Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 261 (2021) 120025

Contents lists available at ScienceDirect

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy

journal homepage: www.elsevier.com/locate/saa

The use of an optimized DRIFTS-FTIR method for the forensic analysis and classification of silicone condom lubricants

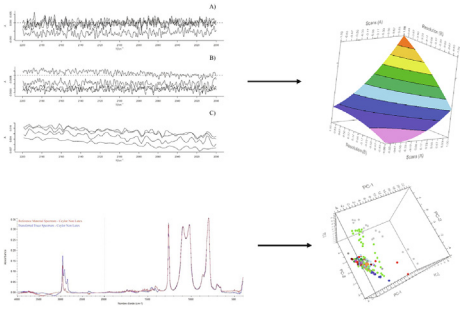
Céline Burnier *, Virginie Favre, Geneviève Massonnet

Ecole des Sciences Criminelles, University of Lausanne, Switzerland

HIGHLIGHTS

- Optimal analytical conditions for DRIFTS-FTIR analysis are 64 scans and 4 cm^{-1}.
- Condom discrimination is not enhanced with DRIFTS compared to ATR-FTIR.
- Traces were found to cluster with the rest of the condom pattern.
- It was not possible to attribute a source to the trace that was recovered.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 11 March 2021
Received in revised form 24 May 2021
Accepted 24 May 2021
Available online 26 May 2021

Keywords:
Chemometrics
Design of experiments
PCA
HCA
Condom traces

ABSTRACT

Condom residues may be encountered in forensic investigations as traces left in sexual assault or rape cases. Considering casework samples analysis, where material from swabs will need to be extracted, Diffuse Reflectance Infrared Fourier Transform Spectroscopy (DRIFTS) was reported as the most relevant method for trace evidence analysis. However, there has been no study to identify which specific parameters were the most suitable for the analysis of silicone-based lubricants, especially in terms of repeatability of the analyses. This study looked at the resolution and number of scans with the aim of optimizing these parameters for polydimethylsiloxane (PDMS) analysis and detection. Experimental parameters were refined while performing a full factorial experimental design (FFD) for the screening and extended to a face centered central composite design (FCCD) for the optimisation. Repeatability of the results was also investigated using principal component analysis (PCA) and hierarchical cluster analysis (HCA) in order to select the most relevant analytical parameters. The optimized DRIFTS parameters were then used to collect data from reference material and from traces after a transfer in a vaginal matrix has occurred. Discrimination models were built with DRIFTS data and compared to pre-existing models built with Attenuated Total Reflectance (ATR)-FTIR data. Condom traces were recovered from volunteers who had sexual intercourse using different types of silicone-lubricated condoms. The corresponding traces characteristics were investigated and analysed.

1. Introduction

Condom trace evidence; including particulates, lubricants and spermicides; has been reported in the forensic investigation of sexual assaults and rape cases for the past 40 years [1]. These traces can be used by forensic investigators to determine whether or not a condom was used during intercourse, potentially supporting the statements of one of the parties.

The majority of condoms consist of latex coated with solid particles and lubricant. The latter is usually polydimethylsiloxane (PDMS), found in more than 85% of condoms on the international

* Corresponding author at: Ecole des Sciences Criminelles, Quartier UNIL-Sorge, Bâtiment Batochime, CH-1015 Lausanne, Switzerland.
E-mail address: celine.burnier@unil.ch (C. Burnier).

https://doi.org/10.1016/j.saa.2021.120025
1386-1425/

market [2–10], or polyethylene glycol (PEG). Solid particles are added to the latex during production to avoid the latex to stick to itself. They are likely to be used as forensic evidence in casework [10,11]. Other additional components such as anaesthetics, colourants or flavourings, can also be found in very small quantities [10]. Condom residues recovered in the investigation of a sexual assault will contain a combination of these compounds along with components of the vaginal matrix. This complexity poses a challenge for detection, which is affected by the initial transfer, including the influence of the source of the trace (i.e. the condom) but also by the vaginal matrix itself and the type of contact (i.e. duration, intensity). Finally, the time elapse between the alleged activity and the sampling as well as the activity of the victim during this period will influence the persistence of the target compounds [8–10,12].

Condom traces are typically sampled using cotton swabs, then extracted into a solvent for analysis [1,12]. Dichloromethane is the most common extraction solvent, although hexane and isooctane have also been reported as dichloromethane may not provide for complete extraction of the condom traces [12]. A screening method is usually performed to determine the classes of compounds present, followed by a confirmatory technique to identify specific components [8–10,12]. Spectroscopy is frequently used for preliminary analysis of condom residues [12–15] due to its rapidity, simplicity of use and availability in forensic laboratories. In particular, Fourier Transform infrared (FTIR) spectroscopy is frequently employed in both research and casework as a screening method, with a preference for diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) when handling caseworks [1,12,15]. DRIFTS is ideal for this purpose as it allows the analysis of polymers such as PDMS, and has been confirmed by several authors as the most relevant method for the analysis of condom residue in solution, based on the spectral quality, the reproducibility, the analysis and preparation time as well as the applicability to trace evidence [1,16].

Although DRIFTS has been established to be suited to the analysis of condom traces [1,12,15], the analytical parameters are multiple and various in the literature. Although this works is for qualitative analysis, analytical conditions should be statistically reproducible and repeatable for an application to casework. Factors influencing the repeatability of DRIFTS analysis have never been reported to the best of our knowledge and there is a need to obtain more reproducible and repeatable results, with a good signal-to-noise ratio (SNR), to be able to work with traces. Development and optimization of the SNR using statistical algorithm, such as design of experiments, is still necessary. Enhancing SNR will ensure that that vibrational bands of the target compounds are distinguishable from the background for qualitative examination. Minimising noise is also mandatory to ensure that detected information can be attributed to the evidence rather than instrumental artefacts, and to improve reproducibility. Forensic sciences attach a strong importance to the identification of the source of the recovered evidence. In the case of condom evidence, inferring on the source means identifying the brand or type of condoms. The most common way to answer such questions is to conduct a statistical exploratory study on a large population set. Such studies have been widely reported with different analytical techniques, and amongst all ATR-FTIR was one of the most important. However, there's no such published available study for DRIFTS analysis, and this can be a serious issue as this is the recommended method for the analysis of condom trace evidence in a forensic context. Therefore, it is necessary to create such a model and compare it with the previously published ATR-FTIR model. In addition, it is of prime importance to see whether real traces are classify the same way pure lubricant samples do.

The present paper aims at first to determine which factors, between the number of scans and the resolution, most significantly affect the SNR, in order to obtain a more adequate understanding of how to analyse silicone-based condom residues. Designs of experiments were used to explore all the possible combinations of parameters variations within the experimental plan [17], with the advantage to allow data collection and analysis with proper statistics to obtain maximal information, such as important factors affecting the targeted response (here, the SNR). In a second part, the optimised parameters were then applied to silicone-lubricated condom extracts and real trace samples, with the aim to build a discrimination model and investigate the classification of real samples in the given model, to observe the differences with pre-existing models and provide relevant information for the forensic scientist, when it comes to casework application.

2. Material and methods

2.1. Material

Hexane (Sigma-Aldrich, Damstadt, Germany) and Methanol (Sigma-Aldrich, Damstadt, Germany) of analytical grade were used as received. Dimethylpolysiloxane 200 cSt was purchased from Sigma-Aldrich (Damstadt, Germany). Cotton swabs (COPAN150C) were purchased from COPAN (Brescia, Italy). Potassium bromide (analytical grade) was purchased from Fluka Chemika (Honeywell International) and was manually grinded before use. eVol XR® syringe from SGE Analytical (Trajan Scientific Australia) was used for spiking the samples on KBr pellets.

2.2. Instrumentation and sample analysis

2.2.1. DRIFTS-FTIR

Infrared **DRIFTS** spectra were acquired with a Digilab FTS 3000 Excalibur FTIR spectrometer (Portmann Instruments AG, Biel, Switzerland), equipped with a Spectra-Tech 0030-05 Collector II diffuse reflectance accessory (Portmann Instruments AG, Biel, Switzerland) and DTGS detector. Spectra were collected from 4000 and 400 cm^{-1}. Potassium bromide (KBr) was manually grinded to obtain a homogenous powder and deposited into metal sample cups for DRIFTS analysis. Manual pressure was applied with a spatula to the pellets to remove residual air, and the pellet batch stored in a 100 °C oven. For analysis, 10 µl of sample in solution were spiked onto a pellet which was then placed in a 100 °C oven for 15 min to evaporate the solvent. Blanks were prepared in the same manner using 10 µl of hexane and analysed every 3 measures to account for background interference. Extraction blanks were also prepared in the same manner using extraction of a clean swab.

2.2.2. ATR-FTIR

Infrared **ATR** spectra were collected using a Nicolet iS50 FTIR spectrometer equipped with single-bounce diamond crystal ATR accessory (ThermoFisher Scientific). Data collection was carried out using the OMNIC software v. 8.2.0.387 (ThermoFisher Scientific). Spectra were collected over the 4000 to 400 cm^{-1} range with 4 cm^{-1} resolution and 32 co-added scans. Samples were deposited or rubbed directly on the crystal as described in [18].

2.3. Optimization of the analytical DRIFTS conditions

2.3.1. Identification of factors influencing the Signal-to-Noise ratio (SNR)

Several experimental designs were conducted in this study, as an iterative process in order to obtain the highest and less variable

SNR. All the designs were realized using a single standard solution of bulk PDMS diluted in hexane (1 mg/ml) and spiked on a KBr pellet. The first experimental cycle used was a two-level FFD (Full Factorial Design) experimental plan, generated using Unscrambler X (Camo Software, Norway) to observe the response surface. The parameters used are described in Table 1. The chosen FFD plan used four replicates of each point including the central point. This resulted in a total of 20 randomized program experiments. The central point was defined at 64 scans and 4 cm^{-1} because it is the closest point to regular practice in forensic laboratories [27].

The second experimental cycle was led to estimate the effects of each factor. A new FFD was designed with new scan number, chosen within a factor 2 from the central point, i.e. 32 and 128 scans. That was decided to reduce the time of analysis as some of the highest level of scan and resolution might significantly increase the latest. The resolution variable was modified, to correspond to a variation of a factor 2 around the value of the central point.

Finally, a third experimental cycle was an extension of the FFD into a Central Composite Design (CCD), more specifically here, a face centered composite design (FCCD). The remaining points were added to the initial FFD to capture the true relation between the factors and the SNR. For each cycle, effect significance, lack of fit, regression significance and curvature were evaluated. A total of 80 experiments were led for these designs.

Data analysis for experimental designs were performed in Unscrambler X v.10.1 (Camo Software, Oslo, Norway) and two-ways ANOVA calculations was used to determine the effects of the factors. For all the models sketched on the data, the significance of the effects, the adjustment of the model (lack-of-fit), the significance of the regression and the curvature of the plans were evaluated. The lack-of-fit was assessed according to a Snedecor's test [19], and the curvature of the plan according to a Student's test [17]. Several regression models of different complexity (from linear to quadratic) were fitted on the data. The model describing the relation between the factors was then selected based on the highest lack-of-fit p-value and the lowest regression significance p-value.

Table 1
Factors and levels used for the identification of the surface response, using a FFD design.

Factor	Level −1	Level 0	Level 1
Scan number	16	64	256
Resolution [cm^{-1}]	8	4	1

Table 2
Factor and levels used for the calculation of the effects, using a FFD design.

Factor	Level −1	Level 0	Level 1
Scan number	32	64	128
Resolution [cm^{-1}]	8	4	2

2.3.2. Repeatability evaluation and parameters selection

Repeatability of the instrument is important, especially when it comes to trace evidence analysis. To evaluate the repeatability, two different statistical methods are usually used: principal component analysis and hierarchical cluster analysis, with distance or correlation measurements. All these techniques will be used to investigate which number of scans provides, for a given resolution, the highest repeatability, with spectra clustered the closest to each other. This will allow the identification of the number of scans which provides the highest repeatability of the data.

2.4. Application to condom samples

In order to evaluate the applicability of the optimized method to real samples, 16 condoms and 2 lubricants from major distributors and manufacturers on the Swiss market were purchased from Swiss supermarkets and pharmacies (Table 3). All the samples were previously categorized as containing a silicone-based lubricant [18] except for Fair Squared Sensitive Dry (P11), which is a non-lubricated condom.

Condom were individually opened and unrolled before being put in a 100 ml glass bottle and covered with 50 ml of hexane. The bottles were then closed and put in the ultrasonic bath for 15 min. Bottles were then stored at −18 °C until analytical. Before analysis, samples were aliquoted and diluted 10 times. 10 µl of the solution were spiked in the quartz tube on the glass wool and the analysis was processed. Three replicate samples were prepared from each condom to probe the composition homogeneity of the sample as well as the variation due to the instrumentation and the sample preparation.

2.4.1. Data pattern recognition – ATR vs DRIFTS analysis

The results of the 16 samples acquired with the 2 different IR methods were analysed using Principal Component Analysis to identify the potential clustering and classification in the data. The hypothesis to corroborate is the following: samples are not clustered according to their analysis type but according to their

Table 3
Samples used for DRIFTS discrimination model. Composition is known based on observations from [18].

N0	Brand	Model	Type	Lubricant	Other Component
P0	Durex	Performa	Latex	silicone	no
P1	Durex	Invisible	Latex	silicone	no
P2	Durex	Natural Feeling	Polyisoprene	silicone	no
P3	Migros	M-Budget	Latex	silicone	no
P4	Migros	Cosano Regular	Latex	silicone	no
P5	RFSU	Profil	Latex	silicone	no
P6	Manix	Contact	Latex	silicone	no
P7	Manix	Skyn Original	Polyisoprene	Silicone	no
P8	Ceylor	Blauband	Latex	Silicone	no
P9	Ceylor	Non-Latex	Polyurethane	Silicone	no
P10	Coop	PrixGarantie	Latex	Silicone	no
P11	Fair Squared	Sensitive Dry	Latex	Non lubricated	Solid Particles
P12	Vitalis	Natural	Latex	Silicone	no
P13	Amor	Nature	Latex	Silicone	no
P14	ESP	Skin	Latex	Silicone	no
L1	Durex	Play play eternal - Perfect glide	Personal lubricant	Silicone	no
L2	Ceylor	Silk sensation	Personal lubricant	silicone	no

sample category. This hypothesis has, to the author's knowledge, never been reported before, these two techniques usually being tested one against the other, and not as the potential of complementary techniques. This investigation aims to verify if samples obtained with DRIFTS analysis can be projected in the market survey discrimination and classification models built with ATR results [18], and if not, to evaluate if the discrimination patterns obtained on the two techniques are similar. This would allow to transpose the discrimination pattern observed on the entire dataset from one technique to the other, although variability could affect the pattern.

2.4.2. Trace vs reference classification

To investigate the classification of real samples, 2 volunteers had sexual intercourse using 10 different condoms, coming from the list presented in Table 1 as well as from other condom types presented in [18]. 3 blank swabs were collected prior to intercourse, and 3 samples swabs were collected right after intercourse. To avoid any cross-contamination, the volunteers were asked to wait one week between each protected intercourse. Each sample was analysed 3 times, as described for the condom samples, resulting in 9 replicates. A total of 132 analysis were run for this purpose. Sample collection was led in accordance with the Swiss Federal Act on Research involving Human Beings (Ordinance on Human Research with the Exception of Clinical Trials) and approved by Swissethics (2018-00690). Informed consent was obtained from all donors, and collected data was entirely anonymized. Based on previous researches [1,12,14], cotton swabs collected from the volunteers were cut from the wooden sticks and individually put in a glass vial and extracted with 1 ml of hexane. The vials were vortexed for 1 min and sonicated for 15 min. The resulting samples were analysed in triplicate.

3. Results and discussion

3.1. Identification of factors influencing SNR

3.1.1. Response surface screening

Analyses carried out on the Full Factorial Design were first visually analysed to evaluate the variability among all the replicates, by looking at the variability around the baseline, and the noise variation. The results showed that both number of scan and resolution affect the SNR, as well as their interaction. The quadratic effects were found to be non-significant. The non-significance of the lack of fit (p-value = 0.805) allowed to assume that all main effects are linear in this model. If the number of scans is increased, SNR increases as well, but so does the analysis time. In the same manner, the lower the resolution, the higher the SNR, but the lower the quality of the spectral information will be. Choosing the adequate analytical parameters is a compromise between the amount of spectral information and the SNR.

The visual comparison of the noise observed in the spectra obtained at 16 scans for the resolution at respectively 1 and 8 cm^{-1} present a good repeatability, in terms of intensity fluctuation. The noise is very important when the resolution increases and reaches 1 cm^{-1}. Similarly, at 256 scans, a very good repeatability is observed, with a noise higher if the resolution is higher. Repeatability is globally higher at 256 scans than at 16 scans for a same resolution (cf. Fig. 1A and B). The scans at 64 scans and 4 cm^{-1} have a rather good repeatability (Fig. 1C). The impact of the resolution is clearly visible compared to the other sets of analysis.

To investigate the variation of the SNR as a function of the number of scans and of the resolution, the noise (RMS) was integrated on the 2200–2000 cm^{-1} region as it is the range where the beam

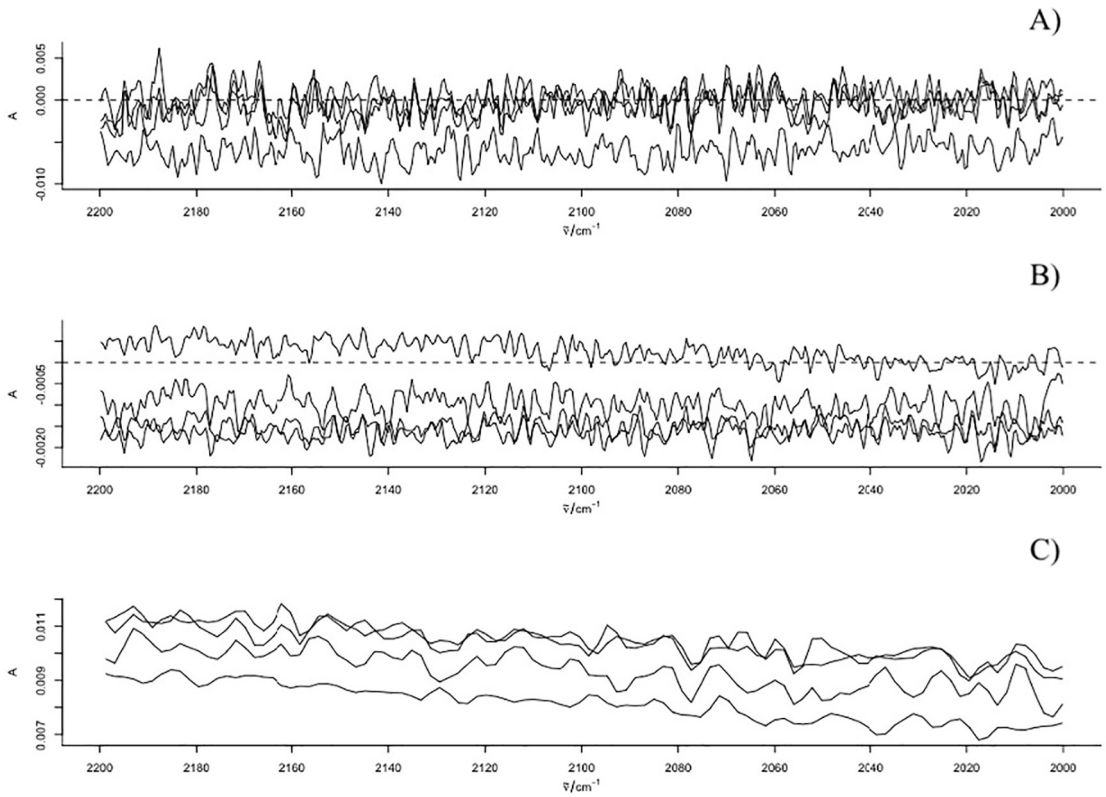


Fig. 1. Illustration of the noise of the spectra acquired under different number of scans and resolution over the 2200–2000 cm^{-1} region, four replicates per design point are presented, A) 16 scans/1cm^{-1}, B) 256 scans/1cm^{-1}, C) 64 scans/4cm^{-1}.

Table 4
Observed coefficient of variation (%) of the SNR as function of the number scans and of the resolution.

Scans	Resolution	SNR	CV %
16	1	283 ± 15	5.60
16	8	883 ± 65	7.43
64	4	1424 ± 91	6.42
256	1	1440 ± 204	14.16
256	8	3662 ± 443	12.09

intensity is the greatest [20]. SNR was calculated as follows: $SNR = \frac{Baseline\ \%\ Transmittance\ Value}{RMS}$. The baseline transmittance value was usually 100% but was sometimes slightly lower (i.e around 99.5%) and therefore was adapted as a function of the value obtained for each spectrum.

Table 4 highlighted the lower variability of the central point (64 scans, 4 cm^{-1}). As illustrated in Table 4, the SNR is lower when a high resolution set up is used than when a lower resolution is used. To this effect, a high resolution at 1 cm^{-1} is not recommended if planning to work on real cases: when dealing with trace evidence, there's a need to have the highest SNR. Similarly, for the number of scans, a too low number of scans provides a lower SNR and therefore is not recommended.

The variability of the SNR was also investigated as it was assumed that the variability detected on the SNR would be observed as well on the data collected from real samples. Variability was found to increase as the number of scans increases (cf. Table 4). Thus, it seems inappropriate to expose the sample to too many scans, since not only the analysis time increases, but so does the variability. The coefficients of variation obtained for the low number of scans are less than 10%, which is rather good considering the mode of analysis and the problems related to diffusion that can be observed. These observations confirm that both number of scan and resolutions are important parameter influencing the reproducibility of the data.

Higher number of scans induce higher variability of the SNR, and the reproducibility is affected as well. Therefore, analyses at 256 scans were discontinued for further investigations. The surface screening showed more repeatable results for a number of scans closer to the central point. At this point, in order to grasp the effects of each variables of interest, a new design of experiments was carried out, focusing the setting values close to this central point.

3.1.2. Calculation of the main effects

The knowledge acquired in the first cycle of experiment allowed reducing the factors closer to the center points. A new two-level factorial design of experiment was run to estimate the effects of each of the factors. Each point was analysed four times to get replicates. Calculation of the main effect of each parameter were realized as described in [17] and respective effects of ~3196 (p-value = 0.0001) for the number of scans and ~−2096 for the resolution (p-value = 0.0008) were obtained. These results show that both effects almost equally influence the response of the SNR. Both factors have positive effects. This means that decreasing the resolution (i.e. 8 cm^{-1}) and increasing the number of scans cause an increase in the SNR. The goal is to increase this ratio, since the signal must be maximized with respect to noise. To achieve the goals set, the number of scans must be maximised, and the resolution minimized. The effect of the interaction has also been calculated and is ~1496 (p-value = 0.0016). The effect here is half as important as the one of the numbers of scan, and almost as important as the one of the resolution, while positively affecting the SNR. This interaction is therefore important for the model because its effect is as important as any of the main effects.

This design was still not sufficient to have a complete coverage and understanding of all the interactions. Thus, an extension to a FCCD design which allows computing more complex interactions and create a final response surface modeling with the best understanding of the impact of each parameter was achieved.

3.1.3. Response surface modeling

FCCD was used to estimate and evaluate first and second order models of regression. The analytical results were used to build a full regression model of the first order, firstly using only the number of scans and the resolution, then considering their interaction

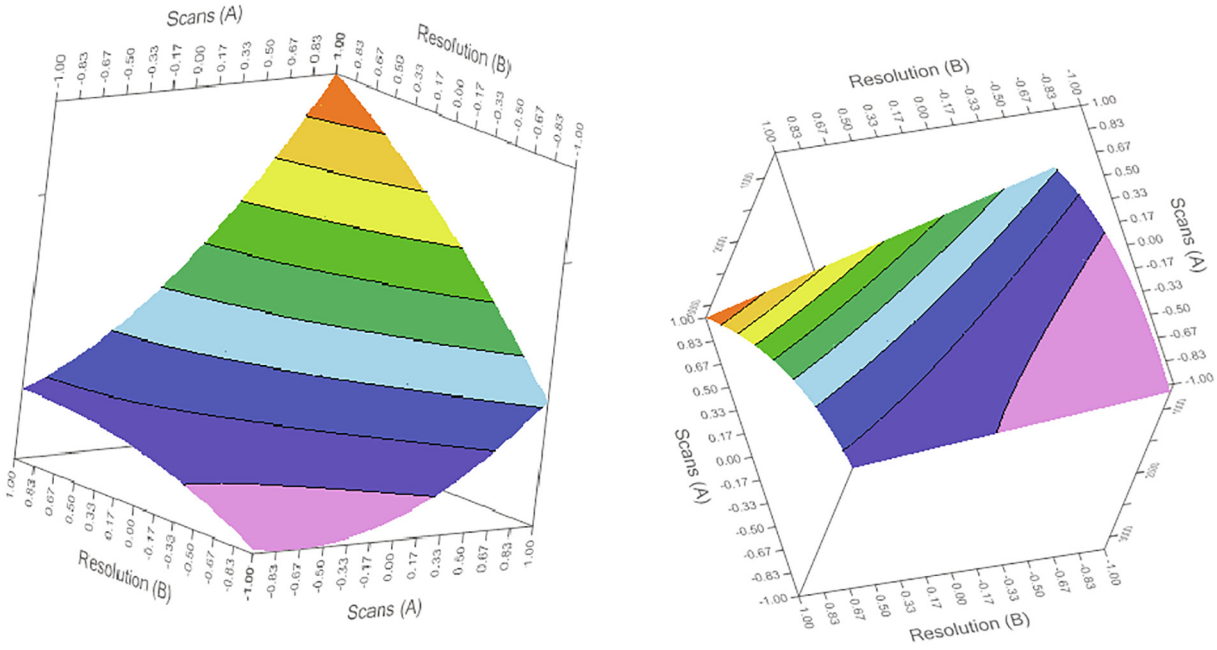


Fig. 2. Surface response obtained for the quadratic model with interaction (left) and for the optimization using only the main effects (right). Factors level are the ones described in Table 2.

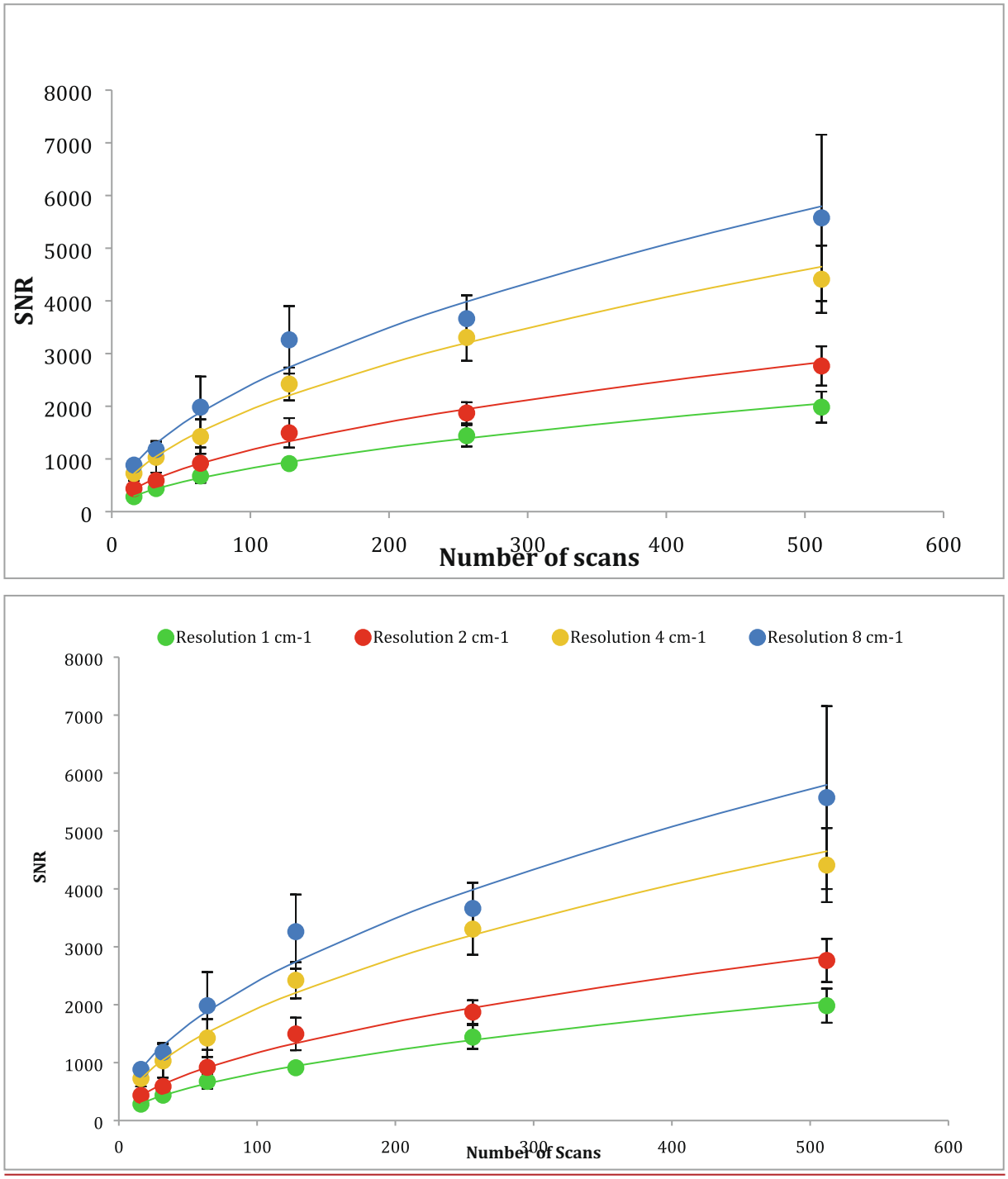


Fig. 3. . Illustration of the relationship between scan number and resolution.

and finally considering second order terms. First order effect were found to be statistically significant, with p-values < 0.001. Within this model, the AB interaction was found to be non-significant, with a p-value of 0.47. Quadratic effects were also investigated but were not found to be significant, with p-values respectively 0.646 for the Scan × Scan (AA) parameter, and 0.388 for the resolution × resolution (BB) parameter. The curvature was not found to be significant, indicating that only a linear model would fit on the data. In addition, the lack of fit was non-significant with a p-value of 0.8051 which indicates the model is not adequate for such a model. The different models were all compared using the adjusted R^2 with a partial Fisher-test. None of the models were found to fit properly. Lack-of-fit values were found to be 0.085 for the first-degree model with interaction, 0.805 for the second-degree model with all the factors and 0.8725 for the second order models without the squared resolution. This indicate that none of these models are likely to fit the data. Surface response obtained for the quadratic interactions and the linear modelling are gathered in Fig. 2.

Multiple regression models were tested and showed out that only the number of scans and the resolution were significant for the surface response. In addition, when removing the quadratic

interactions, the obtained surface response was completely linear, suggesting an increase of the SNR with the increase of the number of scans and the decrease of the resolution. To be able to select adequate analysis parameters, focus was set up on literature reporting that the SNR increases according to an exponential curve of the type $y = ax^b$, the exponential of which is close to a theoretical value of 0.5 [20]. Therefore, optimal analysis parameters will be selected based 1) on the exponential curve for the optimal resolution number, and 2) on the shorter distance or correlation between spectra acquired for a same resolution and different number of scans.

It also needs to be considered that these second and third design cover a very limited area of the global S/N curves. Therefore, it might appear that, on this small area, the global trends are not always respected. However, one of the important parameters to consider is the variability on the sample, and to this extent, there is a need to evaluate the repeatability of the analysis in order to select the most relevant parameters for the forensic analysis of condom evidence.

3.2. Repeatability evaluation and parameters selection

3.2.1. Resolution

As illustrated in [20], optimal resolution can be selected when plotting the resolution as a function of the number of scans, fitting a power function of the type $y = ax^b$, the exponent of which is close to a theoretical value of 0.5 [20]. The optimal analysis parameters are therefore those whose exponential is closest to 0.5, with the regression coefficient the closest to 1.

Relation between resolution and number of scans will be plotted and the parameters of the curves were calculated, as well as the regression coefficients of the latest. The relation between resolution and number of scans is illustrated on Fig. 3. However, it has to be considered that a lower resolution won't allow the optimized separation of the infrared signal. In the practice, a resolution of 4 cm^{-1} usually offers the best compromise between the S/N value and spectral separation.

Parameters of the curves were calculated, as well as the regression coefficients of the latest (Table 5). Although 8 cm^{-1} resolution is the one which offers the highest SNR, both the regression coefficient and the power function parameter *b* are more fitting to the power function when using a resolution of 4 cm^{-1}. As it is also a very common parameter in most forensic laboratories, the resolution of 4 was selected as a final parameter for further analyses.

3.2.2. Number of scans

To select the adequate number of scans, the repeatability of the analysis should be assessed. Therefore, the data were first plotted into a principal component analysis to evaluate the variability of the different measurements. In addition, hierarchical cluster analyses were used with Ward's Linkage, Euclidean Distance measurement and Pearson's correlation measurements, and different linkage were tested, to see whether the results were consistent.

As illustrated on the PCA results on Fig. 4, the spectral data are rather spread out and are not really clustered together. The only ones clustered together are the ones acquired at 64 scans, as represented with a black circle on Fig. 4.

Table 5
Resolution impact.

Resolution cm^{-1}	a	b	R^2
1	88.951	0.4692	0.9334
2	141.65	0.4503	0.899
4	246.51	0.4897	0.9958
8	319.4	0.4253	0.8194

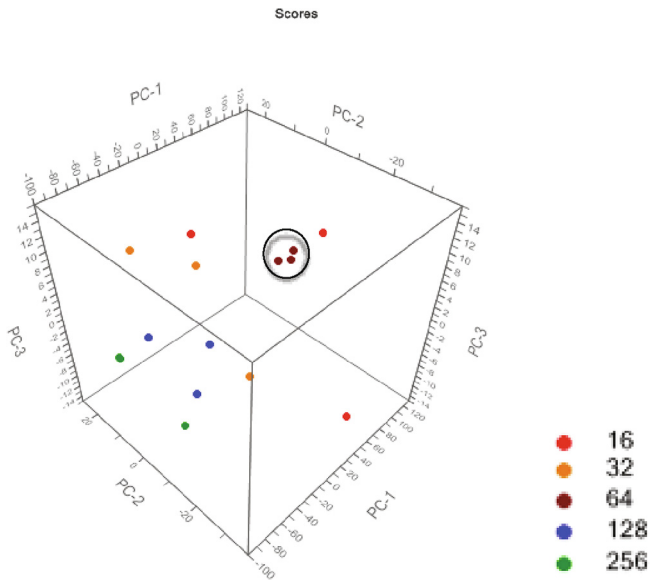


Fig. 4. 3-dimensional scores plot of the data acquired on different number of scans but a same resolution. Samples are coloured as a function of the number of scans.

All the hierarchical clusters provided the same results, with the smallest distance or correlation between the spectra being observed at 64 scans independently of the measurement methods or linkage method, thus suggesting a more appropriate repeatability of the data. An example of the results obtained with a Pearson's correlation and a complete linkage analysis is presented on Fig. 5. The black circle highlights the measurements obtained for 64 scans. The grey circle outline measurements obtained for 256 scans: interestingly, 2 replicates acquired at 256 scans are very close to each other, as close as 64 scans ones together, but the third replicate systematically is more distant to the rest of the samples. 256 scans would be a very interesting option of analysis as some authors have used a larger number of scans (256 or 512) to deconvolute the spectra and classify samples of different types. Increasing the number of scans to 512 would not significantly improve the SNR but increase analysis time by a factor of 10: analysis were found to last around 3 min for 64 scans and up to 15 min for 256 scans. Hence, this is not a cost-effective compromise based on spectral variability, but it might be requested for specific aims (i.e., spectral deconvolution).

The results presented above show that the more relevant conditions regarding the number of scans and the resolution are **64** and **4 cm^{-1}**, respectively. These are the parameters that will be used for further analysis.

3.3. Application to condom samples

3.3.1. Comparison of DRIFTS and ATR discrimination models

Most of the previous classification and discrimination models build for condom evidence analysis were constructed using ATR-FTIR. However, DRIFTS has been reported as the most adequate analytical method for the analysis of condom residues when it comes to casework, as described by [15]. This may be an issue for a forensic scientist facing condom evidence in case work, especially if the main concern is the proper identification of a condom, its brand or a specific model, as previous classification and discrimination models were constructed using ATR-FTIR, which is the reason why a model dedicated to DRIFTS analysis should be considered.

It is therefore of great interest to observe if the clustering identified in ATR data is also highlighted in DRIFTS data. Considering

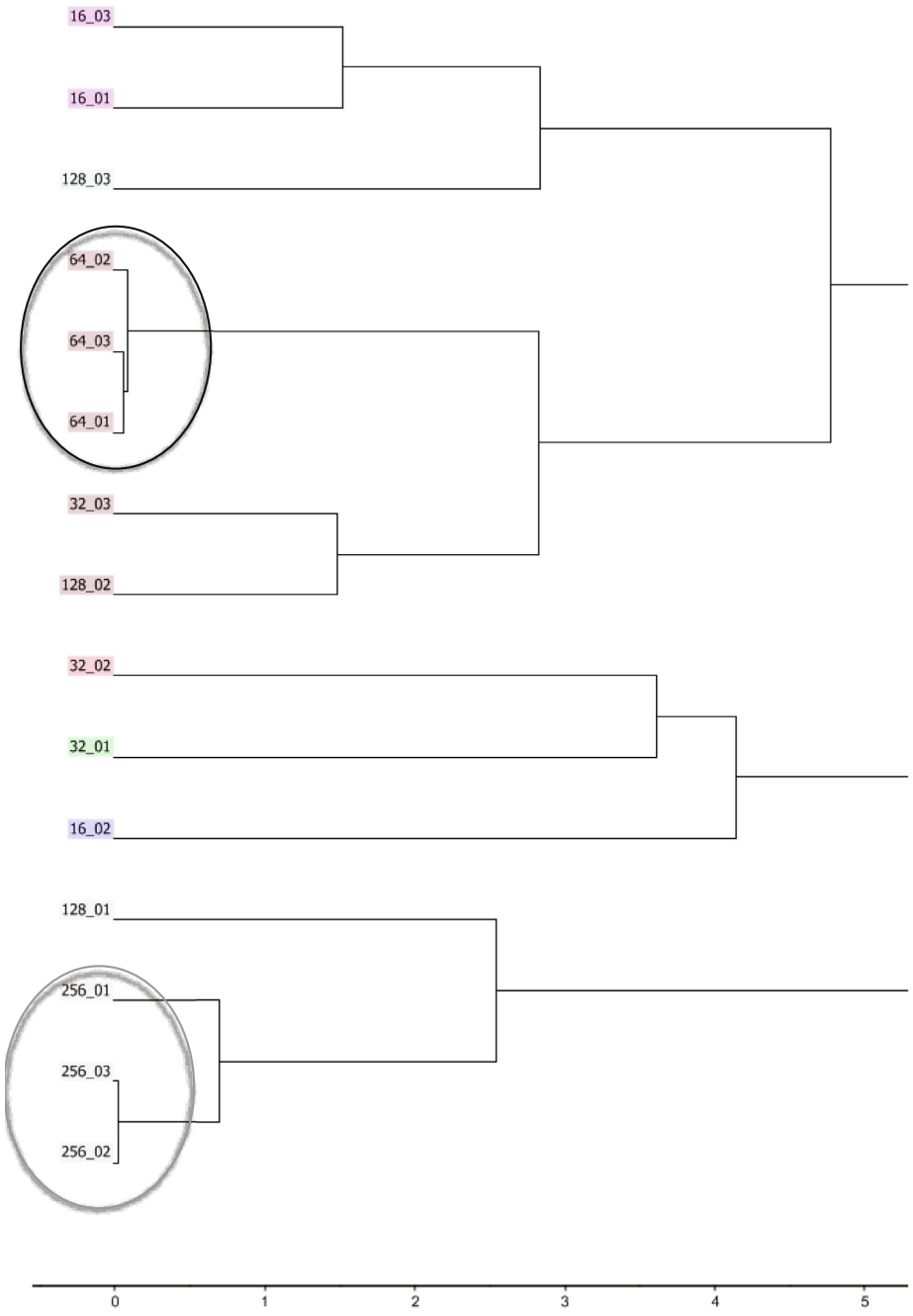


Fig. 5. Hierarchical Cluster Analysis obtained with a Pearson's correlation measurement, and a complete linkage clustering.

the chemical profiles obtained on a batch of different samples using DRIFTS, the main difference between ATR and DRIFTS is based on the silicone-based samples and only on the 1100–1000 cm^{-1} region of the spectra. Indeed, it is where the silicon-oxygen double bond presents its symmetric and asymmetric stretching vibrations, which are better resolved and show higher intensity in DRIFTS than in ATR, as previously illustrated in [15]. There seem to be an opportunity here to possibly differentiate using DRIFTS samples that could not have been distinguished with ATR-FTIR.

At first, PCA was computed on the acquired data sets separately. For the DRIFTS analysis data, the variance explained by PC1 to PC3 was respectively 64%, 15%, and 7%. The first three principal components explain 86% of the variance whereas for the ATR data, the variance explained along PC1 to PC3 were respectively 84%, 8%, and 6%, leading to 98% of variance explained. Fig. 6 shows the scores of the first three principal components (PC1-PC3) plotted against each other for all data. In this figure, replicates of condom of same *model* are represented by the same color. Fig. 6 illustrates the comparison of the PCA scores plot obtained for the ATR (left) and DRIFTS (right) data. Previous research [18] has highlighted specific data patterns using ATR-FTIR analysis. A similar structure in the data can be observed, with samples L2 and sample P11 being separated from the rest of the sample set. Similarly to previous

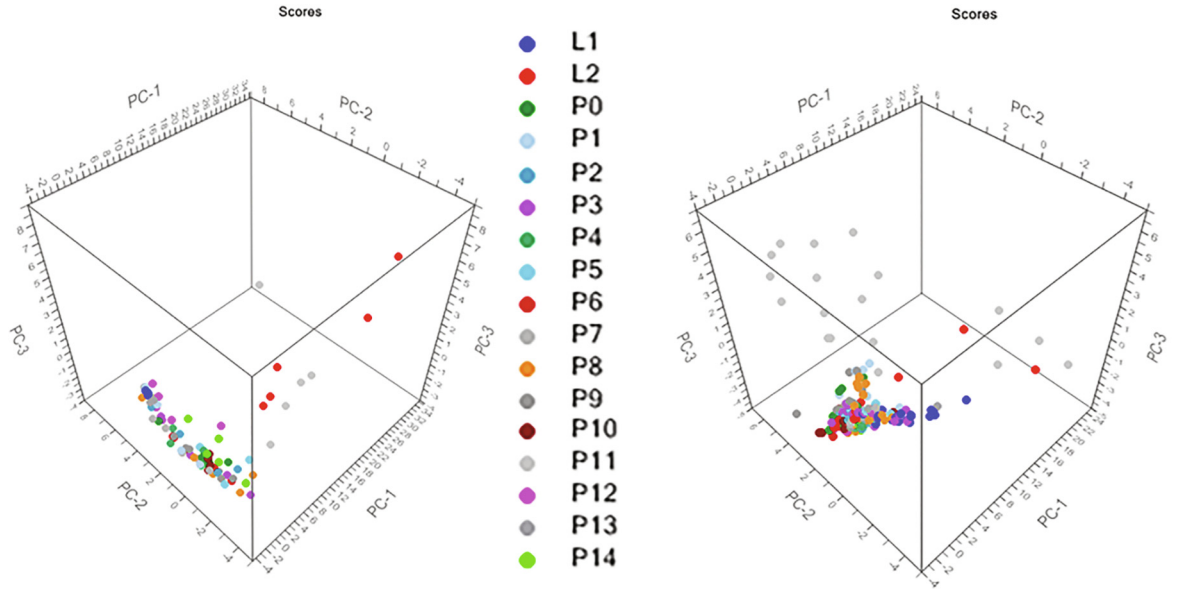


Fig. 6. 3-dimensional PCA scores plots showing the distribution of the samples acquired with ATR-FTIR (on the left) and DRIFTS-FTIR (on the right) for comparison. Legend numbers refers to Table 1.

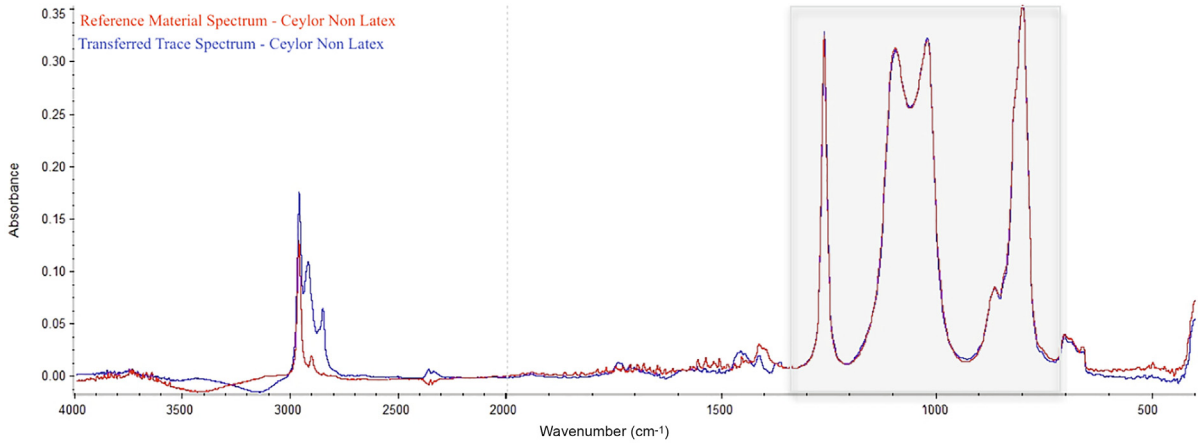


Fig. 7. Illustration of the overlay between a condom reference material spectrum (in red) and a trace evidence spectrum (in blue). The diagnostic peaks of PDMS are highlighted in grey, and are 1263, 1090, 1020 and 900 cm^{-1}. The overlay does not illustrate any significant variations in the area of PDMS, the four bands being present (grey zone), and a similar abundance being observed.

results published on ATR-FTIR results [18], the rest of the silicone containing samples cannot be differentiated of each other, and it is not possible to differentiate samples coming from different brands or models. Further PC did not help enhancing the discrimination of the samples. This suggests that although DRIFTS offer better visualisation of the siloxanes chemical patterns, it does not enable any improved discrimination of samples.

A very important difference that can be noted when comparing the two models presented on Fig. 6 is the dispersion of the data: DRIFTS model presents a far more important dispersion of the data than ATR model. This can easily be explained by two main reasons: the sample preparation and the extraction procedure. Regarding the sample preparation, DRIFTS analysis required the manual preparation of KBr pellets followed by spiking the liquid sample. Not only the KBr grinding might not always be reproducible between all the pellets, but an important modification of the KBr surface is generated at the moment of spiking, which is a source of variation.

Secondly, the samples analysed using DRIFTS were samples extracted from the condom whereas the ATR data were acquired by rubbing the sample on the crystal. Not only will there be variation resulting from the transfer, but also the extraction efficiency has not been evaluated, as it is known from [12] only that the use of a solid-liquid extraction using hexane allows adequate extraction of the target compounds. ATR was not found to offer significant better results for the analysis of condom evidence on swabs. *In situ* analysis was shown to be applicable on reference material (i.e. when condom was rubbed on a cotton before squeezing it on the ATR crystal) but produced no results from real samples, when an external matrix (i.e. vaginal matrix) was present. *In situ* analyses present considerable risk of contamination if DNA extraction has not been previously processed. Previous researchers reported that DNA extraction did not affect silicone-based residues analysis using hexane extraction on the same cotton swab [8]. The use of extracted samples is not adequate for ATR analysis, as at the moment of the deposition, the solvent evaporates, generating ran-

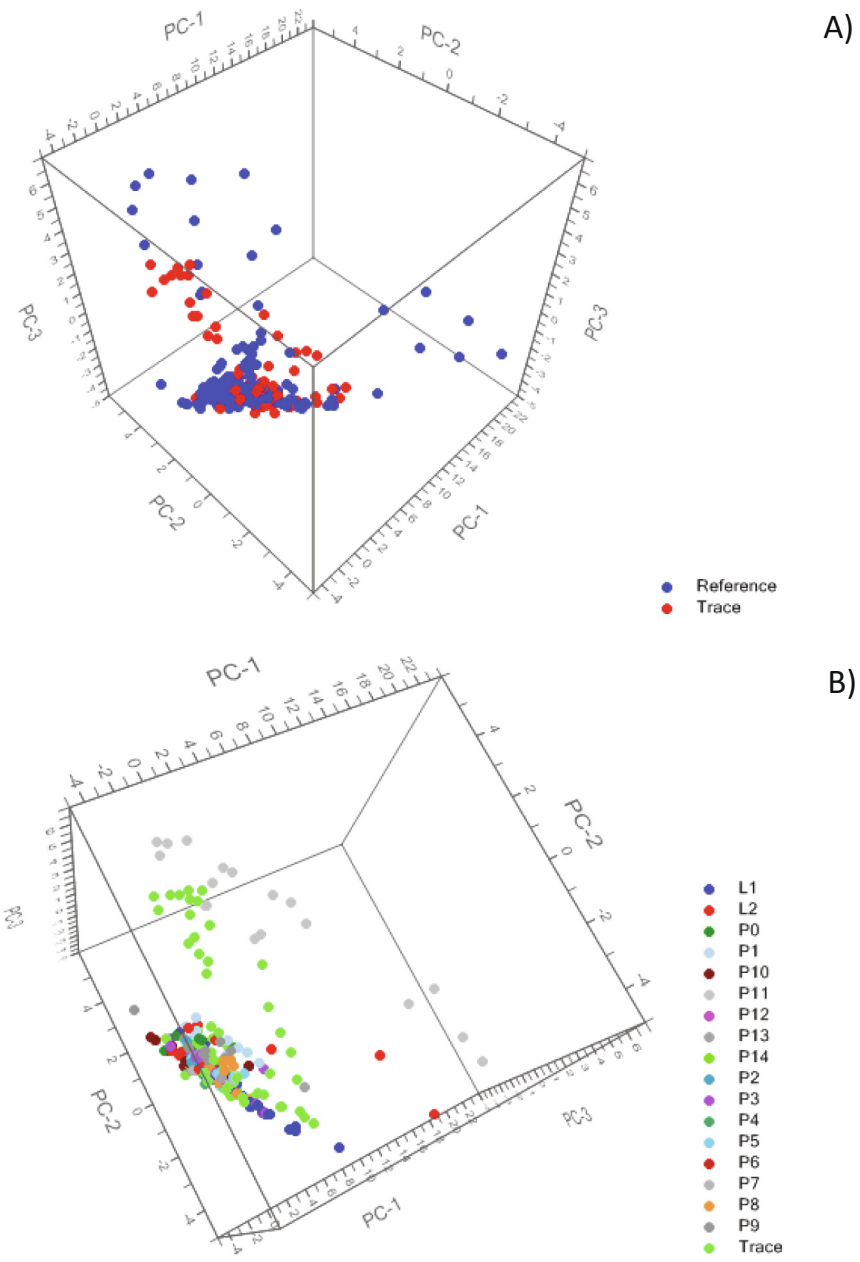


Fig. 8. . 3D-scores plot of the PCA on 16 reference samples (n = 5) and 11 samples collected from human volunteer (n = 8) analysed with DRIFTS, A) distinguishing traces and references, B) distinguishing all the different types of condom used as reference material.

domly deposited PDMS aggregates on the ATR crystal, and thus causing reproducibility issues. On the other hand, DRIFTS was reported as the most adequate technique for the analysis of real samples in case works [1] and a very sensitive technique. In this case, an extraction step is mandatory.

Although the number of samples is limited in this study, the analysis of such a big dataset as presented in [18] would unlikely provide more accurate results, considering that condom production is regulated by international ISO norms [21–24]. The conclusions of a discrimination and clustering investigation of a bigger dataset would be unlikely to provide different results on the discrimination of condoms than the ones observed in [18]. Should further discrimination be required, other analytical instrumentations such as DART-TOF-MS would be recommended given their discrimination ability [5,25].

3.3.2. Trace vs reference classification

Another recurrent and important question is to know whether real traces classify the same way pure lubricant samples do [18]. To this aim, samples collected from real samples were analysed. The visual analysis did not allow to differentiate samples coming from traces from the reference, except for the ones coming from sample P11, which is not a silicone-based condom. An illustration of a reference spectrum obtained from the condom Ceylor Non-Latex (P9), compared to the trace spectrum obtained after a sexual intercourse is presented in Fig. 7. The four diagnostic bands of PDMS are present in the spectra (grey zone in the spectra). The overlay do not highlight any significant visual differences, except in the area around 3000 cm^{-1} where more CH_2-CH_3 vibrations can be observed, due to the presence of the vaginal matrix [26].

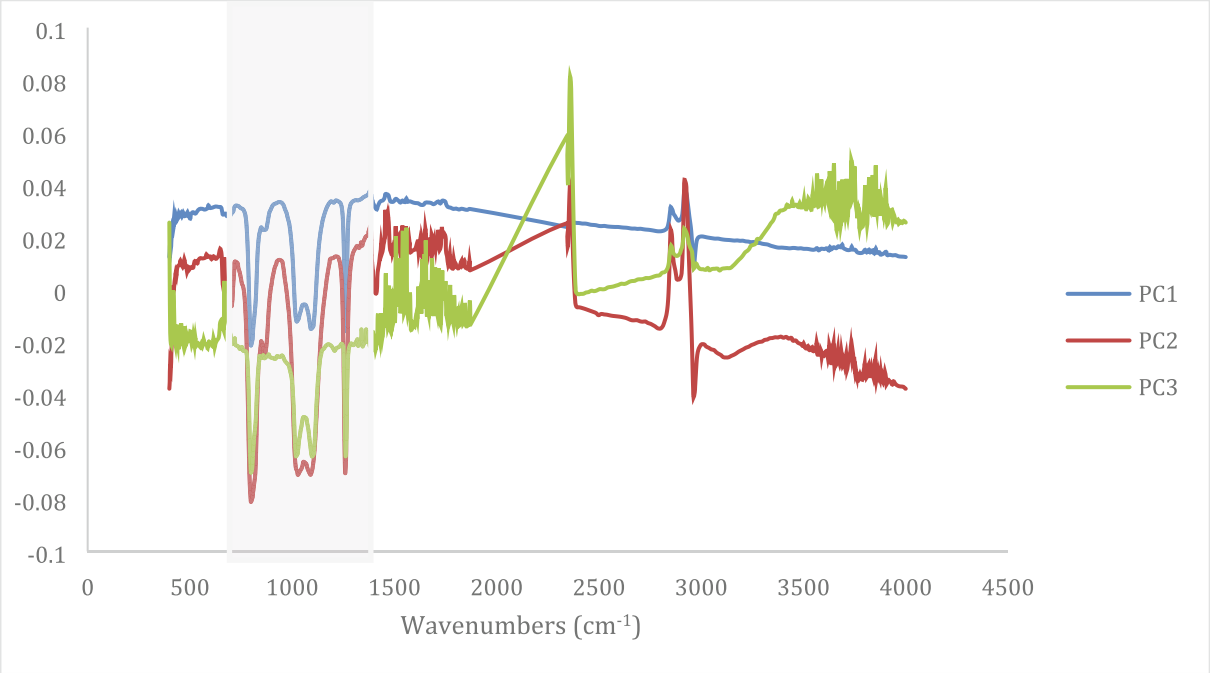


Fig. 9. Factor loadings of PCs 1–3 for the PCA conducted in Fig. 8, based on their DRIFTS-FTIR spectra.

The real samples were then processed and added to the PCA plots to investigate their possible discrimination pattern within the reference material (Fig. 8). As illustrated in Fig. 8A, trace samples are clustered within the same pattern as the one obtained with reference material. Traces present a slightly higher variability than the reference material when it comes to the silicone content (Fig. 8B). Replicates of the same samples were clustered together. Most of the traces were clustered very close to the reference material, thus suggesting that the model would be appropriate to classify traces at the moment of their transfer and that the chemical profiles are not affected by matrix residues which is not surprising considering the type of extraction processed on the samples. Indeed, non-lubricated condom traces were found to be clustered in the same zone (in terms of PC eigenvalues) than the non-lubricated reference P11. In a similar manner, silicone-containing traces were found to cluster within the silicone-containing reference cluster. These observations are very interesting, as they suggest that the chemical profile gathered from condom lubricants after intercourse is not affected at T0 by the receiver or the contact to generate a different pattern or a new specific cluster in the PCA plots.

Samples were clustered together along all the PC, and no clusters were observed differently than from the ones previously observed from the visual inspection of the data. The factor loadings for PC1-3 (Fig. 9) were used to identify the spectral region generating the sample discrimination. Along all the principal components, a strong negative correlation was observed with the peaks at 790, 1020, 1090 and 1260 cm^{-1}, which are linked to the PDMS silicone backbone [15]. A positive correlation was observed with peaks at 2912 and 2925 cm^{-1} consistent with the bonds coming from the vaginal matrix [26], and this discrimination was found to be stronger along PC2, although all PC offer this discrimination. Consequently, the discrimination between the samples is mainly due to the silicone content, most likely based on the concentration content.

In addition to PCA, HCA was also performed, as an unsupervised method to see the similarity between the data, without taking into account any other information. All the traces and the references were used, and 16 clusters were asked. All of the different distance measurements and linkage measurement revealed the same pattern: P11, L2 and the traces coming from these references were clustered together and presented a higher distance to the rest of the samples. Considering the silicone samples, the cluster analysis failed to correctly group the traces with their corresponding condom, independently of the type of linkage or distance/correlation measurements. These results also illustrate that whilst traces can be clustered with the reference material, it is not possible to link a trace to a specific material. This informs that when a trace is recovered, inferring on its exact source might not be possible, although in most casework it might not even be relevant, as the questions usually target the presence or the absence of traces. Whether it is a condom, or another type of sample can also be answered, but the exact source of the trace cannot be inferred, as the chemical profiles are all undistinguishable. This is interesting from the interpretative point of view, as the method is not able to discriminate condoms coming from different brands and models. Interpreting the evidence on a source level would be more relevant in the discrimination from condom vs lubricants rather than between condoms. When considering the use of a given condom vs the use of another one, the likelihood ratio (LR) obtained would be of 1, and therefore uninformative to the Court.

4. Conclusion

The choice of analytical parameters for infrared analysis was assessed using experimental design, as literature was not providing any consistent parameters. Full Factorial design was used and extended to Central Composite Design, using SNR as the target response factor. Distance and correlation measurement were also used to select the more adequate conditions in relation to the repeatability of the data. Most adequate conditions parameters in terms of variability and statistical representation of the data were found to be 64 scans and a resolution of 4 cm^{-1}.

The discrimination pattern of the reference material obtained using DRIFTS was compared to the ones previously published using ATR-FTIR. Similar patterns were observed, with a slightly higher dispersion for DRIFTS analysis, which is attributed to the extraction

and sample preparation step. Data acquired with both models could not be computed together, as the chemical profiles obtained were too different. The discrimination of different brands or model types of condom was similar to the one observed with ATR-FTIR.

Reference and traces were found to be undifferentiated, independently from the condom at the source of the trace. The results indicate that reference condom material and transferred traces do not present distinct chemical profiles. As expected, no differentiation was observed according to the transfer effect itself. The experiments described in this study illustrate the potential to distinguish condoms in a similar way to what was observed during the construction of the infrared profile database.

CRediT authorship contribution statement

Céline Burnier: Writing - original draft, Conceptualization, Methodology, Validation, Investigation, Data curation, Visualization. **Virginie Favre:** Methodology, Investigation, Resources, Data curation. **Geneviève Massonnet:** Writing - reviewing & editing, Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R.D. Blackledge, M. Vincenti, Identification of polydimethylsiloxane lubricant traces from latex condoms in cases of sexual assault, J. Forensic Sci. Soc. 34 (1994) 245–256.

[2] M. Maric, L. Harvey, M. Tomcsak, A. Solano, C. Bridge, Chemical discrimination of lubricant marketing types using direct analysis in real time time-of-flight mass spectrometry: discrimination of lubricant types using DART-TOFMS, Rapid Commun. Mass Spectrom. 31 (2017) 1014–1022, https://doi.org/10.1002/rcm.7876.

[3] M. Maric, C. Bridge, Characterizing and classifying water-based lubricants using direct analysis in real time®□time of flight mass spectrometry, Forensic Science International. 266 (2016) 73–79. https://doi.org/10.1016/j.forsciint.2016.04.036.

[4] Y. Moustafa, C.M. Bridge, Distinguishing sexual lubricants from personal hygiene products for sexual assault cases, Forensic Chem. 5 (2017) 58–71, https://doi.org/10.1016/j.forc.2017.06.004.

[5] A.M. Coon, S. Beyramysoltan, R.A. Musah, A chemometric strategy for forensic analysis of condom residues: identification and marker profiling of condom brands from direct analysis in real time-high resolution mass spectrometric chemical signatures, Talanta 194 (2019) 563–575, https://doi.org/10.1016/j.talanta.2018.09.101.

[6] R.A. Musah, A.L. Vuong, C. Henck, J.R.E. Shepard, Detection of the Spermicide Nonoxynol-9 Via GC-MS, J. Am. Soc. Mass Spectrom. 23 (2012) 996–999, https://doi.org/10.1007/s13361-012-0353-7.

[7] R.A. Musah, R.B. Cody, A.J. Dane, A.L. Vuong, J.R.E. Shepard, Direct analysis in real time mass spectrometry for analysis of sexual assault evidence: DART-MS for analysis of sexual assault evidence, Rapid Commun. Mass Spectrometry 26 (2012) 1039–1046, https://doi.org/10.1002/rcm.6198.

[8] L.S. Tottey, S.A. Coulson, G.E. Wevers, L. Fabian, H. McClelland, M. Dustin, Persistence of polydimethylsiloxane condom lubricants, J. Forensic Sci. 64 (2019) 207–217, https://doi.org/10.1111/1556-4029.13816.

[9] G.P. Campbell, A.L. Gordon, Analysis of condom lubricants for forensic casework, J. Forensic Sci. 52 (2007) 630–642, https://doi.org/10.1111/j.1556-4029.2007.00411.x.

[10] C. Burnier, G. Massonnet, Pre-analytical considerations of condom traces: a review of composition, background, transfer and persistence, Forensic Sci. Int. 302 (2019) 109861, https://doi.org/10.1016/j.forsciint.2019.06.019.

[11] C. Burnier, G. Massonnet, Forensic analysis of condom traces: chemical considerations and review of the literature, Forensic Sci. Int. 310 (2020) 110255, https://doi.org/10.1016/j.forsciint.2020.110255.

[12] P. Maynard, K. Allwell, C. Roux, M. Dawson, D. Royds, A protocol for the forensic analysis of condom and personal lubricants found in sexual assault cases, Forensic Sci. Int. 124 (2001) 140–156.

[13] L.-L. Cho, K.-B. Huang, Identification of condom Lubricants by FT-IR Spectroscopy, Forensic Sci. Int. 11 (2012) 33–40.

[14] R.D. Blackledge, Viscosity comparisons of polydimethylsiloxane lubricants latex condom brands via fourier self-deconvolution of their FT-IR Spectra, J. Forensic Sci. 40 (1995) 467–469.

[15] C.A. Burnier, W. van Bronswijk, G. Massonnet, Comparison of spectroscopic methods in detection of silicone-based condom lubricants evidence, Anal. Methods. (2020) 10.1039.C9AY02471A. https://doi.org/10.1039/C9AY02471A.

[16] R.D. Blackledge, ed., Forensic analysis on the cutting edge: new methods for trace evidence analysis, Wiley ; John Wiley [distributor], Hoboken, N.J. : Chichester, 2007.

[17] G.E.P. Box, J.S. Hunter, W.G. Hunter, Statistics for experimenters: design, innovation, and discovery, second ed., Wiley-Interscience, Hoboken, N.J., 2005.

[18] C. Burnier, S. Coulson, G. Massonnet, K. Pitts, G. Sauzier, S.W. Lewis, A forensic inernational market survey of condom lubricants and personal hygiene products using ATR-FTIR coupled to chemometrics, Sci. Justice (2020).

[19] S.L.C. Ferreira, R.E. Bruns, E.G.P. da Silva, W.N.L. dos Santos, C.M. Quintella, J.M. David, J.B. de Andrade, M.C. Breitkreitz, I.C.S.F. Jardim, B.B. Neto, Statistical designs and response surface techniques for the optimization of chromatographic systems, J. Chromatogr. A 1158 (2007) 2–14, https://doi.org/10.1016/j.chroma.2007.03.051.

[20] J.P. Blitz, D.G. Klarup, Signal-to-noise ratio, signal processing, and spectral information in the instrumental analysis laboratory, J. Chem. Educ. 79 (2002) 1358, https://doi.org/10.1021/ed079p1358.

[21] International Organization for Standardization, Condoms – Determination of nitrosamines migrating from natural rubber latex condoms, 2010.

[22] International Organization for Standardization, Natural rubber latex male condoms – Requirements and test methods, 2015.

[23] International Organization for Standardization, Préservatifs en caoutchouc – Directives sur l'utilisation de l'ISO 4074 dans le management de la qualité des préservatifs en latex de caoutchouc naturel, 2005.

[24] International Organization for Standardization, Lubrifiants additionnels et péparations médicamenteuses ou non, destinés à ou susceptibles d'être mis en contact avec des préservatifs masculins, 2007.

[25] B. Baumgarten, M. Marić, L. Harvey, C.M. Bridge, Preliminary classification scheme of silicone based lubricants using DART-TOFMS, Forensic Chem. 8 (2018) 28–39, https://doi.org/10.1016/j.forc.2017.12.005.

[26] A. Hermelin, L. Fabien, J. Fischer, N. Saric, G. Massonnet, C. Burnier, Analysis of condom evidence in forensic science: Background survey of the human vaginal matrix using DRIFTS and pyrolysis-GC/MS, Forensic Sci. Int. 321 (2021) 110724, https://doi.org/10.1016/j.forsciint.2021.110724.

[27] D. Lambert, G. Massonnet, G. Langer, Project UNIL-BKA Towards Eucap Data Harmonization, presented at the European Paint Group Meeting, Praha, 2019.