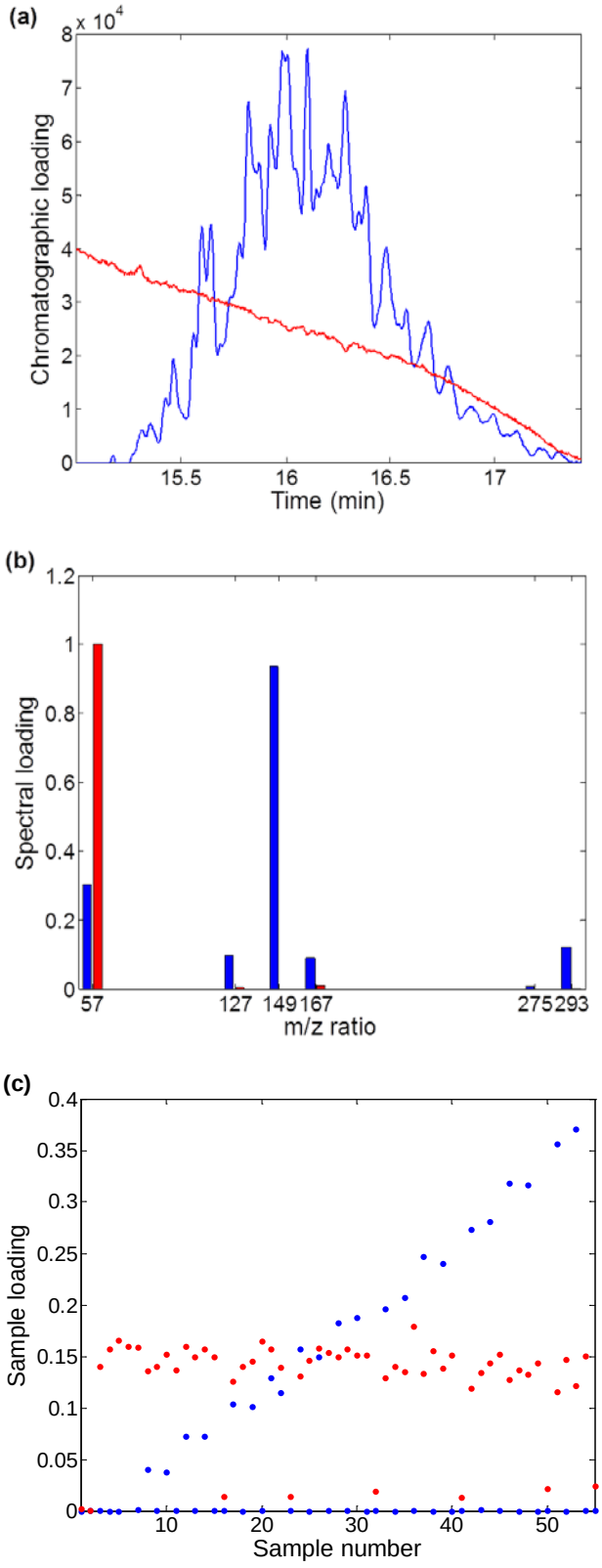


**Figure 5.2.** Loadings of the **(a)** chromatographic profile, **(b)** spectral profile, **(c)** sample profile of the PARAFAC model with two factors built with the data tensor of DiNP for the calibration step (blue: DiNP; red: baseline).

Despite this, as commented in Section 5.2.2, a comparison between the probability distributions of the level of DiBP in solvent blanks and in solvent standards at a fixed concentration could be made. The probability distribution related to the amount of DiBP in solvent blanks would account for the non-constant pattern of this analyte in the laboratory environment. That comparison would be posed through a two-sample hypothesis test; the critical value of the standardized sample loading of DiBP beyond which the concentration of this analyte in a solvent standard would be statistically greater than that in the laboratory environment could be estimated after setting the values of the $\alpha$ and $\beta$ errors that would be taken.

5.2.5.2.1. Data structure

To estimate the two probability distributions for this study, 50 solvent blanks on the one hand and 44 solvent standards with 25 μg $L^{-1}$ of DiBP on the other hand were analysed, both sets containing DiBP-$d_4$ at 25 μg $L^{-1}$. The number of objects in each distribution was selected to prevent the insensitivity of the goodness-of-fit hypothesis tests. According to [42], for the Kolmogorov-Smirnov test, letting the distance between the two experimental distributions be 0.3 at the most, a size of about 50 objects must be considered for each distribution in order to ensure that the power of the test is 90%. The rest of the analytes were also included in these solvent standards to avoid significant differences in the signals of DiBP and DiBP-$d_4$ due to the absence of some compounds in the GC liner. Otherwise, it may result in a different adsorption/evaporation behaviour of DiBP and DiBP-$d_4$ in the inlet, which could lead to results and conclusions that may not be statistically comparable to those obtained in other steps of this work. The concentration ranges of BHT, BP, DEHA and DiNP in these standards are listed in Table 5.1 (third row, third column).

Besides the assessment on the blank amount of DiBP, another study was carried out in parallel. Some rubber latex nipples used to make up the reference standard solutions to the ring graduation mark on the volumetric flask were suspected of being one of the sources of cross-contamination of DiBP. To test that and the possible leaching of the other four compounds from the polymer, two different nipples (referred as *N1* and *N2*) were used to pipette hexane into 5-mL flasks to prepare solvent blanks; the IS at a final concentration of 25 μg $L^{-1}$ had been previously added. Four situations during this stage were simulated, each in duplicate (replicates referred as *R1* and *R2*): on the one hand, solvent (hexane) was not allowed to reach up and be in contact with the nipple polymer (samples named as *N1_0_R1* and *N1_0_R2* for nipple *N1*, while *N2_0_R1* and *N2_0_R2* for nipple *N2*), and on the other hand, there were one, two and three forced contacts, respectively (samples named as *N1_1_R1* and *N1_1_R2*, *N1_2_R1* and *N1_2_R2*, *N1_3_R1* and *N1_3_R2* for nipple *N1* at each of these three situations; the same code was used for the blanks from

nipple *N2*). So a total of 16 solvent blanks prepared in this way resulted, eight for each nipple.

In addition to the 50 solvent blanks and the 44 solvent standards for the study on DiBP and the 16 nipple blanks, both system blanks and solvent blanks without IS were injected to control the performance of the GC-MS equipment. It must be pointed out that no nipples were used in the volume adjustment of any solutions, except for those specifically indicated, either in this stage of the work or in the following. Because of the impossibility to carry out the resulting 130 analyses within the same day, the sequence was performed in two days (55 analyses on the first day, 75 analyses on the second one). After background subtraction, five data tensors whose dimensions are collected in Table 5.1 (third row, columns 4–8) were built.

5.2.5.2.2. Data analysis

The PARAFAC decomposition of the tensor related to BHT showed the existence of one single factor unequivocally linked to this analyte (no constraints imposed, explained variance of 99.59%, no outliers detected regarding the threshold values of $Q$ and $T^2$ indices at a 99% confidence level).

An unconstrained one-factor model was also obtained for BP (explained variance of 92.41%, no outliers), whose presence was also guaranteed in terms of retention time and mass spectrum.

As for the joint three-way tensor for DiBP and DiBP-$d_4$, its PARAFAC decomposition yielded a three-factor model (CORCONDIA of 98%, explained variance of 98.90%, 11 outliers finally removed), where the chromatographic and spectral ways had been non-negativity-constrained. Factor 1 in this model (in dark blue in Figures 5.3.a, 5.3.b and 5.3.c) was unequivocally associated to DiBP and Factor 2 (in light green in Figures 5.3.a, 5.3.b and 5.3.c) was to DiBP-$d_4$, whereas Factor 3 (in red in Figures 5.3.a, 5.3.b and 5.3.c) was considered an unidentified interferent eluting near the beginning of the DiBP-$d_4$ peak. Figure 5.3.c reveals that the sample loadings of DiBP for the nipple blanks were higher than those for solvent blanks, so the suspicions about the leaching of DiBP out from the nipples evaluated were reasonable.

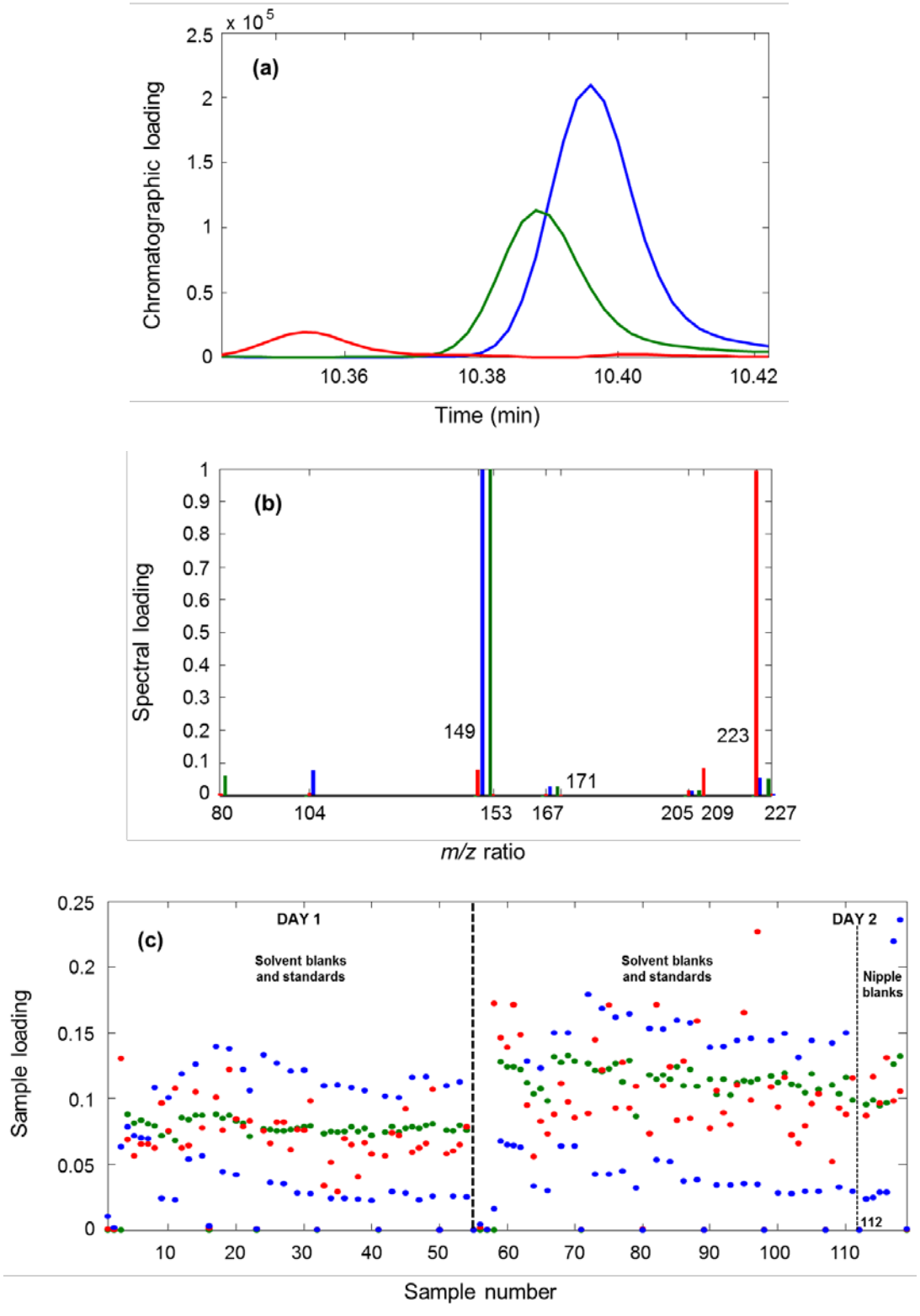


**Figure 5.3** PARAFAC decomposition of the common tensor for DiBP and DiBP-$d_4$ after the removal of the outliers detected. Loadings of the **(a)** chromatographic, **(b)** spectral and **(c)** sample profiles of the resulting PARAFAC model; Factor 1 (DiBP): dark blue; Factor 2 (DiBP-$d_4$): light green; Factor 3: red. The index of sample number 112 is indicated in Figure 5.3.c for an easier understanding of the text.

On the other hand, Figures 5.4.a (solvent blanks) and 5.4.b (solvent standards at 25 μg $L^{-1}$) display the values of the standardized sample loadings of DiBP for the two sets under study. It can be seen that, although greater values of the sample loadings of both DiBP and especially DiBP-$d_4$ were obtained on the second day of analysis (see blue and green circles in Figure 5.3.c), the standardization procedure succeeded in compensating for the variations responsible for that shift.

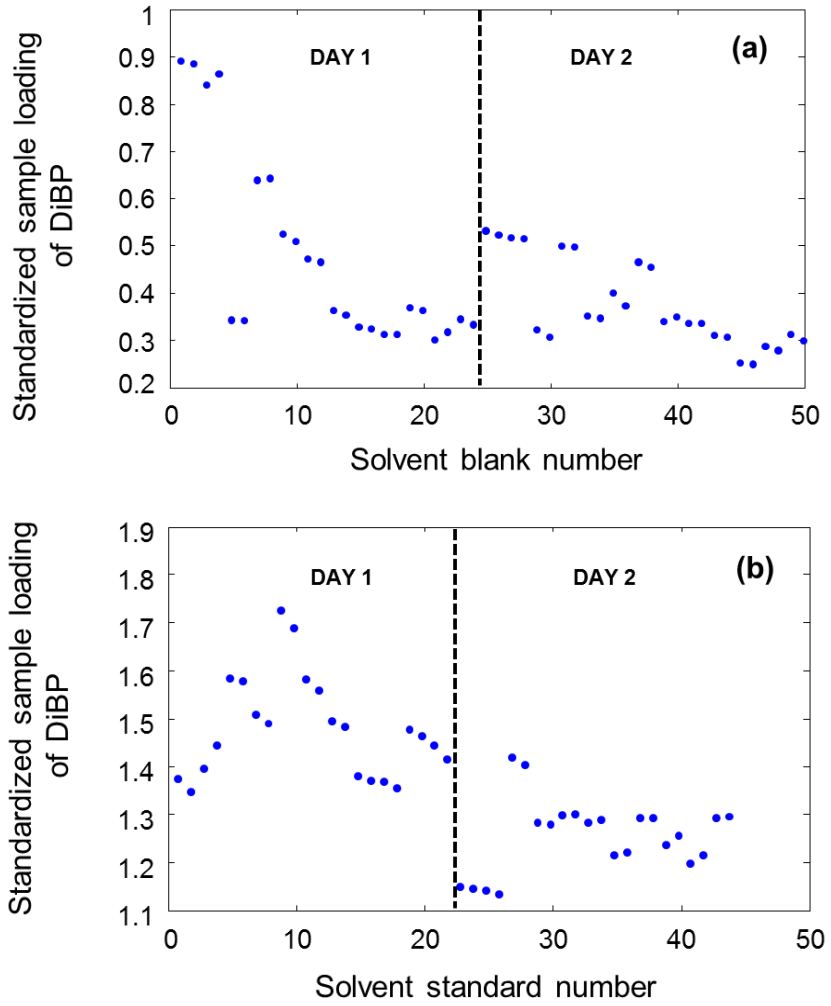


**Figure 5.4.** Standardized sample loadings of DiBP for **(a)** the solvent blank class and **(b)** the solvent standard class.

The 11 outliers detected in the tensor for DiBP and DiBP-$d_4$ that contained the initial 130 samples corresponded to 10 nipple blanks and 1 system blank; to be precise, as Figures 5.5.a and 5.5.b show, only the four solvent blanks where hexane had had no contact with the nipple (samples *N1_0_R1*, *N1_0_R2*, *N2_0_R1* and *N2_0_R2*, which correspond to sample numbers 113, 114, 115 and 116, respectively, in Figure 5.3.c) and those prepared after 1 contact of hexane with nipple *N2* (samples *N2_1_R1* and *N2_1_R2*; sample numbers 117 and 118 in Figure 5.3.c) could not be considered outliers. The rest of the

nipple blanks lay out of the plane defined by the threshold values of $Q$ and $T^2$ indices at a 99% confidence level; this meant that all these outliers were placed far away both from the new space spanned by the three factors of the PARAFAC model in the sample mode and from the centroid of the samples. By having a closer look at Figure 5.5.a, the values of $Q$ and $T^2$ indices for the nipple blanks revealed important differences among these samples, being each class of nipple blanks (regarding the number of contacts with hexane) located separately in the plot. This was confirmed by the fact that no constant values of the sample loadings of DiBP-$d_4$ were obtained for the 16 nipple blanks: the higher the number of contacts between hexane and the nipple polymer, the greater the sample loading of the IS, as shown in Figure 5.3.c when comparing the results for the zero-contact blanks (sample numbers 113 to 116) and those for the one-contact blanks for nipple *N2* (sample numbers 117 and 118). This fact could be responsible for those outlier nipple blanks to be different from the rest of the samples in the tensor in terms of $Q$ and $T^2$ values, and led to think of the possibility of matrix effects that would account for the signal enhancement observed throughout the analysis of the nipple blanks. This hypothesis is assessed in Section 5.2.5.3.

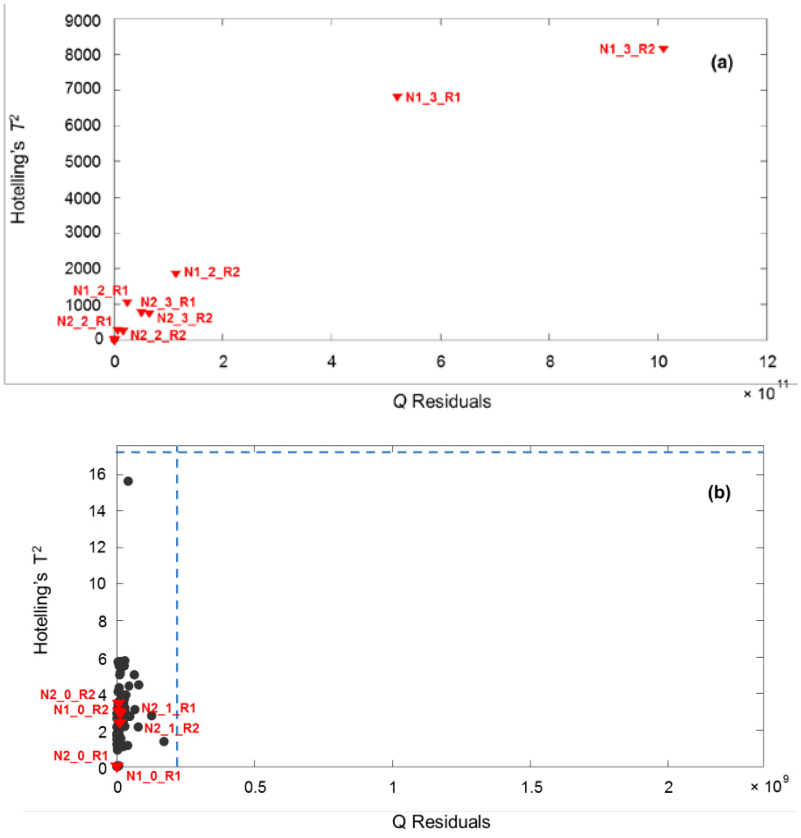


**Figure 5.5.** **(a)** Plot of the indices Hotelling's $T^2$ versus $Q$ of the 119 samples that make up the joint tensor for DiBP and DiBP-$d_4$; the nipple blanks are marked in red triangles, while the rest of the samples are in grey circles. **(b)** Enlargement of Figure 5.5.a to show the threshold values of $Q$ and Hotelling's $T^2$ at a 99% confidence level (in dashed blue lines).

Regarding DEHA, a two-factor PARAFAC model was estimated (CORCONDIA of 100%, explained variance of 98.62%, no outliers found) after a non-negativity constraint had been laid down on the chromatographic, spectral and sample ways. Factor 1 matched DEHA unequivocally, whereas Factor 2 was attributed to an unknown interferent that mainly eluted in the nipple blanks.

Lastly, the PARAFAC decomposition of the DiNP tensor revealed, through a two-factor non-negativity constrained model (CORCONDIA of 100%, explained variance of 98.73%, no outliers rejected), that the baseline signal (Factor 2) could be totally set apart from the chemical information on DiNP (Factor 1), which enabled its unequivocal identification.

Figure 5.6 depicts the values of the standardized sample loadings of the five analytes in the 50 solvent blanks, the 44 solvent standards and the 6 nipple blanks eventually considered. It is clear from this plot that BP, DEHA and DiNP were not found either in the solvent blanks (no evidence of the presence of these three compounds in the lab environment) or in the nipple blanks (no leaching of these substances from the polymer). As for BHT, a slight amount of this analyte seemed to appear on both the solvent and nipple blanks, but displaying a uniform pattern: in fact, the values of the standardized sample loadings for BHT in both kinds of blanks varied between 0.09 and 0.12, being the mean and the standard deviation equal to 0.09 and 0.02, respectively. The concentration of BHT corresponding to that mean value was determined from the regression line "*Standardized sample loading versus True concentration*" estimated in the study in Section 5.2.5.1. As the result was $-0.06 \pm 1.91$ μg $L^{-1}$, it could be concluded that BHT was not present in the lab environment and did not leach out of the nipple polymer either, because its blank concentration after both studies was statistically equal to 0. However, as pictured in Figure 5.6, the situation for DiBP was quite different, since non-zero standardized sample loadings were obtained for both the solvent and nipple blanks; furthermore, as stated above, due to the variability in those values, the estimation of a trustworthy calibration line was not possible, and the strategy based on the assessment of the probability distributions was thus designed.

The distributions coming from the two classes, namely, *Distribution 1* for the solvent blanks and *Distribution 2* for the solvent standards at 25 μg $L^{-1}$, were compared by means of a Kolmogorov-Smirnov test ($H_0$: *Distribution 1 = Distribution 2*; $H_a$: *Distribution 1 ≠ Distribution 2*), which is performed by computing the maximum distance between the cumulative distributions of the two sets. Since the *p*-value for this test was $4.5\ 10^{-9}$, it could be concluded that there was a statistically significant difference between the two distributions at a 95% confidence level. That is, both data sets could not be modelled by the same type of probability density function, which meant that the pattern of DiBP was completely

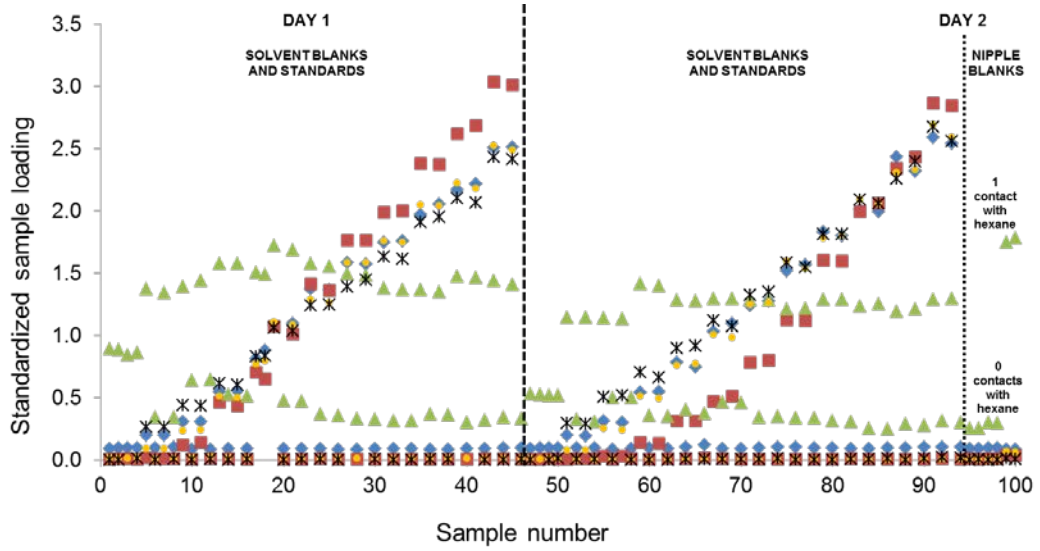


**Figure 5.6.** Values of the standardized sample loadings of the five analytes. BHT: light blue diamonds; BP: red squares; DiBP: light green triangles; DEHA: yellow circles; DiNP: black asterisks.

different at these two concentration levels. The standardized sample loadings from the solvent blanks (*Distribution 1*) could not be adequately fitted to a normal distribution, but to a three-parameter lognormal one with mean 0.43, variance 0.03 and lower threshold 0.23. However, the data from the solvent standards (*Distribution 2*) were contrasted to follow a normal distribution $N(1.37, 0.02)$. The plot of both distributions is shown in Figure 5.7.a.

To decide when the amount of DiBP in a sample could be considered as statistically different (greater) than that at the blank level, a hypothesis test was posed on these two probability distributions. This situation was formally expressed as

$H_0$: DiBP is present at the blank amount: the sample comes from *Distribution 1*. (5.7)

$H_a$: DiBP is present at an amount greater than that in the laboratory environment: the sample comes from *Distribution 2*. (5.8)

The critical standardized sample loading of DiBP that would mean the lowest value of this variable beyond which the null hypothesis $H_0$ in Eq. (5.7) would be rejected was computed after having specified the probabilities of false non-compliance ($\alpha$) and of false compliance ($\beta$) taken when making a decision on this issue. $\alpha$ and $\beta$ values are not independent, but change in opposite directions, as can be seen in Figure 5.7.b, where the operating-characteristic curve of this hypothesis test is displayed. From all the pairs ($\alpha$,$\beta$) drawing this curve, the situation when $\alpha = \beta = 0.012$ was selected (marked with a yellow triangle in Figure 5.7.b), since the probabilities of making a type I error ($\alpha$) and a type II error ($\beta$)

were both low and equal. The critical value (standardized sample loading) of DiBP associated was 1.04, which is also represented in Figure 5.7.a with a dotted vertical line. This means that any sample, after performing PARAFAC or PARAFAC2 and confirming the presence of DiBP unequivocally, that has got a standardized sample loading of this analyte higher than 1.04 will contain DiBP in an amount significantly greater than that considered as blank in the lab environment, being probabilities $\alpha$ and $\beta$ equal to 0.012.

As for the amount of DiBP released from the nipples into the six remaining solvent blanks once outliers had been rejected, only samples *N2_1* (replicates *R1* and *R2*) were concluded to have DiBP at a level higher than the blank one, since their standardized sample loadings were 1.75 and 1.78, respectively. On the contrary, the values of this variable ranged from 0.25 to 0.30 for the four nipple blanks prepared without allowing any contacts between the nipple polymer and hexane.

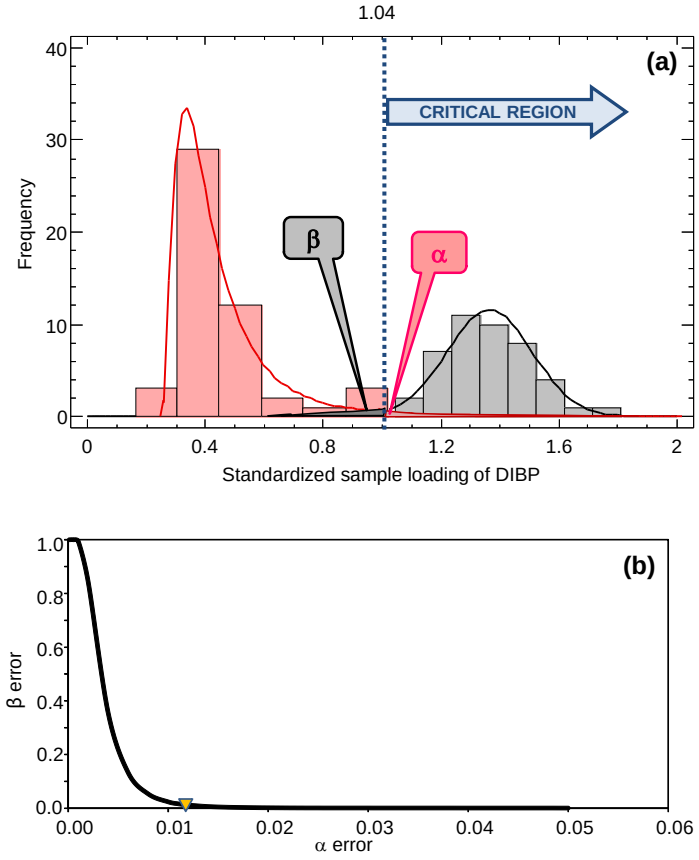


**Figure 5.7.** **(a)** Frequency histogram and probability distribution for the two sets under study: solvent blanks (*Distribution 1*, in magenta) and solvent standards at 25 μg $L^{-1}$ of DiBP (*Distribution 2*, in black). **(b)** Operating characteristic curve of the hypothesis test $H_0$: The sample comes from *Distribution 1* against $H_a$: The sample comes from *Distribution 2*. The values of the $\alpha$ and $\beta$ errors finally considered are represented by the point marked with a yellow triangle.

*5.2.5.3. Evaluation of possible matrix effects*

The existence of matrix effects on the GC signals due to the presence of sample matrix when injecting the blanks prepared from nipples was assessed. The nipple that showed the highest level of DiBP released in the previous stage, namely, nipple *N1* (see Figure 5.5.a), was selected to perform this study. The calibration models estimated for the five analytes from three calibration sets were compared. One of them (*C2* set) was made up of matrix-matched standards prepared by simulating two contacts of liquid hexane with the nipple during preparation; from the four concentration levels considered for the standards, the three of them different from zero were prepared and analysed in duplicate, whereas six blanks were prepared. The other two calibration sets consisted of solvent standards at the same concentration levels and with the same number of replicates as the matrix-matched series. These two solvent-standard sequences were analysed before (*C1* set) and after (*C3* set) the matrix-matched one, respectively, to test the extent of the possible matrix effects. The concentration ranges considered are specified in Table 5.1 (fourth row, third column). A calibration model could be estimated for DiBP this time, since the lowest standard prepared for this study was at 25 μg $L^{-1}$, which had been proved to be statistically different from the blank concentration ($\alpha = \beta = 0.012$) in the previous stage of the work (see Section 5.2.5.2).

After GC-MS analysis, five data tensors, each for every analyte, were built from all these standards, together with solvent and system blanks without containing IS (see Table 5.1, fourth row); as ever, the data from DiBP and the IS were included in the same array. The features of the model estimated from the PARAFAC (or PARAFAC2) decomposition of every tensor are listed in Table 5.4 (second column). By comparing the chromatographic and spectral profiles of the factor associated to every analyte with those used as reference (Table 5.2), it could be stated that all compounds were unequivocally identified.

As can be seen in Table 5.4, peak shifts forced to a PARAFAC2 decomposition of the array for DiBP and DiBP-$d_4$. The loadings in the third way (sample profile in this work) of a PARAFAC2 model are always obtained without being normalized, so the division of the sample loading vector of every factor by its norm had to be done prior to standardization and quantification.The equations of the LS regression models and the accuracy lines related to the three calibration sets analysed are also collected in Table 5.4 (columns 3–8). The criterion used for considering a sample as an outlier in every set with regard to the regression model was to have a studentized residual greater than 3 in absolute value. Trueness was verified in all cases at a 95% confidence level.

**Table 5.4.** Comparison between solvent and matrix-matched calibrations to evaluate matrix effects

| Analyte | PARAFAC / PARAFAC2 model (Outliers) | Solvent calibration (before matrix-matched analysis – *C1* set –) | | Matrix-matched calibration (*C2* set) | | Solvent calibration (after matrix-matched analysis – *C3* set –) | | *p*-values[c] |
|---|---|---|---|---|---|---|---|---|
| | | $SSL^{a} = f(c_{true})$ ($R^2$; $s_{yx}$) (Outliers) | $c_{pred} = b_0 + b_1 c_{true}$ ($R^2$; $s_{yx}$) | $SSL = f(c_{true})$ ($R^2$; $s_{yx}$) (Outliers) | $c_{pred} = b_0 + b_1 c_{true}$ ($R^2$; $s_{yx}$) | $SSL = f(c_{true})$ ($R^2$; $s_{yx}$) (Outliers) | $c_{pred} = b_0 + b_1 c_{true}$ ($R^2$; $s_{yx}$) | |
| BHT | 1 factor (PARAFAC) Unconstrained model EV[b] 99.11% (0/45) | $y = 1.0\ 10^{-1} + 5.1\ 10^{-2}x$ (99.65%; 0.06) (0/12) | $y = -8.3\ 10^{-4} + 1.0x$ (99.65%; 1.12) | $y = 1.3\ 10^{-1} + 2.9\ 10^{-2}x$ (99.73%; 0.03) (1/12) | $y = -7.0\ 10^{-4} + 1.0x$ (99.73%; 0.88) | $y = 1.0\ 10^{-1} + 4.9\ 10^{-2}x$ (99.96%; 0.02) (1/12) | $y = 0.0 + 1.0x$ (99.96%; 0.34) | 0.32 ($a_0$)<br>0.00 ($a_1$)<br>0.83 ($a_2$)<br>0.01 ($a_3$) |
| BP | 1 factor (PARAFAC) Unconstrained model EV 89.89% (0/45) | $y = 1.7\ 10^{-3} - 8.2\ 10^{-4}x + 8.9\ 10^{-4}x^2$ (99.56%; 0.05) (0/12) | $y = 1.44 + 0.96x$ (99.44%; 1.37) | $y = 7.7\ 10^{-2} + 4.0\ 10^{-2}x$ (99.97%; 0.01) (0/12) | $y = 7.8\ 10^{-4} + 1.0x$ (99.97%; 0.31) | $y = -2.5\ 10^{-3} + 1.8\ 10^{-2}x + 5.7\ 10^{-4}x^2$ (99.67%; 0.05) (0/12) | $y = 8.7\ 10^{-2} + 1.0x$ (99.73%; 0.98) | ——— |
| DiBP | 2 factors (PARAFAC2) Non-negativity constraint in mode 2 EV 99.99% CORCONDIA 100% (0/45) | $y = 1.7\ 10^{-2} + 4.2\ 10^{-3}x$ (99.99%; 0.001) (1/12) | $y = 1.5\ 10^{-3} + 1.0x$ (99.99%; 0.31) | $y = 1.14 + 3.2\ 10^{-3}x + 2.2\ 10^{-5}x^2$ (98.96%; 0.02) (0/12) | $y = -1.8\ 10^{-2} + 1.0x$ (98.68%; 3.66) | $y = 1.7\ 10^{-2} + 4.3\ 10^{-3}x$ (99.87%; 0.005) (1/12) | $y = 2.3\ 10^{-3} + 1.0x$ (99.87%; 1.13) | ——— |
| DEHA | 2 factors (PARAFAC) Non-negativity constraint in mode 1 EV 99.50% CORCONDIA 100% (0/45) | $y = 1.3\ 10^{-1} + 2.2\ 10^{-2}x$ (99.25%; 0.05) (0/12) | $y = 2.3\ 10^{-2} + 1.0x$ (99.25%; 2.09) | $y = 6.5\ 10^{-1} + 1.8\ 10^{-2}x$ (98.04%; 0.06) (0/12) | $y = -9.7\ 10^{-3} + 1.0x$ (98.04%; 3.42) | $y = 9.5\ 10^{-2} + 1.9\ 10^{-2}x$ (99.84%; 0.02) (0/12) | $y = -6.8\ 10^{-4} + 1.0x$ (99.83%; 0.99) | 0.22 ($a_0$)<br>0.00 ($a_1$)<br>0.00 ($a_2$)<br>0.00 ($a_3$) |

**Table 5.4. (cont.)**

| Analyte | PARAFAC / PARAFAC2 model (Outliers) | Solvent calibration (before matrix-matched analysis – *C1* set –) | | Matrix-matched calibration (*C2* set) | | Solvent calibration (after matrix-matched analysis – *C3* set –) | | *p*-values[c] |
|---|---|---|---|---|---|---|---|---|
| | | $SSL^{a} = f(c_{true})$ ($R^2$; $s_{yx}$) (Outliers) | $c_{pred} = b_0 + b_1\, c_{true}$ ($R^2$; $s_{yx}$) | $SSL = f(c_{true})$ ($R^2$; $s_{yx}$) (Outliers) | $c_{pred} = b_0 + b_1\, c_{true}$ ($R^2$; $s_{yx}$) | $SSL = f(c_{true})$ ($R^2$; $s_{yx}$) (Outliers) | $c_{pred} = b_0 + b_1\, c_{true}$ ($R^2$; $s_{yx}$) | |
| DiNP | 3 factors (PARAFAC) Non-negativity constraint in all modes EV 96.05% CORCONDIA 89% (0/45) | $y = 1.7\ 10^{-2} + 8.9\ 10^{-4}x$ (99.65%; 0.06) (0/12) | $y = -3.0\ 10^{-2} + 1.0x$ (99.65%; 63.55) | $y = -1.5\ 10^{-2} + 6.6\ 10^{-4}x$ (99.25%; 0.06) (0/12) | $y = -5.7\ 10^{-2} + 1.0x$ (99.25%; 92.83) | $y = 4.0\ 10^{-3} + 8.0\ 10^{-4}x$ (99.46%; 0.06) (0/12) | $y = 3.0\ 10^{-2} + 1.0x$ (99.46%; 79.11) | 0.70 ($a_0$)<br>0.00 ($a_1$)<br>0.34 ($a_2$)<br>0.00 ($a_3$) |

[a] Standardized sample loading

[b] Explained variance

[c] *p*-values for the hypothesis test on the significance of the estimates of the *a*-coefficients in Eq. (5.3)

As explained in Section 5.2.4.2, the possible matrix effects in the determination of the five analytes in the nipple extracts were assessed by checking the statistical equality of the three calibration models "*Standardized sample loading versus True concentration*" estimated. The model in Eq. (5.3) was fitted to the experimental data from the three sets just for BHT, DEHA and DiNP. For BP and DiBP, different mathematical functions (quadratic instead of linear models) were obtained for the solvent calibration curves and for the matrix-matched calibration (see Table 5.4), since a significant lack of fit was found when a linear regression was posed; this fact confirmed that the quantification of BP and DiBP in the nipple samples was affected by the presence of matrix components: the analytical behavior of each of these two compounds was described by different functional models (polynomial instead of linear regressions) depending on the environment where they were in solution.

For BHT, DEHA and DiNP, once the model in Eq. (5.3) had been fitted in each case, a hypothesis test on the statistical significance of the estimated coefficient of every term including the indicator variables $z_1$ and $z_2$ was posed, where $H_0$: $a_i = 0$ against $H_a$: $a_i \neq 0$; the resultant *p*-values are shown in the last column of Table 5.4. At a 99% confidence level, some of the estimates of $a_0$, $a_1$, $a_2$ and $a_3$ were significantly non-null for those three analytes. As a consequence, matrix effects were proved to be present for BHT, DEHA and DiNP, being the analytical sensitivity of every compound significantly different. These joint matrix problems detected for the five compounds under study led to the decision of performing their quantification in future studies following the standard addition method.

#### *5.2.5.4. Extraction from a dummy*

A dummy made of natural rubber latex was purchased from a local supermarket. An extraction from this dummy into hexane was carried out as detailed in Section 5.2.3.3. The standard addition method was used to determine the concentration of the analytes studied in this work in the sample. Each analyte was added within the concentration ranges included in Table 5.1 (fifth row, third column). For the special case of DiBP, the lowest added concentration was 25 μg $L^{-1}$ since the procedure to assess the level of this analyte in solvent blanks (explained in Section 5.2.5.2) was done at this concentration level.

High amounts of BHT were detected in the extracted solution, so it was necessary to perform the standard addition method twice. The first one was carried out using 4.5 mL of the extracted solution into 5-mL volumetric flasks (with 25 μg $L^{-1}$ of DiBP-$d_4$) in order to quantify BP, DiBP, DEHA and DiNP. However, as the quantity of BHT present was too high, the second one was performed to quantify only BHT. In this case, the extract was more diluted (1.3 mL of the same extracted solution into 5-mL volumetric flasks) and

contained a higher amount of the internal standard (225 μg $L^{-1}$ of DiBP-$d_4$) than at any previous stages of this work.

In both analyses, seven matrix-matched standards at seven concentration levels were injected in duplicate. All the analytes were added to these standards (except for the first one), although only some of them would be quantified in each case. The concentration levels of each analyte in both analyses were the same except for BHT, since its concentration range was reduced for the development of the first standard addition regression (0–30 μg $L^{-1}$). One solvent blank without IS was also injected at the beginning of both analytical sequences and a system blank was analysed between each matrix-matched standard.

A three-way data tensor containing all the data from the first standard addition set was built for BP, DiBP (with DiBP-$d_4$), DEHA and DiNP, while only data tensors for BHT and the internal standard (DiBP-$d_4$) were built for the second one. The dimensions of these data tensors are specified in Table 5.1 (fifth row), whereas the characteristics of the PARAFAC decompositions performed for each tensor are collected in Table 5.5.

**Table 5.5.** PARAFAC models estimated with the migration test data, identification of every analyte according to the regulation and results obtained with the standard addition method. The third column shows the relative retention time and the relative abundances (in brackets) estimated from the PARAFAC spectral loadings for each diagnostic ion. In bold, the non-compliant $m/z$ ratios.

| Analyte | PARAFAC model | Identification | LS regression "*Standardized sample loadings versus Added concentration*" | | Accuracy line | Sample concentration and 95% confidence interval (μg $L^{-1}$) | Conclusions |
|---|---|---|---|---|---|---|---|
| | | | Model ($R^2$, $s_{yx}$) | Error[c] (%) | Model ($R^2$, $s_{yx}$) | | |
| BHT | 1 factor<br>Unconstrained model<br>Explained variance 99.88% | $t_{R,rel}$ = 0.808<br>91 (5.98%)<br>145 (10.99%)<br>177 (7.55%)<br>205 (100%)<br>220 (24.06%) | $y = 3.84\ 10^{-1} + 2.34\ 10^{-3}\ x$<br>($n$ = 12)<br>(99.24%, $2.87\ 10^{-2}$) | 4.50%<br>($n$ = 10) | $y = 1.44\ 10^{-1} + 1.00\ x$<br>(99.24%, 12.25) | 37.87<br>(31.23, 45.13) | BHT detected |
| BP | 1 factor<br>Unconstrained model<br>Explained variance 99.33% | $t_{R,rel}$ = 0.882<br>51 (18.83%)<br>77 (58.15%)<br>105 (100%)<br>152 (3.80%)<br>**182** (57.68%) | $y = 6.72\ 10^{-1} + 4.89\ 10^{-2}\ x$<br>($n$ = 14)<br>(96.94%, $3.76\ 10^{-2}$) | 10.77%<br>($n$ = 12) | $y = 5.18\ 10^{-3} + 1.00\ x$<br>(96.93%, $7.69\ 10^{-1}$) | 0.92<br>(0.76, 1.1) | Below its CCα |
| $DiBP\text{-}d_4$ | 1st standard addition method:<br>3 factors[a] (Factor 2: $DiBP\text{-}d_4$)<br>Non-negativity constraint in mode 2<br>Explained variance 99.99%<br>CORCONDIA 99.72% | $t_{R,rel}$ = 1.000<br>80 (5.72%)<br>153 (100%)<br>171 (2.11%)<br>209 (1.47%)<br>227 (5.45%) | | | Internal standard | | |
| | 2nd standard addition method:<br>2 factors (Factor 1: $DiBP\text{-}d_4$)<br>Non-negativity constraint in modes 1 and 2<br>Explained variance 99.71%<br>CORCONDIA 100% | $t_{R,rel}$ = 1.000<br>80 (5.25%)<br>153 (100%)<br>171 (2.70%)<br>209 (1.40%)<br>227 (5.69%) | | | Internal standard | | |

**Table 5.5. (cont.)**

| Analyte | PARAFAC model | Identification | LS regression "*Standardized sample loadings versus Added concentration*" | | Accuracy line | Sample concentration and 95% confidence interval (μg $L^{-1}$) | Conclusions |
|---|---|---|---|---|---|---|---|
| | | | Model ($R^2$, $s_{yx}$) | Error[c] (%) | Model ($R^2$, $s_{yx}$) | | |
| DiBP | 1st standard addition method:<br>3 factors [a] (Factor 3: DiBP)<br>Non-negativity constraint in mode 2<br>Explained variance 99.99%<br>CORCONDIA 99.72% | $t_{R,rel}$ = 1.001<br>104 (6.82%)<br>149 (100%)<br>167 (3.23%)<br>205 (1.65%)<br>223 (6.47%) | $y = 4.00\,10^{-1} + 2.03\,10^{-2}\,x$<br>($n$ = 10)<br>(99.01%, $2.77\,10^{-2}$) | 3.15%<br>($n$ = 8) | $y = -7.39\,10^{-4} + 1.00\,x$<br>(99.01%, 1.36) | 1.31<br>(1.04, 1.62) | Lower than 25 μg $L^{-1}$ |
| DEHA | 3 factors (Factor 1: DEHA)<br>Non-negativity constraint in modes 1 and 2<br>Explained variance 99.33%<br>CORCONDIA 96% | $t_{R,rel}$ = 1.282<br>**112** (23.62%)<br>129 (100%)<br>147 (18.86%)<br>241 (5.41%)<br>**259** (3.31%) | $y = 1.47\ 10^{-2} + 3.26\ 10^{-2}\,x$<br>($n$ = 14)<br>(99.45%, $4.21\,10^{-2}$) | 3.60%<br>($n$ = 12) | $y = -1.07\,10^{-3} + 1.00\,x$<br>(99.45%, 1.29) | 0.03<br>(−0.13, 0.19) | The confidence interval includes 0 |
| DiNP | 3 factors[a] (Factor 2: DiNP)<br>Non-negativity constraint in modes 2 and 3<br>Explained variance 99.99%<br>CORCONDIA 94.95% | $t_{R,rel}$ = —[b]<br>57 (43.88%)<br>127 (8.80%)<br>149 (100%)<br>167 (9.16%)<br>**275** (0.01%)<br>**293** (15.75%) | $y = 5.86\ 10^{-3} + 9.91\ 10^{-4}\,x$<br>($n$ = 12)<br>(95.89%, $9.59\cdot 10^{-2}$) | 12.93%<br>($n$ = 10) | $y = -3.57\,10^{-3} + 1.00\,x$<br>(95.89%, 96.80) | 0.394<br>(−11.44, 14.06) | The confidence interval includes 0 |

[a] PARAFAC2 model.

[b] It is not possible to establish a retention time for DiNP.

[c] Mean of the absolute value of the relative error in calibration.

The PARAFAC decomposition for BHT and BP required just one factor. In the case of DiBP-$d_4$, a three-factor PARAFAC2 model was needed for the data obtained from the first performance of the standard addition method. The first factor was the interferent that eluted before the internal standard as explained in previous sections; the second factor was associated with DiBP-$d_4$ and the third one with DiBP. On the other hand, when the PARAFAC decomposition of the tensor containing DiBP-$d_4$ for the quantification of BHT was performed, a two-factor model that was coherent with the presence of DiBP and DiBP-$d_4$ was obtained.

The PARAFAC decomposition of the DEHA tensor yielded a three-factor model where two interferents coeluted in the matrix-matched standards, while a three-factor PARAFAC2 model was necessary for DiNP. In this last case, the first factor was associated to an interferent that appeared mainly in the system blanks, the second factor was associated to DiNP and the third one was the baseline.

By way of example, the sample loadings for BHT in the second standard addition set are represented in Figure 5.8. There was not a memory effect since the sample loadings for all the system blanks had a null value. As expected, the sample loading for the first matrix-matched standard and its replicate (in which only the internal standard had been added) was different from zero. The loadings for the remaining matrix-matched standards were in increasing order, which was coherent.

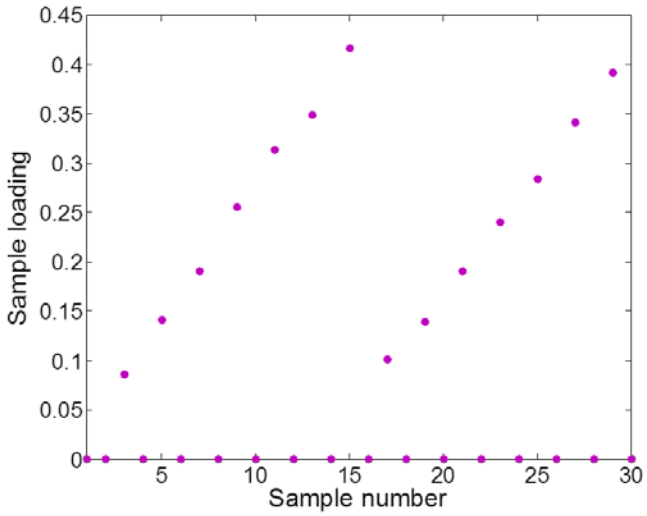


**Figure 5.8.** Loadings of the sample profile of the one-factor PARAFAC model for the data tensor of BHT built with the samples analysed during the second performance of the standard addition method.

It must be noted that the GC column was cut due to maintenance tasks before the analysis of the samples corresponding to this section. This led to a variation in the absolute

retention times regarding the ones obtained with the reference standards. However, there was no problem in the identification of each analyte since the relative retention times obtained in this section for each analyte (see Table 5.5, third column) lay within the tolerance intervals established previously (see Table 5.2). In addition, the unequivocal identification of each analyte was guaranteed since at least the relative abundances of 3 $m/z$ ratios for each analyte (see Table 5.5, third column) were within their tolerance intervals (see Table 5.2).

The sample loadings of each analyte were standardized with the sample loadings of DiBP-$d_4$ obtained from the model required in each case. As commented in Section 5.2.5.3, the sample loadings of DiBP-$d_4$, DiBP and DiNP obtained from the data from the first performance of the standard addition method were numerically high because they came from PARAFAC2 decompositions, so they had to be manually normalized prior to standardization. Next, a LS regression between the standardized sample loadings of every analyte and the added concentration was built. All the regressions were significant and there was not a lack of fit in any case after the removal from the calibration set of the last matrix-matched standard and its replicate for BHT and DiNP and the last two ones and their replicates for DiBP. The parameters of the regression models built for each analyte and the results of this analysis are included in Table 5.5. The trueness of the method was verified for all analytes at a 95% confidence level.

The amount of each analyte released from the dummy in the total volume of hexane together with the corresponding confidence interval appears in Table 5.5. The concentration found for BP was below its $CC\alpha$, the value obtained for DiBP was below the first standard fixed at 25 μg $L^{-1}$, so it could not be exactly determined, while the confidence interval for DEHA and DiNP contained zero. Therefore, BP, DEHA and DiNP were not detected in the extracted solution, while the exact concentration of DiBP could not be estimated. The conclusion reached after this study was that the presence of BHT from the dummy was observed.

#### *5.2.5.5. Future developments*

This work presents a methodology to face the two problems posed, but several issues still remain open. Regarding the unequivocal identification of every compound, one of them is that related to the retention time of finger-peak analytes like DiNP, so a strategy for the estimation of this feature or a similar one for complex signals is needed and currently being devised, as the relative retention time is one of the requirements that must be fulfilled at the identification step. On the other hand, the procedure to assess the blank level of DiBP permits its performance by considering a lower concentration of this analyte in the solvent

standards for *Distribution 2*. Of course, both $\alpha$ and $\beta$ would increase, but they would achieve quite good values anyway, as the results yielded for them in our work were low enough to achieve still low $\alpha$ and $\beta$ errors even if the blank level of DiBP and the concentration value considered to pose the alternative hypothesis in the test were closer.

### 5.2.6. Conclusions

In the multiresidue determination of BHT, BP, DiBP, DEHA and DiNP by GC-MS, using DiBP-$d_4$ as internal standard, PARAFAC has succeeded in the unequivocal identification of every compound according to the performance requirements laid down in the EU legislation in force. Besides, this chemometric tool has offered an improved way to deal with complex finger-peak signals such as that of DiNP by using loadings instead of integrated areas in both identification and quantification steps.

The problem of the ubiquity of DiBP has also been overcome thanks to the statistical comparison of the two probability distributions of the level of this analyte in solvent blanks and standards and the assessment of the $\alpha$ and $\beta$ errors taken. This has enabled to state that any standardized sample loading of DiBP significantly greater than 1.04 will indicate that this analyte is present in that sample at a concentration higher than that in blanks ($\alpha = \beta = 0.012$).

The evidence of matrix effects has forced to quantify every analyte by means of the standard addition methodology after an extraction into hexane from a natural rubber latex dummy. BHT has been the analyte undoubtedly detected as a potential migrant from the dummy if used, with a concentration significantly different from zero in the hexane extract.

### 5.2.7. Acknowledgements

The authors thank the financial support provided by projects of the Ministerio de Economía y Competitividad (CTQ2014-53157-R). M.L. Oca and L. Rubio are particularly grateful to Universidad de Burgos for their FPI grants.

### 5.2.8. References

[1] Office of Environmental Health Hazard Assessment's (OEHHA) - Reproductive and Cancer Hazard Assessment Branch, *Evidence on the Carcinogenicity of Diisononyl Phthalate (DINP)*, California Environmental Protection Agency, October 2013.

[2] P. Ventrice, D. Ventrice, E. Russo, G. De Sarro, *Phthalates: European regulation, chemistry, pharmacokinetic and related toxicity*, Environmental Toxicology and Pharmacology 36 (2013) 88-96

[3] J. Annamalai, V. Namasivayam, *Endocrine disrupting chemicals in the atmosphere: their effects on humans and wildlife*, Environment International 76 (2015) 78-97.

[4] M. Wagner, J. Oehlmann, *Endocrine disruptors in bottled mineral water: total estrogenic burden and migration from plastic bottles*, Environmental Science and Pollution Research 16 (2009) 278-286.

[5] E. Fasano, F. Bono-Blay, T. Cirillo, P. Montuori, S. Lacorte, *Migration of phthalates, alkylphenols, bisphenol A and di(2-ethylhexyl)adipate from food packaging*, Food Control 27 (2012) 132-138.

[6] M.D. Spillmann, M. Siegrist, C. Keller, M. Wormuth, *Phthalate exposure through food and consumers' risk perception of chemicals in food*, Risk Analysis 29 (2009) 1170-1180.

[7] B. Cavaliere, B. Macchione, G. Sindona, A. Tagarelli, *Tandem mass spectrometry in food safety assessment: The determination of phthalates in olive oil*, Journal of Chromatography A 1205 (2008) 137-143.

[8] M.R. Lee, F.Y. Lai, J. Dou, K.L. Lin, L.W. Chung, *Determination of trace leaching phthalate esters in water and urine from plastic containers by solid-phase microextraction and gas chromatography–mass spectrometry*, Analytical Letters 44 (2011) 676–686.

[9] A. Afshari, L. Gunnarsen, P.A. Clausen, V. Hansen, *Emission of phthalates from PVC and other materials*, Indoor Air 14 (2004) 120-128.

[10] C.S. Giam, H.S. Chan, G.S. Neff, *Rapid and inexpensive method for detection of polychlorinated biphenyls and phthalates in air*, Analytical Chemistry 47 (1975) 2319-2320.

[11] Commission Regulation (EU) No 10/2011 of 14 January 2011 on plastic materials and articles intended to come into contact with food, Official Journal of the European Union, L 12, 1-89.

[12] A. Fankhauser-Noti, K. Grob, *Blank problems in trace analysis of diethylhexyl and dibutyl phthalate: Investigation of the sources, tips and tricks*, Analytica Chimica Acta 582 (2007) 353–360.

[13] M. Marega, K. Grob, S. Moret, L. Conte, *Phthalate analysis by gas chromatography–mass spectrometry: Blank problems related to the syringe needle*, Journal of Chromatography A 1273 (2013) 105-110.

[14] A. Reid, C. Brougham, A. Fogarty, J. Roche, *An investigation into possible sources of phthalate contamination in the environmental analytical laboratory*, International Journal of Environmental Analytical Chemistry 87 (2007) 125 -133.

[15] K. Furtmann, *Phthalates in surface water – A method for routine trace level analysis*, Fresenius' Journal of Analytical Chemistry 348 (1994) 291-296.

[16] F. David, P. Sandra, B. Tienpont, F. Vanwalleghem, M. Ikonomou, in: C.A. Staples (Ed.), *Phthalate Esters, Part Q (The Handbook of Environmental Chemistry, vol. 3)*, Springer-Verlag, Berlin, Heidelberg, 2003, pp. 9-56.

[17] M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, *Tutorial on evaluation of type I and type II errors in chemical analyses: From the analytical detection to authentication of products and process control*, Analytica Chimica Acta 674 (2010) 123–142.

[18] M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, A. Herrero, *Quality of analytical measurements: statistical methods for internal validation*, in: S. Brown, R. Tauler, R. Walczak (Eds.), *Comprehensive Chemometrics, vol. 1*, Elsevier, Oxford, 2009, pp. 17–76.

[19] European Union risk assessment report. *1,2-Benzenedicarboxylic acid, di-C8-10-branched alkyl esters, C9-rich and di-"isononyl" phthalate (DINP) CAS Nos: 68515-48-0 and 28553-12-0. EINECS Nos: 271-090-9 and 249-079-5*. European Commission – Joint Research Centre, Institute for Health and Consumer Protection, European Chemicals Bureau (ECB). Volume 35 (2003).

[20] *Plasticisers and Flexible PVC information centre.* Available at: http://www.plasticisers.org/plasticisers/22/93/Diisononyl-phthalate-DINP (Accessed: 24/09/2015).

[21] S. Bratinova, B. Raffael, C. Simoneau, *Guidelines for performance criteria and validation procedures of analytical methods used in controls of food contact materials - EUR 24105 EN (JRC53034)*, Publications Office of the European Union, Luxembourg, 1st ed., 2009.

[22] 2002/657/EC Commission Decision of 12 August 2002, implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Off. J. Eur. Commun. L 221 (2002) 8-36.

[23] SANCO/12571/2013, Guidance document on analytical quality control and validation procedures for pesticide residues analysis in food and feed, EU, Brussels, 2013.

[24] *IARC Monographs – 101. Benzophenone*, pp. 285-304. Available at https://monographs.iarc.fr/ENG/Monographs/vol101/mono101-007.pdf (Accessed: 22/10/2015).

[25] *NIST Mass Spectral Search Program for the NIST/EPA/NIH Mass Spectral Library Version 2.0 a.* (build July 1 2002). National Institute of Standards and Technology, Gaithersburg, USA, 2002.

[26] B.M. Wise, N.B. Gallagher, R. Bro, J.M. Shaver, W. Winding, R.S. Koch, *PLS Toolbox 6.0.1*, Eigenvector Research Inc., Wenatchee, WA, USA.

[27] *MATLAB, version 7.12.0.635 (R2011a)*, The Mathworks, Inc., Natick, MA, USA.

[28] *STATGRAPHICS Centurion XVI Version 16.1.05 (32 bit)*, Statpoint Technologies, Inc., Herndon, VA, USA, 2010.

[29] L.A. Sarabia, M.C. Ortiz, *DETARCHI. A program for detection limits with specified assurance probabilities and characteristic curves of detection*, Trends in Analytical Chemistry 13 (1994)1-6.

[30] R. Bro, *PARAFAC. Tutorial and applications*, Chemometrics and Intelligent Laboratory Systems 38 (1997)149-171.

[31] R. Bro, H.A.L. Kiers, *A new efficient method for determining the number of components in PARAFAC models*, Journal of Chemometrics 17 (2003) 274-286.

[32] H.A.L. Kiers, J.M.F. Ten Berge, R. Bro, *PARAFAC2—Part I. A direct fitting algorithm for the PARAFAC2 model*, Journal of Chemometrics 13 (1999) 275-294.

[33] R. Bro, C.A. Andersson, H.A.L. Kiers, *PARAFAC2—Part II. Modeling chromatographic data with retention time shifts*, Journal of Chemometrics 13 (1999) 295-309.

[34] M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, A. Herrero, S. Sanllorente, C. Reguera, *Usefulness of PARAFAC for the Quantification, Identification, and Description of Analytical*

*Data*, in: A. Muñoz de la Peña, H.C. Goicoechea, G.M. Escandar, A.C. Olivieri (Eds.), *Fundamentals and Analytical Applications of Multiway Calibration*, Elsevier, 2015, pp. 37-81

[35] N.R. Draper, H. Smith, *Applied Regression Analysis*, 3rd ed., John Wiley and Sons, New York, 1998.

[36] M.C. Ortiz, M.S. Sánchez, L.A. Sarabia, *Quality of analytical measurements: univariate regression*, in: S. Brown, R. Tauler, R. Walczak (Eds.), *Comprehensive Chemometrics, vol. 1*, Elsevier, Oxford, 2009, pp. 127–169.

[37] M.L. Oca, L.A. Sarabia, A. Herrero, M.C. Ortiz, *Optimum pH for the determination of bisphenols and their corresponding diglycidyl ethers by gas chromatography–mass spectrometry. Migration kinetics of bisphenol A from polycarbonate glasses*, Journal of Chromatography A 1360 (2014) 23–38.

[38] A.M. Botero-Coy, J.M. Marín, R. Serrano, J.V. Sancho, F. Hernández, *Exploring matrix effects in liquid chromatography–tandem mass spectrometry determination of pesticide residues in tropical fruits*, Analytical and Bioanalytical Chemistry 407 (2015) 3667–3681.

[39] H. Kwon, S.J. Lehotay, L. Geis-Asteggiante, *Variability of matrix effects in liquid and gas chromatography–mass spectrometry analysis of pesticide residues after QuEChERS sample preparation of different food crops*, Journal of Chromatography A 1270 (2012) 235-245.

[40] M.C. Ortiz, L.A. Sarabia, I. García, D. Giménez, E. Meléndez, *Capability of detection and three-way data*, Analytica Chimica Acta 559 (2006) 124-136.

[41] M.S. Pascal Gimeno, S. Thomas, C. Bousquet, A.F. Maggio, C. Civade, C. Brenier, P.A. Bonnet, *Identification and quantification of 14 phthalates and 5 non-phthalate plasticizers in PVC medical devices by GC-MS*, Journal of Chromatography B 949-950 (2014) 99-108.

[42] F.J. Massey Jr., *The Kolmogorov-Smirnov Test for Goodness of Fit*, Journal of the American Statistical Association 253 (1951) 68-78.

# CHAPTER 6

## CONCLUSIONS OF THIS DOCTORAL THESIS

After the exposition of the research that comprises this doctoral thesis, the following conclusions, arranged by chapter and related to the objectives proposed in Section 1.4 of Chapter 1, can be drawn from it:

### *Chapter 2 – Ensuring Food Quality in regulated products through multivariate methodologies: Prediction of Queso Zamorano's quality by near-infrared spectroscopy assessing false non-compliance and false compliance at minimum permitted limits stated by Protected Designation of Origin regulations*

1. The fat, dry matter and protein contents in *Queso Zamorano* items can be simultaneously determined by the fast non-destructive NIR+PLS methodology resulting from this research and developed in accordance both with the Spanish regulation of the PDO "*Queso Zamorano*".and the ISO guideline 21543:2006 for the application of near infrared spectrometry on milk products. Consequently, the fulfilment of the objectives G1, G3 and S1, can be confirmed.

2. The choice of the final PLS regression for every constituent has been made by comparing the features of the calibration models obtained from several spectral preprocessing strategies, so the one that yielded the lowest number of latent variables bearing the evolution of the *RMSECV* in mind together with the highest percentage of explained variance of the response variable without having removed too many outlier data from every calibration set has been selected as the most coherent model. In this sense, obtaining the best analytical quality possible has also been the target during this development, thus fulfilling objective G2.

3. Considering the minimum permitted limits for the percentages in weight of protein, dry matter and fat-to-dry matter ratio laid down in the regulation of the PDO "*Queso Zamorano*", the concepts of *decision limit* and *detection capability* when a NIR+PLS calibration is used have been adjusted to this situation, thus resulting the concepts of $CD\alpha$ and $CD\beta$, respectively, which have been applied here for the first time to a food product granted with a European PDO. Thanks to the goodness of the NIR+PLS models obtained, the concentration of every constituent under control that could be distinguished from its corresponding minimum permitted limit has a very close value to the latter, being the probability of false non-compliance $\alpha$ equal to 0.05 and the probability of false compliance $\beta$ equal to or less than 0.05. As a consequence, objective S2 has been completed.

***Chapter 3 – Ensuring Food Safety in regulated products through Design of Experiment approaches: Effect analysis and robustness assessment of several sample pretreatment factors on the determination of seven sedatives in animal muscle by liquid chromatography-tandem mass spectrometry***

As for ***Section 3.2 – Desirability functions as response in a D-optimal design for evaluating the extraction and purification steps of six tranquillizers and an anti-adrenergic by liquid chromatography-tandem mass spectrometry:***

4. The analysis of the effect of 4 factors related to the extraction and purification steps of azaperone, propionylpromazine, chlorpromazine, haloperidol, xylazine, azaperol and carazolol from pig muscle tissue on its final determination by LC-MS/MS has been performed through a D-optimal design, where the number of experiments has been reduced from 24 to 11, thus reducing the experimental effort and costs by more than 50%. This statement confirms the achievement of objective S3.

5. The optimum conditions for the determination of the analytes have been achieved by using a desirability function as experimental response, so objective S4 has been attained. This variable has been defined by bearing two analytical criteria of interest in mind: *i)* the assessment of a similar chemical behaviour of each analyte in relation to its internal standard, and *ii)* the avoidance of a significant reduction of the peak area of the internal standards.

6. The desirability function defined as experimental response has proved to be a more appropriate and flexible approach for the optimization phase that reflects the analytical problem better, thus enhancing the quality of the devised method. Since the significance of the factors in a study and their optimal levels depend on the experimental response considered, the *ad hoc* procedure built to define that response through desirability functions together with the experimental domain and the connection between factor levels and response by means of a D-optimal design would enable to adapt the methodology of experimental designs exactly to any analytical problem under study. As a consequence, there can be no doubt that objective G2 has been significantly achieved.

7. The unequivocal identification of every analyte and the internal standards used has been ensured in the way the European legislation in force requires, thus fulfilling objective G1.

As for ***Section 3.3 – Robustness testing in the determination of seven drugs in animal muscle by liquid chromatography–tandem mass spectrometry:***

8. The robustness of the sample preparation stage for the determination of the seven sedatives tackled in Section 3.2 in pig muscle tissue by LC-MS/MS has been verified by means of a $2_{III}^{7-4}$ fractional factorial design, following the Youden approach recommended by European legislation in force. Objective S5 has then been fulfilled.

9. The influence of 7 factors involved in the sample preparation prior to the LC-MS/MS analysis has been evaluated from 8 experimental runs by means of different statistical strategies, namely, hypothesis testing based on a previous estimation of the experimental variance, Lenth's method and Bayes' analysis, the latter two allowing to draw conclusions on which factors are active even though no estimate of the residual error is available. Whatever the approach considered for the interpretation of the results of the experimental design, the sample preparation procedure under study has proved to be robust to the small variations defined as factor ranges, which can be expected throughout its routine performance. This evidences the fulfilment of objective S6.

10. The unequivocal identification of every analyte and the internal standards used has been ensured in the way the European legislation in force requires, thus fulfilling objective G1.

## *Chapter 4 – Ensuring Food Safety from packaging through multiway techniques: Development of an analytical method to determine bisphenols and their derivatives by gas chromatography-mass spectrometry and Modelling of their migration into lipophilic/alcoholic foodstuffs*

As for *Section 4.2 – Optimization of a GC-MS procedure that uses parallel factor analysis for the determination of bisphenols and their diglycidyl ethers after migration from polycarbonate tableware:*

11. A GC-MS method fully adhered to the EU food migrant regulations and guidelines has been developed for the simultaneous determination of BPF, BPA, BFDGE and BADGE using BPA-$d_{16}$ as internal standard and PARAFAC decomposition as the chemometric technique to extract the analytical information on the chromatographic, spectral and sample behaviours of every compound taking advantage of the second-order nature of the experimental data. This approach has proved to be successful in: *i)* identifying every compound unequivocally according to the maximum permitted tolerances for relative ion abundances and relative retention time, and *ii)* quantifying every analyte, even in the presence of coeluents. The specificity required for every analytical method and specifically for those dealing with regulated substances is fully ensured by the use of PARAFAC decompositions on GC-MS three-way tensors; that specificity can fail if the Total Ion Chromatogram (TIC) is considered whenever there is poor resolution between some peaks or whether interferents coelute. So, all this being said, it can be concluded that objectives G1, G2, S7 and S8 have been achieved.

12. Regarding objective S9, the effect of changes in some parameters in the oven temperature program on the quality of determinations and identifications of the analytes has been assessed according to a $2^2$ factorial design. After the calculation of the five figures of merit considered as response variables from non-linear calibration curves based on the PARAFAC models resulting from the tensors built for every experimental run, a Principal Component Analysis of the data obtained for the four analytes at the four experimental points has been performed to find the optimum conditions for the GC oven, taking the total run time and the EU restrictions on some of these compounds into account.

13. Being specificity guaranteed by PARAFAC decompositions, detection capability values between 2.65 and 4.71 $\mu g\ L^{-1}$ when acetonitrile is the injection solvent, and between 1.97 and 5.53 $\mu g\ L^{-1}$ for acetone, with probabilities of false positive and false negative fixed at 0.05, have been obtained. In all cases, the trueness of the devised methodology has been confirmed at a 95% confidence. On the other hand, the

sample preparation method has clearly performed better for BPF and BPA than for BFDGE and BADGE, as recoveries were higher for the first two compounds. To sum up, all this shows evidence of the completion of objective S10.

14. The migration of BPA from polycarbonate tableware into food simulant *D1* has been confirmed, and values between 104.67 and 181.46 μg $L^{-1}$ (0.73 and 1.27 μg $L^{-1}$ after dilution correction) of BPA have been estimated. None of the results obtained exceeds the specific migration limit of 600 μg $L^{-1}$ established for BPA in food contact materials different from infant feeding bottles by the legislation in force back when this research was conducted. As a result, objectives G3 and S11 have been attained.

As for ***Section 4.3 – Optimum pH for the determination of bisphenols and their corresponding diglycidyl ethers by gas chromatography-mass spectrometry. Migration kinetics of bisphenol A from polycarbonate glasses:***

15. Starting from the methodology devised and explained in Section 4.2, a statistical assessment of the influence of the pH on the simultaneous determination of BPF, BPA, BFDGE and BADGE has been conducted. The output of this evaluation reflects that the best working pH is 9, which agrees with the epoxide ring-opening of BFDGE and BADGE at extreme pH conditions and, on the other hand, with the $pK_{a,\,25\,^{\circ}C}$ values of BPF and BPA. Objective S12 has then been completed.

16. In connection with objectives G3 and S14, the experimental data from a kinetic study at two different incubation temperatures (50 ºC and 70 ºC) to assess the migration of BPA from polycarbonate glasses into food simulant *D1* have been fitted to an exponential model. The migration equation estimated for every glass tested has evidenced that *i)* the higher the heating temperature, the greater the relative migration rate of BPA, and that *ii)* BPA is hardly significantly transferred from the plastic to the food simulant at 50 ºC, whereas it is at 70 ºC. On the other hand, an identical analytical behaviour regarding the release of BPA has been found for all the glasses evaluated at the same temperature. However, this essentially common migration process cannot be described by the same parameters for all the containers, probably due to some differences in their manufacture.

17. Specifically linked to objective G3, none of the amounts of BPA migrated after every test conducted on every polycarbonate glass has exceeded the specific migration limit of 0.6 mg $kg^{-1}$ set out for this compound back when this research was performed, so the compliance of the six glasses evaluated was ensured.

18. By applying the optimized analytical method to several migration samples fortified with the four analytes under study and being specificity guaranteed by PARAFAC decompositions, the corresponding average recovery rate now turns out to be 92.75% for BPF, 90.62% for BPA, 89.29% for BFDGE and 50.14% for BADGE, thus obtaining significantly greater values for the two diglycidyl derivatives. The trueness of the analytical method has also been verified at a 95% confidence level, being the capability of detection for BPA equal to 2.60 μg $L^{-1}$ ($n$ = 2 replicates, $\alpha = \beta = 0.05$). Consequently, objective S13 has been attained.

19. PARAFAC decomposition has again succeeded in: *i)* identifying every compound unequivocally according to the maximum permitted tolerances for relative ion

abundances and relative retention time, which guarantees the specificity of the analysis, and *ii)* quantifying every analyte, even in the presence of interferents. The traditional methodology based on the mass spectrum recorded at the retention time could never have achieved those unequivocal identification and accurate quantification because of coeluents that share ions with the analyte of interest, which gives clear proof of the benefits of considering the second-order nature of GC-MS data. This fact matches objectives G1 and G2 proposed at the beginning of this doctoral thesis.

**Chapter 5 - Ensuring Food Quality by first ensuring Analytical Quality when complex signals and an uncontrolled source for one of the target migrants are present in the laboratory environment: Dealing with the ubiquity of phthalates in the laboratory when determining plasticizers by gas chromatography-mass spectrometry and PARAFAC**

20. The multiresidue determination of two phthalates, namely, diisobutyl phthalate (DiBP) and diisononyl phthalate (DiNP), together with two plasticizers (2,6-di-*tert*-butyl-4-methyl-phenol (BHT) and bis(2-ethylhexyl) adipate (DEHA)) and a UV stabilizer (benzophenone (BP)) by gas chromatography-mass spectrometry (GC-MS) using DiBP-$d_4$ as internal standard has been achieved in accordance with the EU food migrant policies and the analytical performance criteria currently in force. The use of PARAFAC or PARAFAC2 as the chemometric tools for the analysis of the resulting second-order GC-MS data has been key for succeeding in:

    a. The quantification and the unequivocal identification of all the analytes even in the presence of unknown coeluents that share both elution times and diagnostic ions with the former. As matrix effects are present in the determination of these compounds, their quantification has been performed following the standard addition method using the PARAFAC sample loadings as response variable, which is not a very common approach. As a result, objectives G1, G2 and S16 have been achieved.

    b. Dealing with the complex chromatogram of DiNP, with a multiple-finger-peak shape because of an array of possible $C_9$ isomers: as both PARAFAC and PARAFAC2 do not consider peak areas but the chromatographic and spectral loadings of the factor linked to every compound, no ambiguity depending on variable peak area integration is present, and a significantly most accurate determination, both qualitative and quantitative, is attained, and so are objectives G1, G2, S15*i* and S16.

21. The issue of the ubiquity of DiBP has been overcome thanks to the statistical comparison of the two probability distributions of the concentration of this analyte in solvent blanks and standards, respectively, and the assessment of the $\alpha$ and $\beta$ errors taken: at a probability of making a type I error ($\alpha$) and that of a type II error ($\beta$) both equal to 0.012, computing the critical or rejection standardized sample loading of DiBP beyond which the presence of DiBP in a test sample is concluded to happen at a concentration statistically higher than that in a blank results in a value of 1.04. This estimation and its use to have a reference of the background level of DiBP in the

laboratory environment and thus minimize the risk of false positives lead to affirm that objectives G2 and S15*ii* have been achieved.

22. The application of the devised methodology to the determination of the analytes in a hexane extract from a natural rubber latex dummy intended for infants has concluded that BHT can be considered as a potential migrant from that dummy if used, since its presence has been detected at a concentration significantly different from zero after that exhaustive extraction procedure. Objectives G3 and S17 have then been completed.