Improved Hardware and Software for Quick Gas Chromatography–Olfactometry Using CHARM and GC–“SNIF” Analysis

C. Debonneville,¹ B. Orsier,¹ I. Flament,¹ and A. Chaintreau*,¹

Corporate R&D Division, Firmenich SA, P.O. Box 239, 1211 Geneva 8, Switzerland, and Varian-JMBS SA, 1, Rue Hector Berlioz, Parc d’Activités des Plans, 38600 Fontaine, France

A multisniffing system has been developed to allow three panelists to simultaneously participate in a GC-olfactometric analysis. This device, associated with a computerized data treatment, allows shortening CHARM and GC–“SNIF” analyses to less than 1 week and less than 1 day, respectively. The program was developed as an extension of an existing commercial chromatography data system, as usual GC processing functions are suited to the treatment of olfactograms (plots of odor response versus GC elution time). Because of the improved algorithm, the consequences of gaps in coincident responses were minimized, and the systematic use of a panel improved the repeatability of CHARM olfactograms. Comparing both methods, GC-SNIF repeatability appears to be higher than that of CHARM, as the former method uses a larger panel, but in a shorter lapse of time.

The detection of peaks eluting from a gas chromatographic column by the human nose has now become a classical tool in the field of flavor and fragrance analysis. The acquisition of olfactograms and the data treatment have been rationalized according to three main techniques.

(1) Dilution methods: CHARM¹² (combined hedonic aroma response measurement) and AEDA²³ (aroma extract dilution analysis). An extract of an aroma is injected in increasing dilutions until the panelist does not smell any odor at the column outlet.

(2) Time intensity: OSM²⁴ (osm in Greek), FSCM⁵ (finger span cross-modality matching). The extract is only injected at a single dilution level and the panelists continuously record the odor intensity perceived at the sniff port by moving a rheostat.

(3) Detection frequency: GC–“SNIF” (surface of nasal impact frequency).⁵⁻⁸ From the injection of a single dilution level, the percentage of panelists who detect an odor is computed over the whole duration of the GC run. Units of olfactogram peak heights and areas have been called NIF (nasal impact frequency) and SNIF from which comes the method’s name.

Details about the procedures may be found in the literature cited above. The applications have been reviewed.⁹–¹⁷ Despite their unique capability to determine the key odorants responsible for a given odor, existing techniques still suffer from a major drawback: the duration of the analysis. Compared to a usual GC or GC/MS run that is achieved within 1–2 h, a gas chromatography–olfactometry (GC–O) analysis still takes days or weeks to perform.

Dilution Methods. To avoid gaps in coincident responses,¹⁸ Schieberle recommended AEDA be performed within 2 days.¹⁹ However, since AEDA is normally performed by only one or two panelists, this does not overcome lacks of perception due to specific anosmia.¹⁷ On the basis of a statistical treatment of CHARM results, Acree concluded that there is a need to triplicate the analysis to achieve a satisfactory repeatability of the olfactogram.¹⁹ This was confirmed by Guichard et al.²⁰ For a series of 10 dilutions, 30 GC injections are needed, which requires about 2 weeks. Despite these conclusions, recent CHARM publications about real odors do not state that a panel is used.¹¹²

¹ To whom correspondence should be addressed. E-mail: alain.chaintreau@firmenich.com.
² Firmenich SA.
³ Varian-JMBS SA.
¹¹⁶ Chaintreau, A. In Flavor, Fragrance and Odor Analysis; Marsili, R., Ed.; M. Dekker: New York, 2002; Chapter 12.
Time Intensity. The original OSM procedure involved a panel of four persons. The average coefficient of variation of each panelist was found to be 9–13%.[23] However, the coefficient of variation averaged on peaks and two panels of five members gave a value of 30%.[24] From the statistical analysis of the FSCM results generated by seven persons, Tistlehay[5] observed “poor individual performances” and concluded that a panel was necessary. In addition to the replicated injections needed for the different members of the panel, OSM and FSCM require a preliminary training session whose duration is not indicated.[4,5]

Detection Frequency. Because the GC-SNIF method is based on the response of a whole panel, the role of inattention, specific anosmia, etc., on the final olfactogram is minimized. However, 6–10 panelists are required for a repeatable result, which means the same number of injections. Although GC-SNIF has been said to be more rapid than dilution and intensity methods,[24,25] it still takes 2–4 days.

The present work intends to evaluate the combination of a multisniffing system with a computerized data treatment of GC–O signals to shorten the analysis, without compromising the repeatability of the olfactogram. Only CHARM and GC-SNIF methods will be investigated in this publication. The OSM procedure will be treated in a forthcoming paper.

EXPERIMENTAL SECTION

Compounds and Flavor Model. Eucalyptol, (Z)-3-hexenol, citronellal, and benzyl acetate were obtained from Fluka (Buchs, Switzerland), β-ionone was from BASF (Ludwigshafen, Germany), and octanal was from Firmenich. A stock solution was prepared in ethanol according to Table 1. For the CHARM analyses, this solution was put in a glass cartridge of the ATD400 headspace injector (see below) containing silanized glass wool and injected into the GC column.

GC/MS–Olfactometry. Analyses were performed with a GC/MS system (Varian, Walnut Creek, CA) made of a gas chromatograph, model 3400, connected to a Saturn-3 ion trap detector. Volatiles were thermally desorbed from Tenax (250 °C, 5 min), using an ATD400 thermal desorber (Perkin-Elmer Corp., Norwalk, CT). They were refocused in an internal Tenax cold trap (−30 °C) and desorbed at 250 °C for 5 min into a Supelcowax column (Supelco, Buchs, Switzerland) (30-m length, 0.53-mm i.d., 1.00-μm phase thickness). Helium was used as carrier gas under an inlet pressure of 10.7 psi. The column was kept at 70 °C for 3 min, increased at a 15 °C/min rate to 240 °C, and maintained for 10 min. About one-fourth of the column flow was diverted to the mass spectrometer (Figure 1, splitter I). The remaining flow and the makeup flow (2 mL/min) were directed to splitter II. Both splitters were “universal connectors” (Gestel, Milhein, Germany). The mass spectra were acquired with a source temperature of 170 °C, under a 70-eV ionization potential.

Multisniffing System. The scheme of the multisniffing system is shown in Figure 1. The flow entering splitter II was divided into three equal parts toward three thermostated transfer lines that were maintained at 250 °C. These lines were made of deactivated silica capillaries of identical size (3-m length, 0.32-mm i.d.). The air makeup (4 mL/min) entering the connection box ensured the homogeneous temperature of lines 1–3. Each line was ended by a Dewar-type sniff port (silver-coated double-jacketed glass port under vacuum), to isolate the panelist’s nose from the hot parts of the transfer line.

Each sniff port was equipped with an electric push-button to generate a signal of 1 V when pressed. The three signals were collected by the Borwin digitalizer (Varian-JMBS SA, Fontaine, France) and transmitted to the computer.

Software. The software for the GC–O data treatment was developed as a plug-in extension of the Galaxy chromatography data system (Varian-JMBS SA). The same algorithm was used for GC-SNIF and CHARM analyses (vide infra, eq 3). Variables were initialized according to the method chosen, as described in the Discussion.

The validity of all test olfactograms generated by Galaxy was verified by performing separately their reconstruction from ASCII files using Excel (Microsoft, Redmond, WA). The peak height of reconstructed olfactograms were scaled as NIF percent (proportion of the panel detecting a given odorant) or FD factor (flavor dilution factor above which a given odorant is not perceived...
Corresponding peak areas are given as SNIF units (1 unit = 1% NIF 0.1 s) or CHARM units (1 unit = 1 FD 0.1 s).

Sniffing Experiments. Three panelists participated in the CHARM experiments (CHD, BRA, SCR) and nine in the GC-SNIF experiments (CHD, BRA, SCR, ALC, ALV, SJS, SCF, MGT, AXC). They were considered as trained because they had also previously participated in other GC-O and sensory analysis panels but had not received specific training for this investigation. To avoid lassitude and alteration of the panel's performance, judges only participated in one GC-O session per half-day.

DISCUSSION

Hardware. The advantage of using a panel in GC-O is now well established, but the main drawback of such an approach is the time-consuming repetition of GC injections as mentioned in the introduction. To limit the duration of the analysis, a multisniff port was built to allow three panelists to simultaneously smell the odorants from a single GC run (Figure 1). The performances of the system were compared to the previous experimental protocol based on replications of the GC-O run using a single sniff port. The model mixture was smelled by three panelists, either simultaneously using the three sniff ports or successively using a same sniff port (Figure 1). The 95% confidence interval of the initial odor detection time was comparable for both experiments and remained equal or less than 0.03 min (Figure 2). This suggests that the flows through the three transfer lines between splitter II and the sniff ports were identical and that both GC-O protocols were equivalent. When the end times of odorant detections were

![Figure 1. Scheme of the multisniffing hardware.](image1)

![Figure 2.](image2) Confidence intervals (95%) of the start times of olfactometric peaks: three panelists using three sniff ports (+); three replications by the same three panelists using consecutively the same sniff port (×).
compared, confidence intervals were larger (up to 0.09 min) in agreement with previous observations made with a mono-sniff-port system (data not presented). This fact might result from two reasons: some panelists could be in the adaptation phase and do not perceive the odorant when its intensity begins to decrease, and some others could perceive it after the peak end due to a persistence in the olfactive system or in the sniff port.

As a second test, the olfactogram of the model mixture (Figure 3) was generated with a panel of nine persons (three sessions of three panelists at once) and duplicated. The two average olfactograms were obtained using the software described in the next section. The average standard deviation (SD) between both experiments was 11.2% NIF, corresponding to a variation of (one panelist (contribution of one panelist to the olfactogram, 11.1% NIF) (Table 1). The SNIF repeatability (mean RSD 19.0%) was slightly lower than our previous result (13.8%). This was mainly due to the difficulty of panelists to clearly detect the end of the olfactive peak as mentioned above.

These results are quite similar to those previously obtained with a mono-sniff-port system and shows that the multisniffing device is suitable for the GC-SNIF method.

**Data Treatment Software.** In addition to the timesaving development of the multisniffing hardware, the data treatment had to be automated to minimize the duration of a GC-O analysis. To do this, a software was developed to generate an olfactogram according to two of the three main GC-O techniques: dilution and detection frequency methods. The properties of the reconstructed CHARM or GC-SNIF olfactograms were the same as the usual chromatograms, thus allowing us to save the files, integrate, and calculate the areas, the heights, and the retention indices of the peaks, using the standard chromatographic features of the Galaxie chromatography data system.

When a panel was used, several authors followed the individual responses of the panelists. Although this may bring some additional information, this approach obviously contributes to the time consumption that this study aims to shorten. In the present paper, the response of the whole panel was considered as a single detector signal, whose variability must be known to determine the confidence level the analyst has in the resulting average olfactogram.

**CHARM Analysis.** According to the formula proposed by Acree, the intensity of the kth point of the final olfactogram is given by

$$A_k = c^{n-1}$$

where c is the dilution factor between two consecutive injections, k is the olfactogram point at retention time $t_k$, and n is the number of coincident positive responses at time $t_b$. Such a calculation has two drawbacks: (1) When a gap of coincident responses occurs from one dilution level to another (i.e., when a given odorant has not been perceived at a given dilution, but it is smelled again in a further dilution of the extract), then the impact of all positive responses occurring at lower dilution levels is underestimated. For instance (Figure 4), if $c = 3$ and a gap occurs for a given peak at dilution 9, whereas the odor is still perceived at dilution 27, then $A_k = 9$ (Figure 4, dotted line), instead of 27 if no gap had occurred. In our opinion, it would be logical to apply the correction factor $(f_i - f_{i-1})$ corresponding to the ith dilution to each point of the ith individual olfactogram and then to sum the corrected coincident responses:

$$A_k = \sum_{i=1}^{n} (f_i - f_{i-1}) A_{i,k}$$

where $A_{i,k}$ is the signal from the individual olfactogram i, at time $t_k$ ($A_{i,k} = 0$ or 1), n is the number of dilution levels (if no dilution:

![Figure 3. Averaged olfactograms of the model mixture using the CHARM and the GC-SNIF method.](image-url)
n = 1), i is the index of the dilution level, and \(f_i\) is the dilution factor of the \(i\)th dilution \((f_{i-1} = 0; \text{if no dilution, } n = i = 1)\).

With the previous numerical example and eq 2, \(A_1 = 1 + 2 + 0 + 18 = 21\), (Figure 4, continuous line). Compared to formula 1, such a calculation minimizes the error due to the gap.

(2) Formulas 1 and 2 do not allow us to calculate an olfactogram if each dilution level has been smelled by several panelists. In this case, an average should be made from the individual olfactograms at each dilution level, prior to application of the correction factor corresponding to this dilution level, and combination of the different levels in the final olfactogram. Equation 2 is then modified to give the following algorithm:

\[
A_k = \sum_{i=1}^{n} \left( f_i - f_{i-1} \right) \left( \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{A_{i,j,k}}{A_{\text{max}}} \right)
\]  

(3)

where \(A_{i,j,k}\) is the point of the \(j\)th individual olfactogram, of the \(i\)th dilution level, at time \(t_k\), \(A_{\text{max}}\) is the signal intensity when the push-button is pressed, \(m_i\) is the number of panelists at the \(i\)th dilution (normally, \(m_i\) is the same for all dilution levels)

When no gap occurs, and for a single panelist, algorithm 3 yields exactly the same values \(A_k\) as Acree’s formula 1. The \(A_k\) value at the apex of an olfactive peak represents the FD factor of this odorant for a single panelist with eq 1 and the mean FD factor for the whole panel with eq 3. As it also overcomes the two previously mentioned drawbacks, this new algorithm was chosen for writing of the Galaxie software and used for all CHARM calculations throughout this paper. The time required to computerize the raw GC–O data falls from half a day, using an exportation of individual olfactograms as ASCII files and a treatment with a spreadsheet program, to 1–2 min with the present software.

**GC–SNIF.** To generate a GC–SNIF olfactogram, the same algorithm 3 may be used. As such a method only involves one dilution level, then \(n = i = 1\). This program has been used for all GC–SNIF calculations in this study.

**CHARM and Interindividual Perception Differences.** The large difference in sensitivity between individuals is exemplified in Figure 5. Each of panelists CHD, BRA, and SCR specifically showed lower detection thresholds than their other two colleagues to eucalyptol, benzyl acetate, and \(\beta\)-ionone, respectively. The \((Z)\)-3-hexenol peak that was detected by the three panelists also showed important intensity differences (height or area) of the CHARM peak between individuals (Table 2). This clearly proves that ranking the odorant intensity using a single panelist is only valid for this given panelist and may be dramatically different for another panelist. As the GC–O objective is the determination of impact odorants from a scent according to the mean perception of a population, the need for a panel is evident. Due to the huge time consumption for CHARM analyses, even with a multisniffing system, optimization of the panel size was not undertaken in contrast to our GC–SNIF investigation.8

---

**Table 2. Characteristics of the (Z)-3-Hexenol CHARM Peak for Each Panelist (CHD, BRA, SCR)**

<table>
<thead>
<tr>
<th></th>
<th>SCR</th>
<th>CHD</th>
<th>BRA</th>
<th>mean(^b)</th>
<th>RSD, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSD, %</td>
<td></td>
<td></td>
<td></td>
<td>interindividual</td>
<td>interpanel</td>
</tr>
<tr>
<td>height (FD)</td>
<td>63</td>
<td>243</td>
<td>21</td>
<td>110</td>
<td>109</td>
</tr>
<tr>
<td>area (CHARM units)</td>
<td>1110</td>
<td>13719</td>
<td>294</td>
<td>5041</td>
<td>149</td>
</tr>
</tbody>
</table>

\(^a\) Peak heights and areas were calculated prior to averaging. \(^b\) The three individual CHARM traces were averaged prior to calculation of the height and area of the resulting peak (Table 3).
In the course of the CHARM experiments, a gap of coincident responses was observed for the panelists SCR and BRA, in the (Z)-3-hexenol peak of the second repetition. Although they detected this odorant down to dilutions 1/81 and 1/27, respectively, they missed the odor at dilution 1/27 and 1/9, respectively. Whereas application of the original CHARM algorithm 1 would have yielded FD values of 27 and 9, instead of 81 and 27 without the gap, the modified algorithm gave 63 and 21 (Table 2). Therefore, the mean height for the whole panel would have been 93 with the old algorithm and 117 without any gap. Due to the combined use of algorithm 3 and the panel, the value calculated by the new software limits the consequences of these two gaps (no gap for panelist CHD).

The RSD between three panels of three members obviously appears to be lower than the RSD between three individuals (Table 2). This supports the use of the individual response average, where the panel is considered as a single detector, rather than separately looking at individual olfactograms.19

Comparison of CHARM and GC-SNIF. Repeatability. Up to now, mainly qualitative comparisons among dilution, intensity, and detection frequency methods were published. They are based on the parallel determinations using AEDA, OSM E (or FSCM), and GC-SNIF of the impact aroma compounds of Champagne25 and Gewurztraminer wines27 and mussels.24 A recent study dealt with the relationship between the perceived intensities and odorant concentrations, but not with variabilities.28 The evaluation of CHARM and GC-SNIF variabilities was published separately.8,19 Therefore, the results of the GC-SNIF repeatability reported above (Table 1) were compared to those of the CHARM using the same model mixture, the multisniffing system and the modified algorithm 3 (Table 3). Following Acree's conclusions,19 a panel comprising three members was used.

Table 3. Repeatability of CHARM Olfactograms Using the Model Mixture, the Multisniffing System, and the Same Panel (Three Members)

<table>
<thead>
<tr>
<th>ofactogram</th>
<th>height</th>
<th>RSD, %</th>
<th>area</th>
<th>RSD, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>eucalyptol</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st</td>
<td>653</td>
<td>11</td>
<td>46442</td>
<td>2</td>
</tr>
<tr>
<td>2nd</td>
<td>767</td>
<td>56</td>
<td>2569</td>
<td>93</td>
</tr>
<tr>
<td>(Z)-3-hexenol</td>
<td>47.7</td>
<td>55</td>
<td>1456</td>
<td>78</td>
</tr>
<tr>
<td>citronellal</td>
<td>8.8</td>
<td>27</td>
<td>671</td>
<td>9</td>
</tr>
<tr>
<td>benzyl acetate</td>
<td>54</td>
<td>32</td>
<td>3935</td>
<td>2</td>
</tr>
<tr>
<td>1st</td>
<td>86</td>
<td>131</td>
<td>46</td>
<td>3112</td>
</tr>
<tr>
<td>2nd</td>
<td>27</td>
<td>131</td>
<td>46</td>
<td>3112</td>
</tr>
<tr>
<td>mean SD</td>
<td>1</td>
<td>27</td>
<td>46</td>
<td>1915</td>
</tr>
<tr>
<td>mean RSD</td>
<td>58</td>
<td>32</td>
<td>32</td>
<td>30</td>
</tr>
</tbody>
</table>

A low repeatability in regard to individuals was observed (mean RSDs of each panelist: 39, 122, and 103% respectively). This is comparable to the previously published mean RSD (85%) between two repetitions using one panelist individually.19 The use of a three-member panel lowered the mean RSD of peak heights and areas to 30 and 32% respectively. This clearly exemplifies the improvement in repeatability when a panel is used.

In contrast to Acree's results, the data of Table 3 do not exhibit a clear increase of SDs as a function of peak heights or areas. Therefore, the transformation of peak heights into their logarithms does not seem to be required because of the following: (1) From the statistical viewpoint, the logarithm transformation should be justified by an exponential increase of SDs. Even in Acree's example, it is noteworthy that only 3 out of 26 SDs exhibited high values. Without taking them into account, the SD increase with the peak height would be less obvious. (2) The logarithm function should apply to a continuous scale, not to a scale possessing only discrete values such as CHARM peak heights. (3) The use of logarithms is not clearly related to a sensorial justification. (4) RSDs calculated from the logarithm of peak heights are dependent on the peak height unit, whereas they should remain independent.

In both repetitions, the eucalyptol peak was only detected by one of the three panelists in the two highest dilution levels. This means that the repeatability of the corresponding peak was almost only dependent on his own repeatability, because the two highest \((f_1 - f_{i-1})\) factors were applied to his results. Similarly, the \(\beta\)-ionone peak was only perceived by panelist SCR in the second dilution series (dilutions 1, 1/3, 1/9, 1/27, and 1/81). In the first series, due to unknown factors (temporary anosmia, fatigue?), this panelist only perceived the ionone in the nondiluted solution. It explains the corresponding high SD for this compound. However, the use of a panel limits the impact of perception differences, as the individual FD = 81 for this panelist became 27 after averaging over the whole panel. A larger panel would presumably decrease the SD of the ionone peak, but the experimental time would dramatically increase.

Rapidity. The CHARM analysis of the model mixture (Figure 3) required nine dilution levels with a dilution factor of 3 between two consecutive dilutions. Therefore, achieving the sniffing procedure with the multisniffing system took 4 1/2 days for one olfactogram. This would have required 2 1/2 weeks with a single sniff port. Similarly, the GC-SNIF olfactogram was established in less than 1 day instead of 2 1/2 days. Two GC-O sessions per half-day were performed, as different members of the panel participated in each of them, in contrast to the CHARM analysis with three panelists.

Peak Ranking. In both CHARM repetitions, the eucalyptol peak had the highest intensity. As for the repeatability, this fact was mainly due to panelist CHD, who was the only one to perceive this odorant in the two last dilutions. The use of a panel certainly limits the problem of specific anosmias and detection gaps of panelists, because CHD’s contribution only accounts for 1/3 in the final olfactogram. However, FD factors are not representative of the odorant intensity.29 Therefore, ranking olfactometric peaks according to their FD factors remains debatable, despite the improvements proposed in the present work.

The GC-SNIF olfactogram (Figure 3) does not exhibit such problems as peak intensities result from a global evaluation of the whole panel, and NIFs seem to be linearly correlated to the perceived intensities.27 However, either octanal or \(\beta\)-ionone is ranked as the most important odorant of the mixture, depending on whether NIF or SNIF units are chosen (Table 1 and Figure 1). It is unclear whether this phenomenon is due to a chromatographic reason (late-eluting peaks are larger, or they “stick” on the sniff port) or a sensory problem (odor persistence in the olfactory system). This should be further investigated in parallel with a sensory analysis of the model mixture.

CONCLUSION

The use of the multisniffing device dramatically decreases the duration of GC-O experiments. The newly developed data treatment program allows building CHARM and GC-SNIF olfactograms using a single algorithm which minimizes gaps in coincident responses and improves the repeatability of CHARM olfactograms due to the use of a panel. Associated with a panel, CHARM thus remains a valuable technique when the experimental time available is not a critical issue. However, GC-SNIF repeatability is higher, for a shorter analysis time.

ACKNOWLEDGMENT

We gratefully acknowledge Prof. M. Danzart for fruitful discussions on statistics, Mr. J.-P. Probst for having built the multisniffing system, Mrs. S. Rattier (Varian-JMBS SA) for extending the Galaxie chromatography data system with GC-O processing features, and all panelists who participated in the GC-O experiments.

Received for review October 5, 2001. Accepted February 23, 2002.

AC011064A