JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

J. Agric. Food Chem. **2007,** *55,* 25–31 **25**

GC-MS Quantification of Suspected Volatile Allergens in Fragrances. 2. Data Treatment Strategies and Method Performances

MAUD BASSEREAU,[†] ALAIN CHAINTREAU,*[,‡] STÉPHANIE DUPERREX,[§] DANIEL JOULAIN,[#] HANS LEIJS,[||] GERD LOESING,[⊥] NEIL OWEN,[⊗] ALAN SHERLOCK,[▽] CHRISTINE SCHIPPA,[×] PIERRE-JEAN THOREL,[△] AND MATTHIAS VEY[◇]

Takasago, 23-25 Avenue De la Mare, F-95310 Saint Ouen l'Aumône, France; Firmenich SA, Corporate R&D Division, P.O. Box 239, CH-1211 Geneva 8, Switzerland; Givaudan, B.P. 98, F-95101 Argenteuil, France; Robertet SA, B.P. 52100, F-06131 Grasse Cedex, France; IFF, P.O. Box 5021, NL-5004 EA Tilburg, The Netherlands; Symrise, P.O. Box 1253, D-37601 Holzminden, Germany; Quest International, UK - Ashford, Kent TN24 0LT, United Kingdom; Treatt, Northern Way, Bury, St. Edmunds IP32 6NL, United Kingdom; V. Mane Fils SA, F-06620 Le Bar sur Loup, France; Givaudan SA, 5 chemin de la Parfumerie, CH-1214 Vernier, Switzerland; and IFRA, 6 Avenue de Arts, B-1210 Brussels, Belgium

The performances of the GC-MS determination of suspected allergens in fragrance concentrates have been investigated. The limit of quantification was experimentally determined (10 mg/L), and the variability was investigated for three different data treatment strategies: (1) two columns and three quantification ions; (2) two columns and one quantification ion; and (3) one column and three quantification ions. The first strategy best minimizes the risk of determination bias due to coelutions. This risk was evaluated by calculating the probability of coeluting a suspected allergen with perfume constituents exhibiting ions in common. For hydroxycitronellal, when using a two-column strategy, this may statistically occur more than once every 36 analyses for one ion or once every 144 analyses for three ions in common.

KEYWORDS: Allergens; GC-MS; quantification; skin reaction; fragrances; perfumes; variability; quantification limit

INTRODUCTION

In a previous paper, we proposed a quantification method (*1*) to determine the amount of 24 fragrance components regulated by a European Directive (*2*) as they are suspected of causing skin reactions. This GC-MS method, which will be referred to hereafter as the "IFRA method", is based on selected ion monitoring (SIM). One ion is used for the quantification, and two others are used as "qualifiers"; the ion abundance ratios are used to check the identity of analytes. In this first step, the limit of quantification (LOQ) and the variability of the determination were not investigated. The present work intends to determine these characteristics in the real context of a fragrance concentrate analysis achieved under "blind" conditions.

One of the first papers on the determination of suspected allergens (SAs) reported only a limit of detection (LOD) of 2 mg/L or below (*3*). This could lead to confusion between the LOD and the LOQ and could suggest that the quantification would be feasible down to this limit. In general, the LOQ is estimated at 10 times the signal-to-noise ratio, as exemplified in a recent paper on the analysis of SAs (*4*). This approach takes into account only the influence of the background and assumes that target peaks are well resolved from all others. However, fragrance concentrates are very complex mixtures; the occurrence of coelutions is frequent. In spite of the selectivity of the MS detection, the coelution of compounds exhibiting isobaric ions, in common with the targeted SAs, is not negligible. Therefore, this work will aim at determining a more realistic LOQ in the context of a fragrance concentrate.

Concerning the variability of analytical methods, the European Commission has stated that "the inter-laboratory coefficient of

* Corresponding author and chairman of the IFRA Analytical Working Group (telephone +41-22-780-3471; fax +41 22 780 33 34; e-mail alain.chaintreau@firmenich.com).
[†] Takasago.
[‡] Firmenich SA.
[§] Givaudan, Argenteuil.
[#] Robertet.
[||] IFF.
[⊥] Symrise.
[⊗] Quest International.
[▽] Treatt.
[×] V. Mane Fils SA.
[△] Givaudan SA, Vernier.
[◇] IFRA.

10.1021/jf062028l CCC: $37.00
Published on Web 12/14/2006

variation [...] shall not exceed the level calculated by the Horwitz equation" (*5*)

$$\mathrm{RSD} = 2^{(1-0.5 \log C)} \quad (1)$$

with C = the concentration of the target analyte.

The daily use of the IFRA method in quality assurance (QA) laboratories has shown that the two major causes of bias in the SA determination were coelutions and retention time shifts. The former case, already described above, is the most frequent one and can give rise to only an overevaluation of the target analyte. The second case occurs when a very abundant analyte elutes just before the targeted SA, which delays the elution of this latter component, sometimes out of its SIM window. This can give rise to its underevaluation or to a false negative. To overcome these drawbacks, a GC-MS analysis in scan mode with two different columns has been proposed (*4*). Characteristic ions of SAs are extracted from the scan file to achieve the quantification. If a given analyte ion is disturbed by an ion of another perfume constituent, an alternative ion, free of interference, is selected and used for the determination. As the scan mode does not require defining elution windows, in contrast to SIM, the risk of retention time (RT) shifts is less critical. However, the price to pay is a lower sensitivity and a greater variability in quantitation. With a quadrupole MS in scan mode rather than SIM mode (*6*), the precision of the GC peak integration decreases when the time allocated to monitor the corresponding ion during the peak is shortened [see IFRA pitfalls (*7*)]. Another strategy addressing the resolution of coelution has been proposed, starting from the IFRA method (*8*). As three ions per analyte are monitored simultaneously (one for the quantification and two for the identification), each of them can be used alternatively as the quantification ion to generate three calibration curves without additional injections. If a coelution occurs [as detected by the Q value (*1*)], the lowest amount resulting from the determination with the three calibrations must be the closest to the real amount. As for the RT shifts, the IFRA procedure recommends that a qualitative scan analysis be performed, which allows the detection of a major peak eluting just before the target analyte. If a shift in the retention time of the analyte is detected, then the times of the SIM window may be adjusted (*7*). The present work, based on the IFRA procedure, will use this strategy and extend it to the use of two different columns to evaluate the variability with fragrance concentrates, intentionally compounded, to represent some differing cases of complexity.

EXPERIMENTAL PROCEDURES

Materials. The suppliers of all standards together with their purity and SIM ions have been listed in the previous work (*1*). Fragrance concentrates were provided in blind to the present working group, by IFRA-member companies.

GC-MS Conditions. The quantification of SAs was achieved according to the IFRA method previously described (*1*). Instruments used included GC-MS 5973 from Agilent Technologies (Wilmington, DE), QP-2010 from Shimadzu (Kyoto, Japan), and DSQ from Thermo (Waltham, MA). The nonpolar and medium-polar GC columns were as proposed in the previous work (*1*). The participants in the interlaboratory test used a variety of polar columns: HP-Innowax and DB-Wax (Agilent), CP-Wax 52 CB (Varian, Walnut Creek, CA), and Stabilwax (Restek, Bellefonte, PA), of dimensions 30−60 m × 0.25 mm i.d. × 0.25 μm phase thickness. For these polar columns, the inlet (constant) pressure was 11.7−18.2 kPa. The initial oven temperature was 70 °C, then it was increased at 2 °C/min to 240 °C, and then maintained for 45 min. Each participant performed the SA determination in duplicate using two different phases.

GC×GC-MS Conditions. A Pegasus-4D GCxGC/TOF MS (LECO Corp., St. Joseph, MI) was used. It was equipped with a two-stage modulator, with a DB1 column (20 m × 0.18 mm × 0.18 μm) for the first dimension, and a DB225 column (1 m × 0.10 mm × 0.10 μm) for the second dimension, both from Agilent Technologies. The carrier gas was helium under a constant flow rate of 0.9 mL/min. The injector temperature was 240 °C. The oven program for the first column started at 60 °C for 3 min, then was ramped to 220 °C at a rate of 8 °C/min, with a final hold time of 5 min. The second column was continuously maintained at a temperature 20 °C higher than that of the first; that is, its temperature was ramped in parallel to that of the first column using the secondary oven. The transfer line into the TOF-MS source was heated at 225 °C and the electron impact ionization source itself operated at 200 °C with a collision energy of 70 eV. The data acquisition rate was 200 Hz over a mass range of 35−320 amu and a detector voltage of 1650 V. The modulation period was 3 s.

Calculation. The identification of peaks in SIM mode was based on the calculation of the Q value previously reported (*1*). A simulation of ion ratio deviations from the target values showed that the decision criteria (recognition/rejection) were close to those recommended by the European Commission (EC) (*5*). As the combination of both ion ratios in the calculation of Q results in a single value, this allowed a partial automation of the data treatment in a spreadsheet format, whereas the EC criteria, based on a more complex decision tree, would not make automation easy.

The mean standard deviation (SD) was computed as

$$\mathrm{SD} = \sqrt{\frac{1}{l}\sum_{i=1}^{i=l}(m_i - \overline{m}_i)^2}$$

The mean squared error (MSE) (or deviation from the truth) was calculated as

$$\mathrm{MSE} = \sqrt{\frac{1}{l}\sum_{i=1}^{i=l}(m_i - \mu_i)^2}$$

with l equal to the degree of freedom (number of laboratories minus one or number of analytes), and m_i, $\overline{m}_i$, and μ_i the determined amount, the mean of determinations, and the spiked amount of the analyte i, respectively. To obtain the accuracy profile, the 90% confidence intervals were calculated from the RMSEs of FT04 samples spiked at different SA levels (Supporting Information Table A-1) and for 27 degrees of freedom (28 independent compounds minus 1).

RESULTS

Limit of Quantification. The same fragrance concentrate (FT) was spiked at different levels with all 28 analytes used in the previous work (*1*). Six different spiking levels, plus a "zero level", were analyzed according to the IFRA method, using a DB17 column only.

If the Q value corresponding to the first ion was greater than 89, and if there was satisfactory agreement between the determinations using the three ions, then the amount given by the first ion was chosen.

In the case of coelution, the minimum of the three determinations, with respect to the three ions used for a specific compound, was chosen.

In the case of a RT shift, the peak was manually re-integrated using the correct time. In the present work, no shift was long enough to delay the peak out of the retention time window defined in the SIM conditions.

As the present investigation aims at determining the method reliability under the "blind" conditions of a real fragrance analysis, all determinations are reported without subtraction of the quantities found in the blank (nonspiked) sample (Supporting Information Table A-1). Because the model fragrance FT was

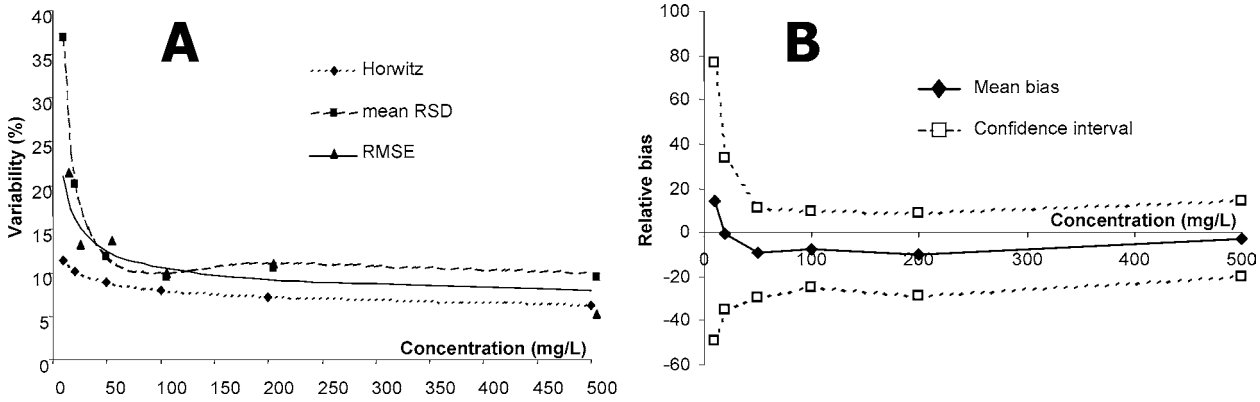

Figure 1. Determination by a single laboratory of a FT sample spiked at different concentrations of all SAs (**A**) experimental mean RSD and RMSE in comparison with Horwitz' prediction; (**B**) accuracy profile.

formulated without allergens, the amounts found in the blank sample are due to interfering compounds exhibiting ions in common with the target analytes. Such interferences must be taken into account to evaluate the method variability. For instance, a scan analysis showed that no hydroxycitronellal was present in the blank, in spite of the 10 mg/L suggested by the quantitation (although with a low *Q* value). Therefore, determinations of samples spiked with hydroxycitronellal were not corrected to allow for the possible interferences with compounds present in the nonspiked fragrance, as such interferences will occur under real conditions.

Interlaboratory Reproducibility. An interlaboratory experiment was conducted using five samples of fragrance concentrates: FT03, FT04, GIV02, IFF01, and SY630. Each of them was spiked with 10−15 SAs from the 28 SAs (the 24 SAs plus the 4 additional analytes investigated in the former work) (*1*) except FT04, which was spiked with the 28 analytes. Spiked amounts were in the range of 10−250 mg/L. All samples were analyzed blind by the different participants according to the IFRA method, but the analysis of each sample was duplicated using two different GC columns chosen among the following: nonpolar (polydimethylsiloxane-type), medium polar (DB17), and polar (Carbowax-type). The latter was provisionally chosen due to its wide availability in analytical laboratories, but more stable polar phases are being tested for future revisions of the procedure. In what follows, the above-described approach will be called the "two-column × three ion" strategy.

Each of the three monitored ions was successively used as a quantifying ion, while the two others became the qualifiers. Starting from a single injection of each combined calibration standard for each column, three analyte-specific calibration tables are thus obtained, as three ions per analyte are monitored simultaneously. From these calibration tables, three determinations are obtained per analyte and per column, that is, a total of six amounts and six *Q* values per compound. For each SA, its amount was determined as follows:

(1) Among all reported amounts, corresponding to a *Q* value higher than 89, the smallest one was chosen.

(2) When no Q value was higher than 89, the chosen determination was the smallest one, whatever the value of *Q* (except 0).

The amounts resulting from this data treatment are presented in Supporting Information Tables A-2 to A-6, together with the relative standard deviations (RSD), the relative mean squared error (RMSE), the number of false negatives, and the median of the *Q* value.

Table 1. Global Results with the Two-Columns × Three-Ions and Two-Columns × One-Ion Strategies (Eight Laboratories, Amounts below 10 mg/L Are Not Included)

		two-column × three-ion strategy			two-column × one-ion strategy	
sample	no. of constituents	mean RSD (%)	RMSE (%)	FN[a] median	mean RSD (%)	RMSE (%)
FT03	57	33	36	0	27	34
FT04	57	55[b]	60[b]	2	69[c]	78[c]
IFF01	39	56	61	0.5	54	60
GIV02	50	49	57	1	65	68
SY630	49	32	32	0	33	40

[a] Median of false negatives. [b] Results derived without farnesol. With farnesol: RSD = 114 and RMSE = 113%. [c] Results derived without including farnesol.

DISCUSSION

Limit of Quantification. A fragrance concentrate (FT) was spiked with all allergens at various levels between 10 and 500 mg/L and analyzed by a single laboratory (Supporting Information Table A-1). Some analytes systematically led to a poorer evaluation than others, such as the second farnesol isomer and amylcinnamic alcohol. The amount of the former was evaluated using its first isomer only, as the second isomer coeluted with diethylphthalate

Using the above-described strategy, global relative standard deviations from the mean (RSD) and from the expected mean value (= relative mean squared error, RMSE) were calculated for the 28 compounds and are summarized in **Figure 1A**. The RMSE remained below 22% for all investigated concentrations. The RSD was ≤20% at all levels except at 10 mg/L, where a significant variability increase occurred.

The present mean RSD does not represent a real repeatability evaluation, because the quantification of each spiked sample was not replicated, as achieving a calibration plus several replications is not feasible within a single day [risk of MS detector drift over time (*1*)]. However, the determination of the seven samples was obtained under repeatability conditions; that is, it was carried out in the same laboratory by the same analyst using the same instrument. The homogeneity of the mean RSD and RMSE in the ranges of 50−500 and 20−500 mg/L, respectively, suggests that the observed variability is close to the repeatability. This variability is itself close to the prediction of the Horwitz equation (**Figure 1A**), although the latter is said to represent reproducibility. This observation will be discussed after evaluation of the real reproducibility (*vide infra*).

The limit of quantitation of an analytical procedure is defined as "the lowest amount of the targeted substance in the sample

Table 2. Probability of Singlet, Doublet, and Triplet (P_1, P_2, P_3) Occurrence with a 150-Constituent Mixture, Using the Three Types of Columns Used in This Work

column	i.d. (mm)	length (m)	N	t_0 (min)	t_n (min)	peak capacity	distinct peaks	P_1	P_2	P_3
OV17	0.18	20	111111	0.57	17	2403	141	0.883	0.053	0.003
OV1	0.25	60	240000	2.00	25	1409	135	0.808	0.082	0.008
CW	0.25	60	240000	4.17	120	3403	144	0.916	0.039	0.002

Table 3. Risk of Coelution of Hydroxycitronellal and Isoeugenol with a Compound Exhibiting One and Three Ions in Common (Π_1 and Π_3), in a Nonpolar Column

	$(1 - P_1)$	m/z	U^a	$\pi(f_i)$	Π_1	Π_3
hydroxycitronellal	0.192	59	2	0.25	0.048	0.024
		71	1	0.5	0.096	
		43	0	1	0.192	
isoeugenol	0.192	164	2	0.25	0.048	0.003
		103	2	0.25	0.048	
		149	2	0.25	0.048	

[a] U value after Pesyna (*19*).

Table 4. Comparison of the Risk of Coelution with Common Ions Using the Three Strategies

	two columns[a] × three ions $\hat{\Pi}_3$	two columns[a] × one ion $\hat{\Pi}_1$	one column[b] × three ions Π_3
citronellol	6.3×10^{-5}	4×10^{-3}	1.2×10^{-2}
hydroxycitronellal[c]	2.5×10^{-4}	1×10^{-3}	2.4×10^{-2}
isoeugenol	3.9×10^{-6}	1×10^{-3}	3.0×10^{-3}

[a] Nonpolar and polar columns. [b] Nonpolar column. [c] Not present in FT03, only as a numerical example.

which can be quantitatively determined under the experimental conditions prescribed with a well defined accuracy" (*9*). To date, this very recent approach has mostly been applied in the pharmaceutical area (*10*−*12*). Using the accuracy profile of the method, it consists of the determination of the lowest concentration for which the corresponding confidence interval remains within a given acceptance limit (*13*). In the present case, the accuracy profile shows that the mean bias remains less than 20% at all spiking levels down to 10 mg/L (**Figure 1B**). For 90% of determinations, the expected bias should be less than 35% down to a level of 20 mg/L and between −49 and 77% at 10 mg/L (i.e., between 5 and 18 mg/L). This range remains acceptable to set the LOQ at 10 mg/L, in view of the SA analysis complexity. It must be emphasized that the present determination validates only the LOQ for the single-column procedure used for this intralaboratory test.

Interlaboratory Variability. From the detailed quantification results presented in the Supporting Information Tables A-2 to A-6 and based on the two-column × three-ion strategy described under Results, RSDs and RMSEs were averaged over all spiked analytes of each sample and over all participants, and the median of false negatives was calculated (**Table 1**). Concentrations below the above-mentioned LOD of 10 mg/L were not taken into account. Eight laboratories with various levels of experience with the IFRA method participated in these ring tests (from a few determinations per year up to a daily practice).

Due to the two-column × three-ion strategy, false positives were very rare. When it occurred for one or several determinations among the six values resulting from this strategy (three ions × two columns), selecting a minimum, as described under Results, led to values below 10 mg/L. As this was below the LOQ, such values were rejected (except those in FT04). Therefore, only the possible occurrence of false positives is discussed hereafter.

FT03 and SY630 exemplify the degree of complexity of typical samples analyzed in quality control laboratories in the fragrance industry. Very few false negatives were observed in these two samples, except for amylcinnamic alcohol in SY630 (Supporting Information Table A-6). The recognition of analytes was good as shown by the excellent Q values, except for geraniol and amylcinnamic alcohol. Both the global RSD and the global RMSE were in the range of 32−36% (**Table 1**). This experimental reproducibility is about 3−4 times greater than the prediction of Horwitz's equation. This is due to the fact that the peaks of analytes are rarely free from interference in spite of the selectivity of the mass spectrometer. This will be discussed in detail hereafter with two critical examples, GIV02 and IFF01.

FT04. This sample was spiked with all analytes at a concentration of 10 mg/L to check whether the limit of quantification previously proposed was applicable under interlaboratory conditions. The RSD and the RMSE lay in the range of 55−60%, without taking into account the determination of farnesol. This latter caused a huge increase in variability that doubled the global RSD and RMSE. These results (Supporting Information Table A-3) suggest that 10 mg/L is an acceptable provisional LOQ in this interlaboratory context, as it corresponds to a reasonable increase of the RSD and the RMSE compared with their values in samples FT03 and SY630, where the target analytes are present in the range of 25−250 mg/L. This provisional LOQ will have to be reassessed in a more structured way after choosing which of the strategies is most appropriate among the options that were investigated in the present work.

GIV02. In this sample, the two poorest determinations concerned hydroxycitronellal and isoeugenol (occurrence of false negatives and poor recognition, see Q value, Supporting Information Table A-4). Whatever the column, the three SIM ions of hydroxycitronellal were hardly visible simultaneously, which biased the Q value calculation. As the ring-test procedure required a scan analysis to be performed in parallel to the quantification in SIM, the presence of hydroxycitronellal was searched *a posteriori* in the raw data corresponding to the three false negatives mentioned in Supporting Information Table A-4. Using the three GC stationary phases in this test, hydroxycitronellal coeluted with other compounds. Its base peak at (m/z 59) was scarcely distinguishable sometimes from the background noise. As the scan mode cannot achieve the same sensitivity as the SIM quantification for which an overvoltage of the photomultiplier was applied, the concerned laboratories logically concluded that it was absent.

A similar situation arose in the case of isoeugenol.

IFF01. The global RSD and RMSE were as high as for the sample at the LOD level (FT04) (**Table 1**), mainly because of the poor RMSE of amylcinnamic alcohol and the RSD of citronellol (Supporting Information **Table A-5**). The latter analyte coeluted with a component exhibiting a similar MS

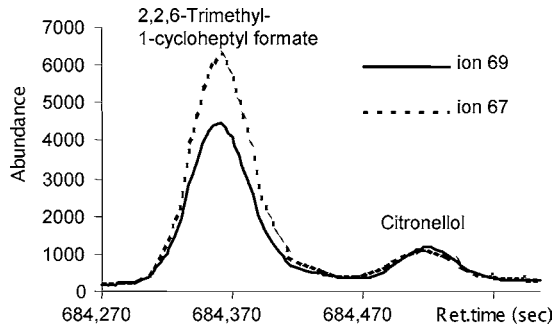

Figure 2. GCxGC/TOF-MS analysis of IFF01: second-dimension chromatogram (DB1) corresponding to the citronellol retention in the first dimension (DB225).

spectrum, which biased the *Q* value calculation and consequently its recognition. The fact that this fragrance was formulated with the lowest number of constituents among the five samples and was spiked with SAs in a similar concentration range as the others (except FT04) strongly suggests that the difficulty of the analysis is more related to the nature of the ingredients than to their number. This case is discussed in detail hereafter (Risk of Coelution).

Risk of Coelution. The experimentally observed risk of coelution reported in the Introduction can be roughly predicted from published relationships. In a chromatogram, the number of distinguishable peaks has a theoretical limit, *p* (*14*)

$$p = me^{-m/n_c} \tag{2}$$

where *m* is the number of components in the sample and n_c is the peak capacity of the column.

For a programmed temperature elution, the peak capacity can be estimated as (*15*)

$$n_c = 1 + \frac{\sqrt{N}}{4}\left[\frac{t_n}{t_0} - 1\right] \tag{3}$$

where *N* is the number of plates of the column and t_0 and t_n are the retention time of an unretained compound and that of the last eluted peak, respectively; *N* can be estimated from the column length and the theoretical plate height, which is approximately equal to the inner column diameter (*16*).

The probability P_n of observing a peak consisting of *n* component(s) can be calculated using Giddings' theory (*17*):

$$P_n = (e^{-2\frac{m}{n_c}})(1 - e^{-\frac{m}{n_c}})^{n-1} \tag{4}$$

These values have been calculated for the three different columns that have been used in this work (**Table 2**). If we assume that the number of constituents, at the 5 ppm level, is about 150 (57 ingredients + 28 standards + 2 internal standards + impurities of all these compounds), the DB17 column would be capable of distinguishing 141 peaks (in practice, 137 peaks were found after integration of the GC file of FT03). From these 141 only 132 would be composed of a single compound, 8 should be doublets, and 1 would be an *n*-uplet with $n \geq 3$. This calculation represents an ideal situation in which the constituents are chemically independent and statistically distributed along the whole chromatogram. In the real context of fragrances, many analytes belong to a small number of functional classes, such as mono- or sesquiterpene alcohols. As members of the same class elute in a short portion of the chromatogram, the probability of observing multiplets is much higher than predicted by eq 4.

It is commonly recognized that an MS detector tolerates GC coelutions due to its capability of selectively monitoring characteristic ions. This is true as long as the overlapping peaks do not themselves exhibit similar spectra. To achieve quantification, a minimum number of ions are monitored (three in the present case). If the abundance of one of them is altered due to the presence of an isobaric ion at the same retention time, the *Q* value calculation, and hence the analyte recognition, is jeopardized. Such a risk can also be approximately evaluated due to the statistical investigation of MS abundances published by McLafferty (*18*): "*The probability of finding a peak of a particular mass at any abundance level is defined as*"

$$\pi(f_i) = {}^1/_2{}^{(U+A)} \tag{5}$$

where *U* reflects the "uniqueness" of the fragment f_i and *A* reflects the influence of the fragment abundance. If we consider the risk that a given analyte coelutes with a compound in which the abundance level of the common fragment is at least 1%, then $A = 0$. The uniqueness values were calculated using libraries of 18806 spectra. Although they were not updated with the advent of the current large libraries, the old values were assumed to be still applicable to the relatively "small" molecules used in fragrances as the latest spectra mainly represented high molecular weight molecules. Therefore, the probability of finding either this single fragment f_i, or the three ions f_i, f_j, f_k, simultaneously in the spectrum of the coeluted compound is $\pi(f_i)$ and $\pi(f_i)\pi(f_j)\pi(f_k)$, respectively.

If we now combine the coelution probability in a single column to that of the occurrence of identical fragments in both spectra, the risk Π of observing a biased ion abundance of a given analyte is

for 1 ion: $$\Pi_1 = (1 - P_1)[\pi(f_i)] \tag{6}$$

for the three ions, simultaneously:

$$\Pi_3 = (1 - P_1)[\pi(f_i)\pi(f_j)\pi(f_k)] \tag{7}$$

Using the previously discussed examples of hydroxycitronellal and isoeugenol that were sometimes missed in sample GIV02, the risk of coelution was calculated (**Table 3**). The risk that a compound coelutes with hydroxycitronellal in the apolar column *and* exhibits a fragment in common with its spectrum base peak was approximately 5%. The risk was still approximately 2.4% that the three SIM ions also occurred in the spectrum of the same unknown. This situation cannot be improved by choosing alternative ions from the hydroxycitronellal spectrum, as these were not abundant enough to achieve a satisfactory LOD.

Using a nonpolar column, the risk of coeluting isoeugenol with a compound exhibiting three ions in common was lower than in the previous case (0.3%, **Table 3**), due to the greater specificity of the three SIM ions.

The present discussion also justifies the two-column × three-ion strategy as the risk, $\hat{\Pi}$, that a coelution with isobaric ions occurs on both columns for the same analyte becomes

for a single ion: $$\hat{\Pi}_1 = (1 - P_1)(1 - P_1')[\pi(f_i)]^2 \tag{8}$$

for the three ions:

$$\hat{\Pi}_3 = (1 - P_1)(1 - P_1')[\pi(f_i)\pi(f_j)\pi(f_k)]^2 \tag{9}$$

where P_1' is the probability of observing a singlet in the second column.

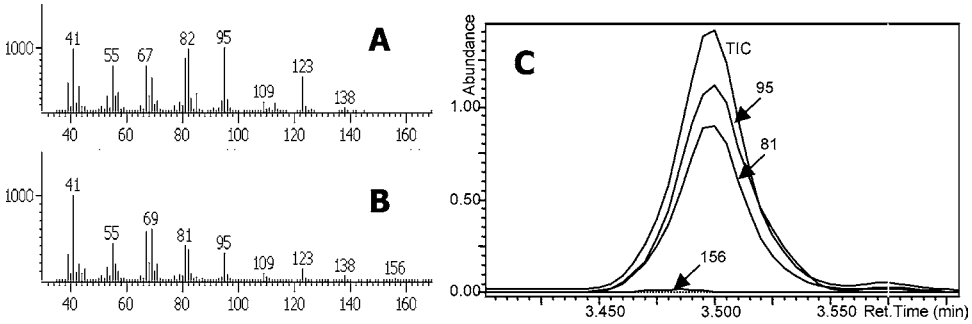

Figure 3. Spectra of (**A**) 2,2,6-trimethyl-1-cycloheptyl formate (?) and (**B**) citronellol and (**C**) GC-MS trace of the citronellol zone (in scan) in the IFF01 sample.

Table 5. FT03 Results with the One Column × Three Ions Strategy (Eight Laboratories, Amounts below 10 mg/L Are Not Included)

column	mean RSD	RMSE	no. of labs
DB1	35	42	4
DB17	42	49	4
CW-type	135	154	5

In the case of hydroxycitronellal in GIV02, and using as an example the nonpolar and polar columns (**Table 2**), the probability was not negligible: $\hat{\Pi}_1 = 1 \times 10^{-3}$ or $\hat{\Pi}_3 = 2.5 \times 10^{-4}$ for just the quantification ion or the three SIM ions, respectively. This represents a potential risk of observing a biased *Q* value with both columns for 1 of the 28 analytes every 36 and 144 samples, respectively. In practice, however, this occurs much more frequently due to the small number of chemical families used in fragrance formulations as previously mentioned. In the case of isoeugenol, this risk was even lower: $\hat{\Pi}_1 = 1 \times 10^{-4}$ and $\hat{\Pi}_3 = 3.9 \times 10^{-6}$, respectively (**Table 4**).

Comparing Π_3 and $\hat{\Pi}_3$ values shows the reliability improvement when using two columns as compared to a single column. Once again, these numerical values do not accurately reflect the real probabilities as found in fragrances, as compounds of the same chemical family are often present in a given formula. They may not only have similar retention times, as discussed above, but their fragmentation patterns can also exhibit some common similarities, as demonstrated in the following example.

Sample IFF01 well exemplifies the occurrence of coelutions on different phases. Whereas the use of two different phases is often considered as a valid chromatographic strategy to overcome this risk, citronellol was missed by five of the eight laboratories in the analysis of this particular sample. In this example, and irrespective of the column combination chosen, citronellol was masked to a lesser or greater degree by a coeluting larger analyte. Using a polar phase (e.g., CP 52CB), it was found to coelute with (*E*)-citronellyl nitrile, which exhibited ions at *m*/*z* 69 and 81 and which interfered with those used for its quantification and identification due to their influence on the *Q* value. Using a medium-polar phase (DB17), citronellol was coeluted with 1-(3,3-dimethyl-1-cyclohexyl)ethyl formate. Again, both compounds have ions in common that interfered with those of citronellol.

To investigate the reason for its being missed with a nonpolar column, the sample was reanalyzed by comprehensive two-dimensional GC (GC×GC), using TOF-MS detection. The first GC column was identical to the nonpolar column recommended in the IFRA method. The retention time elution zone of citronellol in the first dimension resulted in two peaks being separated by the second (polar) dimension (**Figure 2**). One of these was confirmed as citronellol; the second one exhibited a mass spectrum having all of the key fragment ions in common with citronellol (**Figure 3A**,**B**). It was tentatively identified as 2,2,6-trimethyl-1-cycloheptyl formate, a compound that occurs as a secondary, minor, component in the above-mentioned 1-(3,3-dimethyl-1-cyclohexyl)ethyl formate. With the conventional GC-MS system (**Figure 3C**), not only were the ion abundance ratios (hence the *Q* value) biased, but its perfect coelution with citronellol made the latter analyte undetectable. As a consequence, the choice of alternative quantification ions, or the use of whole mass spectrum in scan mode, would not solve the problem as there are no specific ions that appear to be unique to only one of the two components.

Other Strategies. To shorten the data treatment, two simplified strategies were compared to the two-column × three-ion approach: two columns and a single quantification ion per column (and two qualifiers) and one column and three quantification ions. Raw data resulting from the analysis of the ring test samples were re-treated.

Two Columns × One Ion. For each of the two columns, only the first ion proposed as the quantification ion in our previous work (*1*) was used to determine the analyte amount, the two other ions being used only as qualifiers. As expected, the uncertainty increased for concentrations at the LOQ (FT04, **Table 1**) and also for one of the complex samples (GIV02). However, in comparison to the two-column × three-ion treatment (**Table 1**), the RSD and the RMSE remained stable for the other complex sample (IFF01), as well as for SY630 and FT03.

From the viewpoint of the coelution risk of analytes exhibiting a fragment identical to the quantification ion, the probability was higher in the three cases calculated in **Table 4.**

This and the experimental results suggest that the one-ion × two-column strategy could be used if a somewhat higher uncertainty is accepted. This could be true particularly at low concentrations.

One Column × Three Ions. This previously proposed strategy (*1*) was compared to the two others. The data retreatment was achieved for only one sample, FT03. Only four to five determinations could be used per column due to an incomplete data set. With nonpolar and low-polar phases, results were close to those of the two-column × three-ion strategy. With the Carbowax-type columns, results were disappointing. This might in part be due to the variations in phase chemistry and performance between columns from different manufacturers. As the five laboratories mentioned in **Table 1** used columns from four different manufacturers, this might have had an influence on the accuracy and precision of the determinations. Therefore, these results do not allow a clear conclusion to be drawn on the one-column × three ion strategy without further standardization of the analytical conditions with such columns. However,

the numerical simulation using three analytes (**Table 5**) seems to confirm that this strategy would be the less efficient one among the three investigated approaches to overcome the risk of coeluted compounds with isobaric ions.

From these three strategies, the two-columns × three-ion option best minimizes the consequences of coelution on the determination of SAs. However, whatever the strategy, some exceptions may persist that lead to either false negatives or an increase in the determination variability. The one-column × three ion results suggest a less reliable performance than the two other approaches. The performances of the three strategies are confirmed by the statistical evaluation of the coelution risk. Coelution cannot be avoided and can sometimes occur simultaneously on two different columns. This also demonstrates that the determination of such coeluted compounds with related structures would be unlikely to be overcome by the application of more sophisticated techniques such as signal deconvolution (similar mass spectra) or high-resolution MS (same raw distribution of fragments with a weak molecular ion).

ACKNOWLEDGMENT

We gratefully acknowledge Prof. F. W. McLafferty (Cornell University, Ithaca, NY), Dr. D. Staufer (Palissade Co., Newfield, NY), and Dr. S. Rudaz (University of Geneva, Switzerland) for their advice in mass spectrometry and method validation.

Supporting Information Available: Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) Chaintreau, A.; Joulain, D.; Marin, C.; Schmidt, C. O.; Vey, M. GC/MS quantitation of fragrance compounds suspected to cause skin reactions. 1. *J. Agric. Food Chem.* **2003**, *51*, 6398−6403.

(2) European Parliament Directive 2003/15/EC of the European Parliament and of the Council of 27 February 2003 amending Council Directive 76/768/EEC on the approximation of the laws of the Member States relating to Cosmetic Products. *Off. J. Eur. Communities* **2003**, *L 66*.

(3) Ellendt, K.; Hempel, G.; Köbler, H. Analysis of sensitizing fragrances by gas chromatography−mass spectrometry. *SOFW J.* **2001**, *127*, 29−34.

(4) Leijs, H.; Broekhans, J.; Van Pelt, L.; Mussinan, C. Quantitative analysis of the 26 allergens for cosmetic labeling in fragrance raw materials and perfume oils. *J. Agric. Food Chem.* **2005**, *53*, 5487−5491.

(5) 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. Communities* **2002**, *L221*, 8−36.

(6) Gross, J. H. *Mass Spectrometry. A Textbook*; Springer: Berlin, Gemany, 2004.

(7) International Fragrance Association (IFRA). GC/MS quantitation of potential fragrance allergens in fragrance compounds; http://www.ifraorg.org/, 2003.

(8) Debonneville, C.; Thomé, M.-A.; Chaintreau, A. Hyphenation of quadrupole MS to GC and comprehensive-two dimensional GC for the analysis of suspected allergens: review and improvement. *J. Chromatogr. Sci.* **2004**, *42*, 450−455.

(9) Hubert, P.; Nguyen-Huu, J. J.; Boulanger, B.; Chapuzet, E.; Chiap, P.; Cohen, N.; Compagnon, P. A.; Dewe, W.; Feinberg, M.; Lallier, M.; Laurentie, M.; Mercier, N.; Muzard, G.; Nivet, C.; Valat, L. Validation of quantitative analytical procedure, harmonization of approaches. *STP Pharma Prat.* **2003**, *23*, 101−138.

(10) Boulanger, B.; Chiap, P.; Dewe, W.; Crommen, J.; Hubert, P. An analysis of the SFSTP guide on validation of chromatographic bioanalytical methods: progresses and limitations. *J. Pharm. Biomed. Anal.* **2003**, *32*, 753−765.

(11) Jacobs, P.; Dewe, W.; Flament, A.; Gibella, M.; Ceccato, A. A new validation approach applied to the GC determination of impurities in organic solvents. *J. Pharm. Biomed. Anal.* **2006**, *40*, 294−304.

(12) Rbeida, O.; Christiaens, B.; Chiap, P.; Hubert, P.; Lubda, D.; Boos, K.-S.; Crommen, J. Fully automated LC method for the determination of sotalol in human plasma using restricted access material with cation exchange properties for sample clean-up. *J. Pharm. Biomed. Anal.* **2003**, *32*, 829−838.

(13) Feinberg, M.; Laurentie, M. A global approach to method validation and measurement uncertainty. *Accredit. Qual. Assur.* **2006**, *11*, 3−9.

(14) Bertsch, W. Two dimensional gas chromatography. Concepts, instrumentation, and applications−part 1: fundamentals, conventional two-dimensional gas chromatography, selected applications. *J. High Resolut. Chromatogr.* **1999**, *22*, 647−665.

(15) Grushka, E. Chromatographic peak capacity and the factor influencing it. *Anal. Chem.* **1970**, *42*, 1142−1147.

(16) Tranchant, J.; Arpino, P.; Prévôt, A.; Serpinet, J.; Vergnol, A.; Witier, P. *Manuel Pratique de Chromatographie en Phase Gazeuse*, 4th ed.; Masson: Paris, France, 1995.

(17) Davis, J. M.; Giddings, J. C. Statistical theory of component overlap in multicomponent chromatograms. *Anal. Chem.* **1983**, *55*, 418−424.

(18) Pesyna, G. M.; McLafferty, F. W.; Venkataraghavan, R.; Dayringer, H. E. Statistical occurence of mass and abundance values in mass spectra. *Anal. Chem.* **1975**, *47*, 1161−1164.

(19) Pesyna, G. M. Computerized structure retrieval and interpretation of mass spectra: the design and evaluation of a probability based matching system using a large data base. Ph.D. Thesis, Cornell University, 1975.

Received for review July 19, 2006. Revised manuscript received October 18, 2006. Accepted October 20, 2006.

JF062028L