

Titre: Optimisation d'un système portable de spectroscopie Raman et développement de protocoles expérimentaux pour la détection de biomarqueurs viraux dans les biofluides
Title: Optimisation d'un système portable de spectroscopie Raman et développement de protocoles expérimentaux pour la détection de biomarqueurs viraux dans les biofluides

Auteur: Antoine Noe
Author: Antoine Noe

Date: 2024

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Noe, A. (2024). Optimisation d'un système portable de spectroscopie Raman et développement de protocoles expérimentaux pour la détection de biomarqueurs viraux dans les biofluides [Mémoire de maîtrise, Polytechnique Montréal].
Citation: Noe, A. (2024). Optimisation d'un système portable de spectroscopie Raman et développement de protocoles expérimentaux pour la détection de biomarqueurs viraux dans les biofluides [Mémoire de maîtrise, Polytechnique Montréal].
PolyPublie. <https://publications.polymtl.ca/61907/>

Document en libre accès dans PolyPublie

Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/61907/>
PolyPublie URL:

Directeurs de recherche: Frédéric Leblond, & Mathieu Dehaes
Advisors:

Programme: Génie physique
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**Optimisation d'un système portable de spectroscopie Raman et
développement de protocoles expérimentaux pour la détection de
biomarqueurs viraux dans les biofluides**

ANTOINE NOE

Département de génie physique

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

Génie physique

Décembre 2024

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

Optimisation d'un système portable de spectroscopie Raman et développement de protocoles expérimentaux pour la détection de biomarqueurs viraux dans les biofluides

présenté par **Antoine NOE**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

Caroline BOUDOUX, présidente

Frédéric LEBLOND, membre et directeur de recherche

Mathieu DEHAES, membre et codirecteur de recherche

Ludvik MARTINU, membre

DÉDICACE

À mes frères, mes parents, ma famille...

À ma chérie, mes amis.

REMERCIEMENTS

Je tiens tout d'abord à remercier mon directeur de recherche, Frédéric Leblond, de m'avoir offert l'opportunité de réaliser ce projet de maîtrise dans son laboratoire. Son expertise du domaine et sa grande expérience m'ont guidé et inspiré pour mener à bien ce projet.

J'aimerais remercier ensuite Katherine Ember, responsable de l'axe biofluides du laboratoire. Sa bienveillance, son écoute attentive et sa disponibilité ont été d'une aide précieuse pour faire progresser le projet.

Je remercie ensuite toute l'équipe du LUMEDLAB, pour leur bonne humeur et leurs précieux conseils tout au long de ma maîtrise, qui m'ont permis de progresser et d'évoluer dans ce monde de la recherche scientifique.

Enfin, je remercie toutes ces personnes qui ont participé à ma recherche et m'ont permis de grandir, d'une manière ou d'une autre, au cours de ces deux années.

RÉSUMÉ

La spectroscopie Raman est une technique prometteuse pour identifier des biomarqueurs spécifiques de maladies dans les biofluides, sans nécessiter de préparation complexe des échantillons. Dans ce contexte, le LUMEDLAB a développé un système de spectroscopie portable, rapide et précis ainsi que des modèles d'apprentissage automatique, ayant permis de valider ce résultat en détectant des biomarqueurs du COVID-19 dans la salive.

Tout en continuant à améliorer ce système, l'objectif est de déterminer des biomarqueurs caractéristiques de maladies virales (COVID-19, RSV, Influenza) dans des échantillons sanguins. Mon travail a consisté à améliorer le système de spectroscopie pour le rendre mieux adapté aux besoins de diagnostic clinique, les développements auxquels j'ai contribué incluent la miniaturisation de composants, une approche innovante pour le contenant de l'échantillon, fondée sur l'élaboration d'un puits formé par une cassette creuse recouverte d'une feuille d'aluminium, et une nouvelle étape de calibration pour assurer la qualité du signal Raman. Ces améliorations techniques ont permis de maintenir la robustesse et la fiabilité du système tout en garantissant une haute résolution spectrale, essentielle pour l'identification de biomarqueurs moléculaires subtils.

Afin de tester ce dispositif et valider le protocole expérimental, une étude a été menée sur un ensemble de 284 échantillons issus de 168 participants de la biobanque « Multi-virus », collectée au Centre Hospitalier Universitaire de Sainte-Justine (CHUSJ) auprès d'enfants présentant des infections virales, à savoir le COVID-19, RSV et Influenza. Cette étude de nature exploratoire a intégré l'ensemble des échantillons disponibles, simulant ainsi une situation réelle de dépistage avec une importante variété d'échantillons. Des différences spectrales ont été observées entre les échantillons positifs aux différents virus et les échantillons de contrôle, aussi bien qualitativement, par comparaison des spectres normalisés, que quantitativement par des analyses statistiques univariées. De plus, le facteur de l'âge, pour une population pédiatrique a été identifié comme responsable de différences spectrales notables entre les enfants de moins de 6 mois et ceux âgés de plus de 6 mois. Finalement, cette étude confirme la possibilité de détection de maladies virales dans le sang par spectroscopie Raman et valide l'objectif initial.

Enfin, les résultats obtenus soulignent le potentiel de cette technique pour un diagnostic non invasif, offrant une approche prometteuse pour un dépistage rapide et efficace de plusieurs virus. À terme, cette méthode pourrait devenir un outil précieux en clinique, permettant de surveiller les infections virales et d'adapter rapidement les stratégies de prise en charge.

ABSTRACT

Raman spectroscopy is emerging as a promising technique for identifying specific disease biomarkers in biofluids without requiring complex sample preparation. In this context, LUMEDLAB developed a portable, rapid, and precise spectroscopy system along with machine learning models that validated this capability by detecting COVID-19 biomarkers in saliva. While continuously enhancing this system, the current goal is to identify characteristic biomarkers of viral diseases (COVID-19, RSV, and Influenza) in blood samples.

My work involved improving the spectroscopy system to better meet clinical diagnostic needs, with developments I contributed to, including component miniaturization, an innovative sample containment approach based on a well created within a hollow cassette covered by an aluminum sheet, and a new calibration step to ensure Raman signal quality. These technical improvements helped maintain the system's robustness and reliability, while ensuring high spectral resolution, essential for identifying subtle molecular biomarkers.

To test this device and validate the experimental protocol, a study was conducted on a set of 284 samples from 168 participants of the "Multi-virus" biobank, collected at the Sainte-Justine University Hospital Center (CHUSJ) from children with viral infections, specifically COVID-19, RSV, and Influenza. This exploratory study included all available samples, simulating a real screening scenario with a wide variety of samples. Spectral differences were observed between samples positive for the various viruses and control samples, both qualitatively, through comparison of normalized spectra, and quantitatively through univariate statistical analyses. Additionally, the age factor in a pediatric population was identified as responsible for notable spectral differences between children younger than 6 months and those older than 6 months. Ultimately, this study confirms the feasibility of detecting viral diseases in blood via Raman spectroscopy, validating the initial objective.

Finally, the results underscore the potential of this technique for non-invasive diagnostics, offering a promising approach for rapid and efficient multi-virus screening. Eventually, this method could become a valuable clinical tool, enabling viral infection monitoring and facilitating prompt adjustments in care strategies.

TABLE DES MATIÈRES

DÉDICACE.....	III
REMERCIEMENTS	IV
RÉSUMÉ.....	V
ABSTRACT	VII
LISTE DES TABLEAUX.....	XI
LISTE DES FIGURES	XII
LISTE DES SIGLES ET ABRÉVIATIONS	XVI
LISTE DES ANNEXES.....	XVII
CHAPITRE 1 INTRODUCTION.....	1
CHAPITRE 2 REVUE DE LITTÉRATURE	4
2.1 Spectroscopie Raman : fondements et applications	4
2.1.1 Diffusion Raman dans les biofluides	4
2.1.2 Composants et instrumentation d'un système Raman.....	6
2.1.3 Applications aux biofluides : avantages et défis	7
2.2 Détection de maladies dans les biofluides.....	10
2.2.1 Détection de cancers.....	10
2.2.2 Détection de maladies virales dans le sang	15
2.2.3 Travaux du LUMEDLAB liés à l'étude des biofluides.....	16
CHAPITRE 3 DÉVELOPPEMENT DU SYSTÈME RAMAN ET OPTIMISATION DES PROTOCOLES	18
3.1 Présentation du système	18
3.1.1 Système Raman du laboratoire détaillé	18
3.1.2 Mise en fonctionnement du système et traitement spectral	24

3.2	Étude cassette/cuvette sans échantillons	27
3.2.1	Enjeux et méthode	27
3.2.2	Résultats et choix final	28
3.3	Développement des cassettes de calibration	31
3.4	Étude avec des échantillons de salive	33
3.4.1	Technique de filtres : méthodes	33
3.4.2	Technique de filtres : résultats	35
3.4.3	Expériences préparatoires : influence de la congélation, entonnoirs	37
CHAPITRE 4	ÉTUDE PRINCIPALE « MULTI-VIRUS ».....	40
4.1	Étude préliminaire – Répétabilité des mesures d'échantillons sanguins.....	40
4.1.1	Méthodes	41
4.1.2	Résultats	41
4.1.3	Discussion	43
4.2	Méthodes – Étude principale : « Multi-Virus ».....	45
4.2.1	Contexte et données	45
4.2.2	Protocole expérimental.....	48
4.3	Résultats – Étude principale : « Multi-Virus »	50
4.3.1	Nature exploratoire de la recherche	50
4.3.2	Premier résultat – Comparaison entre échantillons infectés et échantillons de contrôle	51
4.3.3	Résultats intermédiaires – Comparaisons entre chaque maladie	56
4.3.4	Dernier résultat – Comparaison des échantillons en fonction de l'âge des patients	61
4.4	Discussion – Étude principale : « Multi-Virus ».....	63
CHAPITRE 5	DISCUSSION GÉNÉRALE	67

CHAPITRE 6 CONCLUSION	69
RÉFÉRENCES	70
ANNEXE	75

LISTE DES TABLEAUX

Tableau 2.1. Articles démontrant l'utilisation de la spectroscopie Raman pour la détection des cancers de la vessie ou de la prostate dans les biofluides humains. Ce tableau présente un ensemble d'articles ayant exploré l'application de la spectroscopie Raman (RS) dans la détection de ces cancers. Les mesures de performance utilisées pour évaluer l'efficacité de la méthode employée sont indiquées : <i>Sens</i> (sensibilité), <i>Spec</i> (spécificité), <i>Acc</i> (précision). .	12
Tableau 3.1. Liste des filtres utilisés dans le cadre de l'étude d'optimisation du protocole de mesure d'échantillons de salive. Les filtres possèdent une taille de pores identique, mais sont fabriqués à partir de matériaux différents.	34
Tableau 4.1. Données de l'étude « Multi-Virus ».	47
Tableau 4.2. Identification de certains pics Raman observés dans nos résultats. Ces pics sont valables pour des échantillons sanguins et correspondent à des longueurs d'onde responsables de différences significatives entre deux groupes de notre étude. Ils sont tirés de la littérature scientifique (1 : [53], 2 : [54], 3: [55], 4: [56], 5: [57]).....	65

LISTE DES FIGURES

Figure 2.1. Diagramme de Jablonski.	Les principales interactions lumière-matière sont représentées dans ce diagramme. Le phénomène de diffusion Raman Stokes, qui constitue la base de ce mémoire, est indiqué en rouge.	5
Figure 2.2. Biofluides étudiés avant et après centrifugation.	Ce schéma présente les différents biofluides analysés au sein du laboratoire, à différents stades de préparation : dans leur état initial avant centrifugation, ainsi qu'après centrifugation, lorsque les composants cellulaires et les débris ont été séparés, laissant un surnageant clair pour les analyses spectroscopiques. (Image conçue par Antoine Noé aux fins d'un article de revue qui sera soumis au <i>Journal of Biomedical Optics</i>)	8
Figure 3.1. Schéma du système de spectroscopie Raman, tiré de [51].	Ce schéma illustre le trajet de la lumière dans un système en épi-illumination, qui se rend du laser à l'échantillon puis au spectromètre. L'ordinateur permet de contrôler l'activation du laser en entrée et de collecter les données du spectromètre en sortie.	19
Figure 3.2. Premier système biofluides. a)	Le système est composé d'une boîte pour l'acquisition des échantillons et d'un chariot mobile transportant le laser et le spectromètre. b)	20
À l'intérieur de la boîte, les composants optiques, la plateforme de translation et l'échantillon étudié sont placés dans un environnement optique isolé.		
Figure 3.3. Second système biofluides – version actuelle. a)	Ce système compact est logé dans une valise facilement transportable, intégrant le laser et le spectromètre. b)	22
La boîte lecteur permet une acquisition simple de l'échantillon grâce à un système de tiroir. Elle est reliée à la valise par les fibres d'excitation et de collection, ainsi le trajet de la lumière est respecté entre le laser, l'échantillon et le spectromètre.....		
Figure 3.4. Dispositifs en aluminium conçus pour contenir l'échantillon de biofluide. a)	La cassette, disponible avec différentes dimensions de puits, permet de mesurer de petits volumes de biofluide (jusqu'à 60µL). b)	23
La cuvette a été conçue pour accueillir des volumes d'échantillon plus importants.....		

Figure 3.5. Capture d'écran de l'interface de contrôle du système. La partie gauche de l'écran permet de choisir les paramètres d'acquisition, de suivre l'état d'activation du laser, ou encore de récupérer les fichiers en sortie du spectromètre. La partie droite apporte une première visualisation du spectre brut de l'échantillon, permettant de vérifier en temps réel la bonne qualité de la mesure.....	24
Figure 3.6. Spectres d'un échantillon de sérum sanguin à différentes étapes du traitement spectral. a) Spectre brut moyenné, représentant les données initiales après acquisition. b) Spectre après calibration des axes et ligne de base, permettant de retirer le signal de fluorescence et de standardiser les données. c) Spectre Raman normalisé final, pour les études de comparaison et l'analyse et la discussion des résultats.	26
Figure 3.7. Spectres SNV d'aluminium, des dispositifs sans échantillons, de cassette a) et de cuvette b). La légende des figures inclut la date de mesure, le numéro d'identification et le temps d'exposition pour chaque échantillon mesuré et présenté.	29
Figure 3.8. Spectres SNV d'aluminium, moyennés sur les données cassette et cuvette. Cette étude de comparaison a pour but de rendre compte des différences entre les deux techniques et de choisir le dispositif final pour les mesures de l'étude « Multi-Virus ».	30
Figure 3.9. Cassettes conçues pour l'étape de calibration. a) Vue de face des cassettes pour la mesure de calibration du Tylenol et du NIST. b) Vue de côté de la cassette du Tylenol, pour illustrer le dispositif de fermeture.	32
Figure 3.10. Spectres SNV de salive, moyennés sur 7 participants. Les quatre filtres étudiés ont été utilisés par tous les participants à la suite d'une collection d'échantillons salivaires.	36
Figure 3.11. Spectres SNV de salive, d'un seul participant, avec deux filtres différents. Ces mesures illustrent l'impact des variations inter-individuelles sur les échantillons de salive.	36
Figure 3.12. Spectres SNV de salive, avec et sans entonnoir, pour un participant. La courbe noire issue de l'échantillon 8003, représente un échantillon collecté sans entonnoir, tandis que la courbe rouge correspond à l'échantillon 8009, collecté à l'aide d'un entonnoir. Ce résultat démontre l'absence de signal Raman supplémentaire à l'ajout d'un entonnoir dans le protocole de collecte d'échantillons de salive.	38

Figure 4.1. Illustration d'une goutte de sang déposée à l'intérieur du puits formé à partir de la cassette. Le volume de la goutte étant de 30 μ L, et alors que ce puits peut contenir des gouttes allant jusqu'à 60 μ L, le liquide doit se trouver au centre du puits pour obtenir une mesure fidèle à cet échantillon.	41
Figure 4.2. Spectres SNV de sang, provenant de différents échantillons de rat. La sous-figure a) représente les spectres associés aux échantillons de plasma alors que la sous-figure b) montre ceux des échantillons de sérum. La légende des figures inclut la date de mesure, le numéro d'identification et le temps d'exposition pour chaque échantillon mesuré et présenté.	42
Figure 4.3. Spectres SNV de sang, provenant de la même mesure et utilisés pour discuter de l'influence d'un paramètre du code de traitement. a) Ce spectre Raman, obtenu avec les paramètres utilisés jusqu'à précédent, met l'accent sur l'intensité des pics. b) Ce spectre Raman met davantage l'accent sur la forme générale du spectre, correspondant ainsi de plus fidèlement aux spectres Raman déjà obtenus dans le laboratoire et dans la littérature.	43
Figure 4.4. Spectres SNV de sang, moyennés sur des mesures d'échantillons provenant d'un seul rat. La courbe bleue représente la moyenne des échantillons de sérum et la courbe rouge montre la moyenne des échantillons de plasma. Ce résultat souligne le faible impact des variations inter-individuelles pour les échantillons sanguins, ainsi que l'absence de différences significatives entre des échantillons de plasma et de sérum.	44
Figure 4.5. Système Raman dans le cas de mesures en conditions réelles. a) Le système est situé sur un chariot mobile, tandis que la boîte lecteur se trouve sous une hotte, essentielle dans le cadre de manipulation d'échantillons biologiques. b) Les éléments nécessaires aux manipulations que l'on retrouve sous la hotte sont illustrés ici. De plus, les éléments importants sont indiqués et détaillés à l'aide d'une numérotation.	46
Figure 4.6. Résultats de l'étude de comparaison des échantillons testés positivement à une des infections virales, soit COVID-19, RSV et/ou Influenza, et des échantillons testés négativement à ces trois maladies. a) Comparaison des spectres SNV de cette étude. Les moyennes et écart-types sont affichés respectivement en rouge et en bleu. b) Différence entre les moyennes des spectres précédents, en pourcentage du maximum de l'intensité. Les régions grisées correspondent aux longueurs d'onde pour lesquelles la p-value calculée est inférieure	

à 0.05. c) P-values inférieures à 0.05 pour chaque longueur d'onde, illustrant les régions spectrales significativement différentes entre les groupes d'échantillons. d) Diagrammes en boîtes associés à des régions spectrales où plusieurs longueurs d'onde partagent des p-values < 0.05. Les seuils de significativité statistique sont indiqués par des étoiles *: $0.01 < p < 0.05$, **: $0.005 < p < 0.01$, ***: $p < 0.005$	54
Figure 4.7. Résultats de l'étude de comparaison des échantillons testés positivement au COVID-19 et des échantillons testés négativement aux trois maladies virales étudiées.	56
Figure 4.8. Résultats de l'étude de comparaison des échantillons testés positivement au RSV et des échantillons testés négativement aux trois maladies virales étudiées.	57
Figure 4.9. Résultats de l'étude de comparaison des échantillons testés positivement à Influenza et des échantillons testés négativement aux trois maladies virales étudiées.	58
Figure 4.10. Diagrammes en boîtes illustrant les différences entre tous les groupes d'échantillons. Les diagrammes sont associés à des régions spectrales où plusieurs longueurs d'onde provenant de différents groupes partagent des p-values < 0.05.	60
Figure 4.11. Résultats de l'étude de comparaison des échantillons provenant d'enfants âgés de moins de six mois et ceux provenant d'enfants de plus de six mois.	62

LISTE DES SIGLES ET ABRÉVIATIONS

AUC	<i>Area Under the Curve</i> (Aire sous la courbe)
BPH	<i>Benign Prostatic Hyperplasia</i> (Hyperplasie bénigne de la prostate)
CCD	<i>Charged-Coupled Device</i> (Dispositif à transfert de charge)
CHUSJ	Centre Hospitalier Universitaire de Sainte-Justine
CNN	<i>Convolutional Neural Network</i> (Réseau de neurones convolutionnel)
FTIR de Fourier)	<i>Fourier Transform Infrared Spectroscopy</i> (Spectroscopie infrarouge à transformée de Fourier)
GA	<i>Genetic Algorithm</i> (Algorithme génétique)
LDA	<i>Linear Discriminant Analysis</i> (Analyse discriminante linéaire)
PCA	<i>Principal Component Analysis</i> (Analyse par composante principale)
PLS	<i>Partial Least Squares</i> (Moindres carrés partiels)
QF	<i>Quality Factor</i> (Facteur de qualité)
RBF	<i>Radial Basis Function</i> (Fonction à base radiale)
RS	<i>Raman Spectroscopy</i> (Spectroscopie Raman)
RSV	<i>Respiratory Syncytial Virus</i> (Virus respiratoire syncytial)
SBR	<i>Signal to Background Ratio</i> (Rapport signal sur signal de fond)
SERS de surface)	<i>Surface Enhanced Raman Spectroscopy</i> (Spectroscopie Raman augmentée par effet de surface)
SNR	<i>Signal to Noise Ratio</i> (Rapport signal sur bruit)
SNV	<i>Standard Normal Variate</i>
SVM	<i>Support Vector Machine</i> (Machine à vecteurs de support)

LISTE DES ANNEXES

Figures supplémentaires.....	75
------------------------------	----

CHAPITRE 1 INTRODUCTION

En 2024, les maladies virales continuent de représenter un défi majeur pour les systèmes de santé mondiaux. Avec l'augmentation des échanges internationaux et la résistance croissante de certains pathogènes, la détection rapide et précise de ces virus est cruciale. Parmi les techniques d'analyse prometteuses, la spectroscopie Raman se distingue par son potentiel unique à fournir des informations moléculaires précises à partir de biofluides [1] ou de tissus [2]. Ces informations s'avèrent très utiles, permettant une identification des signatures moléculaires caractéristiques des agents pathogènes viraux ou des réponses de l'hôte [3]. De plus, la spectroscopie Raman dans les biofluides a également montré un potentiel considérable dans le domaine de l'oncologie, où elle est utilisée pour détecter et caractériser différents types de cancers. Plusieurs études ont mis en évidence sa capacité à identifier des signatures spectrales spécifiques associées à des biomarqueurs de cancers, notamment dans le cancer du sein [23], prostate [27] ou vessie [28]. La spectroscopie Raman permet ainsi de différencier des tissus contenant des cellules cancéreuses de tissus sains grâce aux variations dans les compositions chimiques, offrant un diagnostic non invasif et rapide.

Cependant, bien que la spectroscopie Raman soit déjà largement utilisée pour identifier des substances dans le domaine chimique notamment, son application à la détection de maladies virales ou cancéreuses reste relativement limitée. En effet, les biofluides, tels que le sang, l'urine ou la salive, permettent d'obtenir un faible signal Raman, en raison de l'intensité relative du phénomène de diffusion Raman et de la composition, complexe et instable, de ces biofluides. En parallèle, les défis cliniques et épidémiologiques actuels soulignent l'urgence d'améliorer les diagnostics : les erreurs ou retards de diagnostic peuvent entraîner des traitements inadaptés, tels qu'une prescription non justifiée d'antibiotiques pour des infections virales ou une absence de traitement antiviral pertinent, ce qui peut aggraver la morbidité et favoriser des problèmes majeurs de santé publique, la résistance aux antibiotiques entre autres.

Pour passer à travers ces défis, le LUMEDLAB a développé une plateforme de diagnostic rapide et précis pour la détection de pathologies dans différents biofluides (sang, urine, salive). Les premiers travaux du laboratoire dans ce domaine ont démontré la capacité du système à détecter des biomarqueurs liés au COVID-19 dans la salive, avec des critères de sensibilité et spécificité pertinents [1].

Un des objectifs principaux dorénavant est de démontrer l'utilité polyvalente de ce dispositif expérimental pour le diagnostic d'un large éventail de maladies, par exemple différents types de cancer ou maladies virales. Il faut alors rassembler des échantillons de biofluides et acquérir leurs spectres Raman, ainsi que des variables démographiques/cliniques pour tous ces échantillons associés à différentes pathologies, afin de tenir compte de potentiels facteurs confondants. Ce travail de maîtrise s'inscrit dans cette continuité de recherche, en contribuant au développement du système Raman du laboratoire et à l'amélioration de sa performance pour le diagnostic de pathologies.

Le développement de cette plateforme basée sur la spectroscopie Raman et des modèles d'apprentissage automatique est au cœur de ce projet. Cependant, la positivité d'un échantillon à une certaine maladie est établie par des méthodes de classification dont les résultats dépendent grandement de la qualité des spectres Raman.

L'objectif de mon projet est d'optimiser le signal Raman pour la détection de biomarqueurs dans des échantillons de biofluides, puis de démontrer son utilité potentielle dans le cadre d'une détection d'infections virales chez les enfants.

L'hypothèse de ce travail repose sur les résultats préalablement obtenus par le laboratoire, et se formule comme suit. Le système de spectroscopie développé par le LUMEDLAB permet d'obtenir un signal Raman reproductible et faiblement bruité, tout en révélant des différences spectrales entre des échantillons provenant d'une population malade et des échantillons de contrôle.

Pour organiser le travail qui nous attend et proposer une démarche rigoureuse, il est préférable de découper l'objectif principal en quelques objectifs intermédiaires :

- Objectif n°1 : Implémenter de nouvelles fonctionnalités au système de spectroscopie afin d'alléger sa structure et d'optimiser sa rapidité et sa facilité d'utilisation pour l'utilisateur.
- Objectif n°2 : Mettre en place un protocole expérimental garantissant la reproductibilité des mesures d'échantillons salivaires et sanguins.
- Objectif n°3 : Appliquer les choix de développement apportés au système et les protocoles expérimentaux élaborés, dans le cadre d'une étude clinique à partir d'une biobanque d'échantillons sanguins de participants affectés de maladies virales.

Ce document est divisé en six chapitres, incluant celui-ci. La partie principale du rapport va suivre les objectifs intermédiaires, avec un chapitre dédié exclusivement au système de spectroscopie Raman, aux améliorations apportées et études ayant permis son développement. Un autre chapitre présentera en détails l'étude finale de ce projet. De plus, ce document inclura également une introduction comprenant une mise en contexte et une présentation des enjeux du projet, la définition de l'hypothèse de travail et des objectifs de recherche, ainsi qu'une bibliographie permettant de se familiariser avec les différents concepts physiques et de passer en revue le travail déjà effectué sur ce sujet. Enfin viendront une partie discussion, dirigée par un bref retour sur les objectifs définis et quelques perspectives intéressantes pour la suite de la recherche, et finalement une conclusion.

CHAPITRE 2 REVUE DE LITTÉRATURE

2.1 Spectroscopie Raman : fondements et applications

2.1.1 Diffusion Raman dans les biofluides

La spectroscopie Raman repose sur le phénomène physique de la diffusion Raman, l'une des interactions possibles entre la lumière et la matière. Lors de ces interactions, un photon est mis en jeu, d'énergie $h\nu_p$ et de longueur d'onde λ_p , régies par l'équation suivante [4] :

$$E_{photon} = h\nu_p = h\frac{c}{\lambda_p} \quad (2.1)$$

Le résultat de cette interaction lumière-matière est un échange d'énergie : le photon transmet son énergie à la molécule, puis, en fonction de la quantité d'énergie dont dispose le photon, deux phénomènes peuvent se produire au niveau de la molécule.

Si l'énergie du photon correspond exactement à la différence d'énergie entre l'état fondamental de la molécule et un état excité, alors celui-ci est absorbé et la molécule se retrouve dans un état excité : c'est le phénomène d'absorption. Sinon, si l'énergie du photon ne permet pas à la molécule de se rendre dans un état excité, cette dernière se rend alors dans un état virtuel et réemet un nouveau photon : c'est le phénomène de diffusion. Ce photon peut avoir exactement la même énergie que le photon incident, ce qui correspond à la diffusion élastique, aussi appelée diffusion Rayleigh. En revanche, si l'énergie du photon réémis est différente, il s'agit de la diffusion Raman ou diffusion inélastique. Dans ce cas, le photon réémis peut avoir une énergie inférieure, ce qui équivaut à une longueur d'onde plus élevée d'après l'équation précédente (diffusion Raman Stokes) ou une énergie supérieure (diffusion Raman anti-Stokes) [5]. Ces différentes interactions sont représentées schématiquement sur le diagramme de Jablonski de la figure 2.1.

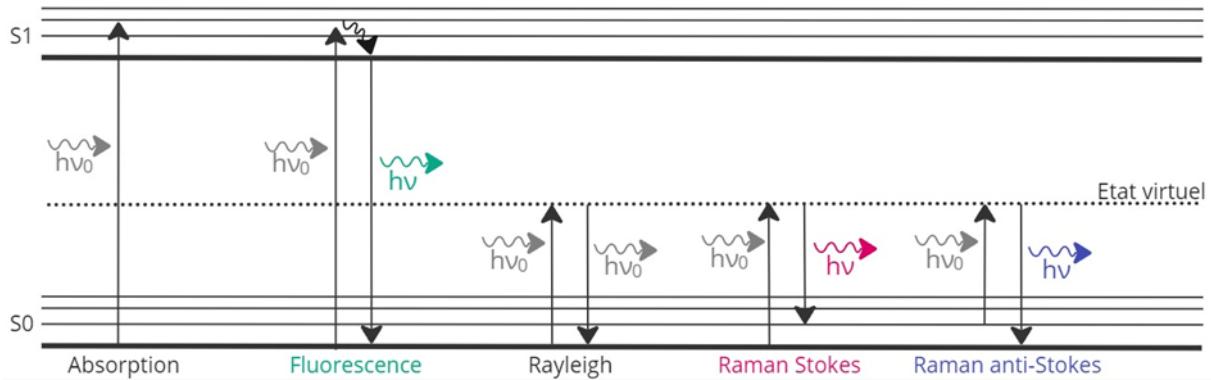


Figure 2.1. Diagramme de Jablonski. Les principales interactions lumière-matière sont représentées dans ce diagramme. Le phénomène de diffusion Raman Stokes, qui constitue la base de ce mémoire, est indiqué en rouge.

Physiquement, le phénomène de diffusion Raman est possible grâce aux différents niveaux d'énergie vibrationnelle de la molécule. Ces niveaux d'énergie disponibles permettent au photon réémis d'accéder à plusieurs longueurs d'onde disponibles, caractéristiques des niveaux vibrationnels de la molécule. Le spectre Raman, soit le nombre de photons diffusés en fonction de la longueur d'onde, permet ainsi d'obtenir une signature moléculaire de l'échantillon.

La dernière interaction lumière-matière, également présente sur le diagramme de Jablonski de la figure 2.1, est la fluorescence. Suite à l'absorption du photon par la molécule, celle-ci se retrouve dans un état excité, puis va dissiper une partie de l'énergie absorbée sous forme de vibrations internes (dissipation de chaleur), dépendamment de sa structure électronique. Une fois cette énergie libérée, la molécule émet un nouveau photon pour revenir à son état fondamental : c'est la fluorescence. Ainsi, de même que pour la diffusion Raman Stokes, le photon réémis aura une longueur d'onde plus élevée que celle du photon incident, raison pour laquelle on retrouvera un signal de fluorescence superposé aux spectres Raman.

En termes d'ordre de grandeur, il est bon de rappeler qu'environ 1 seul photon sur 1000 photons incidents est sujet à la diffusion élastique, mais seulement 1 photon sur 10 millions pour la diffusion Raman Stokes et 1 sur 100 millions pour la diffusion anti-Stokes [6]. Cette dernière se fait plus rare car pour avoir la chance de se produire, la molécule doit déjà se trouver dans un état vibrationnel excité. Le phénomène sur lequel se concentre l'équipe de recherche du LUMEDLAB

est la diffusion Raman Stokes, que l'on simplifiera en parlant de « diffusion Raman » dans ce mémoire.

De nombreuses applications dans le domaine biomédical utilisent le principe de la spectroscopie Raman. En effet, cette partie théorique a illustré qu'il était possible d'identifier à l'aide d'un spectre certaines molécules ou liaisons moléculaires d'un échantillon. En pratique, ce phénomène est étudié au sein du laboratoire pour caractériser, par exemple, des échantillons de salive positifs au coronavirus [1], ou encore pour détecter des tumeurs lors d'une opération chirurgicale [2].

2.1.2 Composants et instrumentation d'un système Raman

Cette partie a pour objectif de présenter les principaux instruments mis en jeu pour l'acquisition d'un spectre Raman. Tout d'abord, le premier élément essentiel à tout système Raman est une source de lumière, permettant de mettre en place les photons incidents nécessaires au phénomène de diffusion. Cette lumière incidente est obtenue à partir d'un laser, qui présente de nombreux avantages. En effet, un laser produit une lumière quasi-monochromatique, c'est-à-dire une longueur d'onde unique et bien définie, ce qui permet de mesurer précisément les changements de longueur d'onde de la lumière diffusée. De plus, le laser peut fournir une puissance lumineuse élevée, maximisant ainsi le signal Raman, qui est généralement très faible. Son faisceau est cohérent et focalisé, offrant une excitation précise au niveau de l'échantillon. Enfin, la stabilité temporelle du laser garantit des mesures reproductibles et fiables.

Le laser utilisé dans le système du laboratoire possède une longueur d'onde de 785nm, caractéristique en spectroscopie Raman dans les tissus biologiques [7], notamment pour deux raisons. Premièrement, l'intensité du phénomène de diffusion Raman dépend de cette longueur d'onde d'excitation, selon la formule [8] :

$$I_{Raman} \propto \frac{1}{\lambda^4} \quad (2.2)$$

D'après cette relation de proportionnalité correspondant à une réalité physique, il faudrait prendre la longueur d'onde la plus petite possible pour maximiser la diffusion Raman. Cependant, il convient également de prendre en considération un autre élément pour choisir la longueur d'onde de notre source d'excitation. En effet, le phénomène de fluorescence, expliqué précédemment, est

hautement présent dans les tissus biologiques : les fluorophores endogènes, présents intrinsèquement dans les biofluides, émettent entre 350 et 600nm [9]. Pour minimiser le signal de fluorescence et éviter que celui-ci domine le signal de diffusion Raman qui nous intéresse, la longueur d'onde utilisée devra se trouver dans le proche infra-rouge. Finalement, une longueur d'onde d'excitation de 785nm semble optimale afin d'atteindre un équilibre entre ces deux critères.

Le second élément qui compose un système Raman est le spectromètre, permettant de séparer la lumière diffusée par l'échantillon en différentes longueurs d'onde. Il est composé d'un réseau de diffraction, dispositif optique qui permet la décomposition spectral, et d'un détecteur CCD, qui convertit le nombre de photons arrivant sur la caméra en un signal électronique exploitable numériquement. En complément, d'autres composants optiques comme des miroirs ou lentilles, permettent de guider la lumière de façon efficace à travers le spectromètre, maximisant ainsi la résolution spectrale et le ratio signal-sur-bruit (SNR), donc la qualité spectrale. De ce fait, un spectre Raman est communément représenté comme l'intensité du signal (nombre de photons de la diffusion Raman) en fonction du décalage en longueur d'onde par rapport à la source, plutôt qu'en longueur d'onde absolue. On parle alors de décalage Raman exprimé en nombre d'onde (en cm^{-1}) à partir de la longueur d'onde d'excitation, qui sera toujours égale à 785 nm dans ce travail.

Les autres éléments qui entrent en compte dans un système Raman sont principalement des composants optiques, propres à chaque système, qui ont comme buts principaux de stabiliser le faisceau lumineux et de maximiser la puissance au niveau de l'échantillon. On peut aussi noter l'utilisation de différents filtres optiques : par exemple, pour éliminer la lumière obtenue mais non désirée par diffusion Rayleigh qui étoufferait le signal Raman, un filtre passe-haut autour de la longueur d'onde d'excitation pourra être intégré au système.

2.1.3 Applications aux biofluides : avantages et défis

La spectroscopie Raman permet alors de caractériser certaines molécules ou liaisons moléculaires, ce qui présente un véritable attrait pour les tissus biologiques et les biofluides en particulier. Les biofluides principaux tels que le sang, pour lequel on peut différencier sérum et plasma, l'urine et la salive, sont des milieux très complexes contenant une multitude de biomolécules, considérés comme des biomarqueurs lorsqu'ils sont interrogés à l'aide de la spectroscopie Raman [10]. Pour homogénéiser ces milieux complexes, l'étape de centrifugation (ou même de filtration pour la salive), telle que détaillée dans la figure 2.2, est primordiale [11].

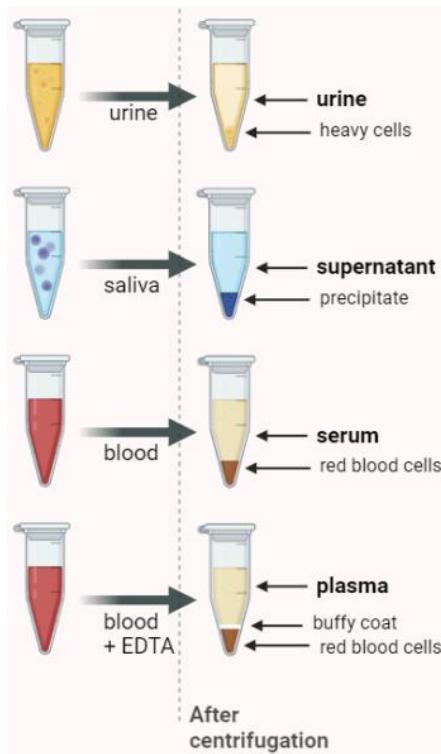


Figure 2.2. Biofluides étudiés avant et après centrifugation. Ce schéma présente les différents biofluides analysés au sein du laboratoire, à différents stades de préparation : dans leur état initial avant centrifugation, ainsi qu'après centrifugation, lorsque les composants cellulaires et les débris ont été séparés, laissant un surnageant clair pour les analyses spectroscopiques. (Image conçue par Antoine Noé aux fins d'un article de revue qui sera soumis au *Journal of Biomedical Optics*)

C'est une étape importante avant de prendre les mesures, pour clarifier les biofluides en éliminant les particules indésirables tout en conservant les biomolécules d'intérêt. Voici des détails spécifiques pour chaque biofluide mentionné. Le sérum est récupéré dans un échantillon de sang après coagulation : la centrifugation du sang coagulé permet de séparer la phase liquide et claire, le sérum, des cellules sanguines et des caillots. Cela permet d'éliminer les globules rouges, les globules blancs et les plaquettes, tout en conservant les protéines, les électrolytes et les autres biomarqueurs potentiels présents dans le sérum. Le plasma, contrairement au sérum, contient les facteurs de coagulation, car il est extrait après un traitement à l'anticoagulant, généralement de l'EDTA. La centrifugation de ce sang non coagulé permet de séparer le plasma des cellules sanguines. L'urine, quant à elle, est un biofluide moins visqueux que le sang, mais elle contient également des particules lourdes, comme des cellules ou des bactéries. La centrifugation permet de clarifier l'urine et de séparer le sédiment urinaire de la partie surnageante, qui est ensuite utilisée

pour les mesures. Enfin, la salive peut être assez visqueuse dépendamment des échantillons, et peut contenir des particules solides, des cellules buccales et des débris alimentaires. Pour l'homogénéisation, la centrifugation est souvent précédée ou remplacée par une étape de filtration. Cela permet d'obtenir un surnageant clair, qui contient les biomolécules importantes pour les spectres Raman, comme les protéines et les enzymes. En résumé, la centrifugation des biofluides est essentielle et permet d'éliminer les particules lourdes et d'obtenir des échantillons homogènes pour optimiser la stabilité des mesures, tout en préservant les biomarqueurs recherchés en spectroscopie Raman.

Ces biofluides représentent une grande promesse dans le domaine biomédical pour la spectroscopie Raman, offrant des avantages indéniables. Premièrement, cette technique est non invasive et non destructive, c'est-à-dire que l'échantillon du biofluide étudié n'est aucunement altéré par la mesure. L'échantillon peut ainsi être réutilisé à d'autres fins, ce qui représente une valeur ajoutée par rapport à d'autres techniques nécessitant des marqueurs ou des colorants pour pouvoir étudier l'échantillon. Deuxième argument, la collecte des échantillons de biofluides est très simple et surtout peu contraignante pour le participant, ce qui en fait une méthode viable à long terme et hautement prometteuse dans le cadre d'un test de dépistage. De plus, au niveau de la préparation des échantillons, en plus de l'étape de centrifugation, il n'existe peu ou pas d'étapes supplémentaires. Pour conclure, il est possible de travailler avec de toutes petites quantités de biofluides, de l'ordre de la goutte, soit $10\mu\text{L}$ de liquide.

Cependant, l'utilisation de la spectroscopie Raman pose également des défis importants à surmonter pour assurer des résultats fiables et exploitables. En effet, l'importante fluorescence intrinsèque de ces biofluides, associée au faible signal Raman, peut masquer les biomarqueurs d'intérêt ou rendre leur détection difficile, ce qui nécessite des systèmes de détection très sensibles et exige une précision accrue. De plus, la composition des biofluides étudiés peut varier considérablement entre les personnes, et même chez une même personne en fonction de plusieurs critères : heure de la journée, alimentation, hydratation, sport, hygiène, sommeil, etc. [12] Le protocole de collecte d'échantillons doit alors être le plus constant et précis possible, pour éviter les biais de mesure systématiques. Malgré tous les efforts effectués dans cette étape, le signal Raman d'un échantillon de biofluide reste particulièrement faible, c'est pourquoi il est primordial d'optimiser au maximum le système de spectroscopie et d'utiliser des méthodes de traitement de données standardisées et éprouvées.

2.2 Détection de maladies dans les biofluides

Cette partie a pour objectif premier de faire correspondre le travail déjà existant dans la littérature pour la détection de quelques maladies dans les biofluides, avec le travail effectué dans le laboratoire. Cette analyse permettra par la suite d'identifier les méthodes et résultats les plus prometteurs afin d'optimiser les performances du système Raman utilisé et ainsi d'acquérir des spectres pertinents.

2.2.1 Détection de cancers

Le potentiel de la spectroscopie Raman pour la caractérisation d'échantillons de biofluides de personnes atteintes de cancer est très prometteur. En ce sens, les trois types d'échantillons, à savoir sang, urine et salive, présentent des avantages intéressants et sont susceptibles d'être utilisés dans cette application.

La méthode de spectroscopie Raman améliorée par effet de surface (SERS) est plébiscitée pour l'analyse de biofluides, car cette technique permet d'amplifier considérablement le signal Raman en ajoutant des nanoparticules à l'échantillon, ce qui permet de détecter des concentrations moléculaires plus faibles. Bien que cette technique soit très développée [13], il s'agit de s'intéresser ici à une utilisation de la spectroscopie Raman sans étiquette, c'est-à-dire sans recourir à des substances externes pour améliorer le signal ou marquer les molécules ciblées, car le développement de cette technique est compliqué à mettre en place et peu reproductible pour une réalité clinique.

Même si les trois biofluides peuvent être utilisés, c'est le sang et notamment le sérum qui semble le plus populaire pour la détection de cancer. Cela peut s'expliquer par une facilité de prélèvement et de conservation, mais aussi une présence importante de nombreux biomarqueurs dans le sang et de plus, des changements moléculaires dans le sang peuvent refléter directement des modifications dues à la présence de cancer chez l'individu. La spectroscopie Raman à l'épreuve du sang a ainsi été utilisée avec succès pour détecter plusieurs types de cancer : foie [14], poumon [15], col de l'utérus [16], bouche [17], colon [18], œsophage [19]...

Pour la salive maintenant, le cancer de la bouche est le plus plausible à identifier ([20], [21], [22]), car directement en contact avec le biofluide concerné, mais la détection d'autres types de cancer est néanmoins possible (sein [23], poumon [24], ovaire [25] ou estomac [26] par exemple).

Enfin, l'urine est également utilisé pour identifier la présence de cellules cancéreuses, permettant de mettre en lumière les cancers de prostate [27], vessie [28], pancréas [29] ou sein [30].

Le tableau suivant résume le travail de revue de la littérature effectué, en mettant en avant les articles scientifiques qui démontrent la capacité de détecter les cancers de la vessie ou de la prostate à partir de biofluides humains. Dans chaque article, plusieurs éléments clés ont été relevés afin d'illustrer les principales méthodes et résultats : le type de biofluide étudié, la méthode d'imagerie utilisée, l'intégration de la spectroscopie SERS, la taille et les détails de la population étudiée, les modèles de classification appliqués pour analyser les données ainsi que des indicateurs de performance témoignant des résultats. Bien que nos recherches se soient étendues à d'autres types de cancers, ce tableau répertorie ici uniquement les résultats relatifs à la vessie et à la prostate pour mieux structurer l'analyse. Il est important de noter qu'un grand nombre d'études confirment la faisabilité de détecter divers types de cancer à partir de biofluides en utilisant la spectroscopie Raman.

Tableau 2.1. Articles démontrant l'utilisation de la spectroscopie Raman pour la détection des cancers de la vessie ou de la prostate dans les biofluides humains. Ce tableau présente un ensemble d'articles ayant exploré l'application de la spectroscopie Raman (RS) dans la détection de ces cancers. Les mesures de performance utilisées pour évaluer l'efficacité de la méthode employée sont indiquées : *Sens* (sensibilité), *Spec* (spécificité), *Acc* (précision).

Référence	Biofluide	Méthode d'imagerie	SERS - non SERS	Taille population	Détails population	Modèle de classification	Résultats-Performance
Li. 2014 [31]	blood (serum)	RS Renishaw	SERS	93 prostate cancer 68 healthy		SVM, PCA-LDA	Sens 97,8% Spec 100% Acc 98,1%
Del Mistro 2015 [27]	urine	RS Renishaw	SERS colloid	9 prostate cancer 9 healthy	Patients undergoing radical prostatectomy	PCA-LDA	Sens 100%, Spec 89% Acc 95%
Medipally 2020 [32]	blood (plasma lymphocytes)	RS FTIR	non SERS	43 prostate cancer 33 healthy	Plasma 37cancer /33h Lymph.29cancer /26h	PCA, PLS-DA, CLS	Sens/Spec ranging from 90% to 99%
Ma 2020 [33]	urine	RS	SERS colloid	75 prostate cancer	12 Re / 63 NRe patient cohorts	GA-PLS-LDA	Sens 86,0% Spec 87,1% Acc 86,6%
Chen 2021 [34]	urine	RS	SERS colloid	39 prostate cancer 45 healthy (BPH)	only included T1 and T2 stages	CNN	Sens 77,32% Spec 72,46% Acc 74,95%
Chen 2017 [35]	blood (serum)	RS Renishaw	SERS colloid	40 prostate cancer 40 healthy (BPH)	> 65 y.o and total PSA within 4–10 ng/mL	PCA-LDA	Sens 95,0% Spec 93,8% Prec 88,4% Acc 94,2%
Medipally 2017 [36]	blood (plasma)	RS	non SERS	10 prostate cancer 10 healthy	male with an age range from 58 to 85	PCA-LDA	Sens 96,5% Spec 95%
Shao 2017 [37]	blood (serum) EPS	RS Renishaw	SERS colloid	20 prostate cancer 20 healthy (BPH)		PCA-LDA	Sens (75%,60%) Spec (75%,76,5%) Acc (75%,68%) (EPS, serum)
Stefancu 2018 [38]	blood (serum)	RS Renishaw	SERS colloid	30 prostate cancer 24 healthy	All patients presented clinical suspicion for prostate cancer (either PSA > 4 ng/ml or abnormal digital rectal examination)	PCA-LDA	Acc 94%

Munteanu 2022 [39]	blood (plasma serum)	RS Renishaw	SERS	29 prostate cancer 14 healthy	both serum/plasma samples collected	PCA-LDA, PLS-DA, PCA-SVM	Serum Sens 96,6% Spec 100% - Plasma Sens 96,3% Spec 71,4%
Phyo 2021 [29]	urine	RS portable	SERS	44 cancer 30 healthy	22 pancreatic cancer 22 prostate cancer	PCA OPLS-DA	Sens 100% Spec 100%
Gao 2022 [40]	blood (serum)	RS Renishaw	SERS	72 cancer 30 healthy	40 liver cancer 32 prostate cancer	PLS-SVM	Acc 100% cancer/healthy
Moisoiu 2019 [41]	blood (serum)	RS Renishaw	SERS colloid	214 cancer 39 healthy	42 breast, 109 colorectal, 33 lung, 17 oral 13 ovarian cancers	PCA-LDA	Sens 98% Spec 91%
Cui 2020 [28]	urine	RS	SERS colloid	101 cancer 57 healthy	51 bladder cancer 50 prostate cancer	PCA-LDA	Acc 91,9%
Hu 2021 [42]	urine	RS Horiba	SERS colloid	161 bladder cancer 87 healthy	not previously received any radiotherapy or chemotherapy	PCA-LDA	Sens/Spec 100/98.85% for high-grade tumours 97.53/90.80% for low-grade tumours.
Chen 2019 [43]	blood (serum)	RS Horiba	SERS colloid	60 bladder cancer 30 healthy	28 NMIBC/32 MIBC	PLS-LDA	Acc 97.8% healthy/cancer 93.2% NMIBC/MIBC
Li 2015 [44]	blood (serum)	RS Renishaw	SERS colloid	55 bladder cancer 36 healthy	cancer confirmed with histopathology	GA-LDA	Sens 90,9% Spec 100%
Zhang 2018 [45]	blood (serum)	RS Renishaw	SERS colloid	48 bladder cancer 40 healthy	23 low-grade 25 high-grade	RBF-SVM	Acc 96.4% low grade Acc 95.4% high-grade

Parmi les travaux remarquables, l'étude de Li et al de 2014 [31] se démarque avec une sensibilité (*Sens*), soit la capacité à détecter la maladie (vrais positifs) de 97,8 %, une spécificité (*Spec*), soit la capacité à exclure la maladie (vrais négatifs) de 100 % et une précision (*Acc*), soit la proportion globale de résultats corrects du test, de 98,1 % pour la détection du cancer de la prostate à partir d'échantillons de sérum, en utilisant la technique de spectroscopie SERS et des méthodes d'analyse telles que les machines à vecteurs de support (SVM) et l'analyse discriminante linéaire (PCA-LDA). Ce travail démontre la possibilité de détecter des échantillons de type cancer de la prostate avec des résultats quantitatifs satisfaisants, d'autant plus avec une quantité d'échantillons significative, ainsi que l'efficacité de la spectroscopie SERS à amplifier le signal Raman dans des biofluides. D'autres études ont également montré des résultats prometteurs, comme celle de Chen et al de 2017 [35], qui a utilisé la spectroscopie SERS sur du sérum pour différencier des patients atteints du cancer de la prostate par rapport à des patients atteints de BPH, agrandissement bénin de la prostate, avec une sensibilité de 95,0% et une spécificité de 93,8%, ou encore celle de Ma et al de 2020 [33], qui a utilisé 75 échantillons d'urine positifs au cancer de la prostate, avec une sensibilité de 86,0% et une spécificité de 87,1%, démontrant que la classification du cancer de la prostate était également possible avec des échantillons d'urine. De plus, d'autres chercheurs se sont intéressés à la détection du cancer de la vessie, comme l'étude de Cui et al de 2020 [28] qui a réussi à collecter des échantillons de ces 2 types de cancer et obtenant une précision pour ces résultats de 91,9%, ou l'étude de Hu et al de 2021 [42] avec 161 échantillons d'urine de personnes atteintes du cancer de la vessie, qui ont présenté leurs résultats en fonction du stade d'avancement de la tumeur et qui ont obtenu des données plus concluantes pour les tumeurs de haut-grade (sensibilité/ spécificité 100/98,85 % pour les tumeurs de haut grade, 97,53/90,80 % pour les tumeurs de bas grade), démontrant qu'il était effectivement possible mais significativement plus difficile de détecter une tumeur dans son état initial que dans un état plus avancé.

Certes, la spectroscopie Raman appliquée aux biofluides permet de détecter plusieurs types de cancer à un stade plus ou moins avancé de la maladie, mais elle permettrait de détecter aussi d'autres maladies importantes, au niveau sociétal, avec précision.

2.2.2 Détection de maladies virales dans le sang

Après avoir réussi à caractériser des spectres Raman provenant de salive de personnes porteuses de la maladie COVID-19, il semble pertinent de s'intéresser à la détection du COVID-19 dans le sang cette fois [46]. Pour élargir l'étude, il convient aussi de s'intéresser à d'autres maladies virales.

Les maladies virales sont des infections causées par des virus, des organismes microscopiques ne pouvant se reproduire qu'en envahissant les cellules d'un hôte. Les virus se propagent principalement par contact direct avec des fluides corporels ou par inhalation de gouttelettes respiratoires contenant des particules virales. Ils peuvent également se transmettre facilement par des surfaces contaminées, ce qui facilite leur diffusion rapide dans la population.

Parmi les virus les plus courants dans les pays développés, on retrouve le virus SARS-CoV-2, responsable de la pandémie de COVID-19, le virus de la grippe (Influenza), causant des épidémies saisonnières et pouvant entraîner des symptômes graves, et le virus respiratoire syncytial (RSV), touchant principalement les jeunes enfants.

De ce fait, la spectroscopie Raman a été explorée comme outil prometteur pour la détection de maladies virales, telles que le COVID-19 et le virus respiratoire syncytial (RSV), grâce à son potentiel de différenciation des échantillons en fonction de leurs signatures spectrales spécifiques. Pour le COVID-19, une étude de Yin et al. [47] a utilisé la spectroscopie Raman pour analyser le sérum sanguin, en discriminant les patients confirmés (63), suspects (59), et les individus sains (55) à partir de données spectrales. : l'analyse globale, combinant les techniques d'analyse en composantes principales (PCA) et d'analyse discriminante linéaire (LDA), ainsi qu'un algorithme de machine à vecteur de support (SVM) et une courbe ROC, a permis une classification des échantillons avec une sensibilité, spécificité et précision respectivement de 91 %, 88 % et 90 %. De même, d'autres études (Carlomagno et al. [48] et Desai et al. [49] notamment) présentent des conclusions similaires, ce qui prouve que la spectroscopie Raman appliquée au sérum sanguin peut discriminer efficacement les différents groupes de santé en lien avec le COVID-19. Concernant le virus RSV, Shanmukh et al. [50] ont appliqué une technique de spectroscopie SERS sur des échantillons de patients positifs au RSV, obtenus dans un milieu de culture sans sérum. L'utilisation combinée de PCA et d'analyse de clusters hiérarchiques (HCA) a permis de différencier les souches du RSV, avec des résultats de spécificité et de sensibilité atteignant respectivement 96 % et 88 %. Finalement, ces études mettent en avant la capacité de la spectroscopie Raman, notamment en

combinaison avec des outils d'apprentissage automatique, de détecter les signatures spectrales de virus dans les biofluides, offrant des perspectives de développement de méthodes de diagnostic rapides et peu invasives pour les infections virales.

L'étude principale de ce projet portera alors sur la détection de ces trois maladies virales (COVID-19, RSV et Influenza) sur une population pédiatrique. L'objectif est de caractériser par spectroscopie Raman la signature moléculaire d'échantillons de sang prélevés chez des patients pédiatriques diagnostiqués avec des virus respiratoires, et de les comparer à des échantillons de témoins présentant des symptômes similaires mais ayant testé négatif pour ces infections virales.

2.2.3 Travaux du LUMEDLAB liés à l'étude des biofluides

La première recherche utilisant un biofluide menée par le LUMEDLAB, a employé la salive soumise à la spectroscopie Raman pour détecter la maladie COVID-19. Cette étude [1] avait pour but de mettre au point un moyen de détecter le COVID-19 dans la salive sans ajout de réactif, et dans un contexte réel, tout en minimisant les contraintes en matière d'acquisition d'échantillons. En organisant une collecte d'échantillons pendant la pandémie à grande échelle, l'équipe a pu avoir accès à de nombreux échantillons de salive, testés positif ou négatif pour le COVID-19. Ces échantillons de salive séchée ont ensuite été mesurés à l'aide d'un système de spectroscopie Raman industriel : les résultats obtenus permettent de mettre en évidence des changements biomoléculaires entre les échantillons de salive de personnes atteintes du COVID-19 et les personnes saines, et les modèles prédictifs développés se sont révélés être plus précis en prenant en compte le sexe de la personne testée. Cependant, les contraintes d'hétérogénéité de la goutte de la salive séchée, et de temps d'acquisition trop long pour une technique de dépistage viable, limitent les résultats.

L'idée est alors de développer un système portable de spectroscopie Raman à point unique, similaire à la sonde déjà développée et devant permettre d'imager un échantillon liquide en seulement quelques secondes. L'objectif ultime de ce système est d'être utilisé comme moyen de dépistage performant à haut débit d'échantillons en situation de contexte épidémiologique. Plusieurs aspects de ce rapport couvriront le développement de ce système et l'optimisation du protocole de collecte d'échantillons. Dans une étude plus récente, le LUMEDLAB, avec l'aide d'une autre équipe de recherche spécialiste du domaine, s'est intéressé au plasma sanguin pour approfondir la détection de la sévérité du COVID-19 grâce à la spectroscopie Raman. Les résultats

obtenus montrent une capacité remarquable à discriminer les cas critiques des cas non-critiques de COVID-19 avec une précision de 83 %. Les résultats de la comparaison entre les individus en bonne santé et les patients atteints de formes critiques du COVID-19 atteignent une précision de 98 %, tandis que ceux entre les individus en bonne santé et les cas non-critiques atteint 99 %. De plus, les modèles prédictifs développés dans cette étude permettent non seulement de prédire avec une précision de 90 % le risque de décès après l'infection au COVID-19, mais aussi d'anticiper le type de ventilation requis (non-invasive ou intubation) avec une précision de 94 %. Ces résultats ouvrent des perspectives très intéressantes pour le suivi mais aussi le tri des patients en contexte clinique, en offrant un moyen potentiellement rapide et précis de surveiller l'évolution des cas de COVID-19 et d'ajuster les interventions médicales selon la gravité de la maladie.

Afin de positionner les travaux du LUMEDLAB dans un contexte actuel, il est pertinent de comparer la spectroscopie Raman à d'autres approches couramment utilisées ou émergentes pour le dépistage du COVID-19. Les tests RT-PCR, considérés comme le standard de référence, offrent une sensibilité et une spécificité très élevées, mais nécessitent des réactifs et des équipements coûteux, en plus de délais d'analyse plus longs que ceux requis par la technique développée par le LUMEDLAB. Les tests antigéniques, bien qu'accessibles et pratiques, présentent une sensibilité moindre, ce qui limite leur fiabilité. D'autres techniques récentes, bien qu'affichant des indices de performance prometteurs, demeurent peu adaptées à un dépistage rapide dans des conditions cliniques réelles. En revanche, la spectroscopie Raman comme moyen de dépistage semble s'imposer comme une alternative efficace, combinant rapidité d'analyse, coût réduit et portabilité.

CHAPITRE 3 DÉVELOPPEMENT DU SYSTÈME RAMAN ET OPTIMISATION DES PROTOCOLES

Cette partie a pour but de détailler le système de spectroscopie développé par le LUMEDLAB pour obtenir le spectre Raman d'un échantillon, et d'expliquer comment chaque caractéristique du système a été optimisée. Par la suite, les études menées pour le développement du système et l'optimisation du protocole de mesure d'échantillons de salive seront détaillées. Mon travail a consisté à mesurer une biobanque d'échantillons salivaires au début de ma maîtrise avec le premier système biofluide, me permettant de prendre en main le système et de comprendre en profondeur son fonctionnement. Par la suite, j'ai pris part à la conception du second système, en participant à des expériences visant à en tester les nouveaux paramètres, puis j'ai développé les techniques de cassette et de cuvette, conçues pour contenir l'échantillon de biofluide, ainsi que les cassettes spécifiques à l'étape de calibration.

3.1 Présentation du système

3.1.1 Système Raman du laboratoire détaillé

Les deux principales pièces qui composent le système sont une source laser de longueur d'onde d'excitation 785nm (*Model IO785MM1500M4S, Innovative Photonic Solutions, USA*) d'une puissance de 1,5W et d'une largeur de bande spectrale inférieure à 2 nm, ainsi que le spectromètre (*modèle HT, EmVision, USA*), composé d'un réseau de diffraction et d'une caméra CCD (*Newton 920, Oxford Instruments, USA*) permettant d'obtenir une résolution spectrale inférieure à 8,7 cm⁻¹. Un logiciel d'acquisition interne personnalisé MATLAB (*MathWorks, MA, USA*) et un microcontrôleur (*Arduino*) sont utilisés pour contrôler l'entrée laser et récupérer les fichiers contenant les données (paramètres d'acquisition, spectres) à la sortie du spectromètre.

Entre ces éléments, de nombreux composants optiques sont utilisés, dans le but d'exciter l'échantillon avec la puissance lumineuse maximale et de récupérer le maximum de lumière diffusée. Pour réussir ces défis, le système optique sera monté en épi-illumination, c'est-à-dire que la lumière d'excitation et la lumière diffusée suivent en partie le même chemin lumineux, grâce à un miroir dichroïque permettant de réfléchir certaines longueurs d'onde et d'en transmettre d'autres.

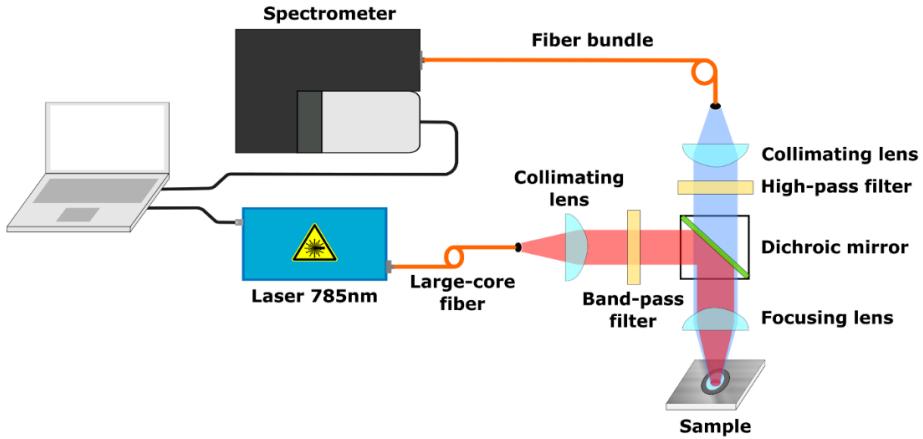


Figure 3.1. Schéma du système de spectroscopie Raman, tiré de [51]. Ce schéma illustre le trajet de la lumière dans un système en épi-illumination, qui se rend du laser à l'échantillon puis au spectromètre. L'ordinateur permet de contrôler l'activation du laser en entrée et de collecter les données du spectromètre en sortie.

A partir de la figure 3.1, il est intéressant de suivre le trajet de la lumière pour bien comprendre le fonctionnement du système : la lumière est confinée dans une fibre optique à la sortie du laser puis passe dans une lentille de collimation. Le but de cette lentille est de collimer le faisceau lumineux, pour éviter qu'il diverge et permettre ainsi une meilleure précision ensuite au niveau de l'échantillon. Elle permet aussi de contrôler le chemin de la lumière et de faciliter l'alignement des autres composants optiques. En conclusion, la collimation est tout simplement primordiale pour contrôler la lumière dans le système et obtenir les effets désirés des composants optiques. Ensuite, un filtre passe-bande autour de la longueur d'excitation va permettre de filtrer la lumière incidente : en effet, un signal indésirable peut apparaître dans la fibre d'excitation en raison du matériau même de la fibre, i.e. la silice. Ensuite, la lumière d'excitation est réfléchie par le miroir dichroïque puis passe par une lentille de focalisation, pour concentrer toute la lumière en un point précis, au cœur de l'échantillon. C'est à ce moment-là que le phénomène de diffusion Raman se produit, il faut alors récupérer le plus de lumière diffusée possible. Cette diffusion est un phénomène isotrope, elle se produit dans toutes les directions de l'espace. Une partie de cette lumière diffusée repasse alors par la lentille de focalisation (dans le sens inverse par rapport à la lumière d'excitation), puis est transmise par l'intermédiaire du miroir dichroïque. Ensuite, un filtre passe-haut autour de la longueur d'excitation permet de filtrer les longueurs d'ondes inférieures non voulues, dans le but

de conserver seulement la diffusion Raman et supprimer la diffusion Rayleigh. Celle lumière passe enfin dans une lentille de collimation puis dans un faisceau de fibres de collection jusqu'à l'entrée du spectromètre. Ce faisceau se compose de 9 fibres optiques, réparties de manière circulaire, avec un diamètre de cœur de $300\mu\text{m}$ et une ouverture numérique de 0,22. L'optique de focalisation était constituée d'une lentille de 25mm de diamètre avec une longueur focale de 50mm. Cela permet d'obtenir une distance de focalisation d'environ 12,6mm à une longueur d'onde d'excitation de 785nm.

Par la suite, le système optique en entier sera remplacé par une sonde commerciale, de la marque *Wasatch Photonics*, simplifiant considérablement le montage et permettant de gagner beaucoup d'espace, mais le principe du parcours de la lumière reste le même. Cette première version du système, comme présentée dans la figure 3.2, est la première mise au point dans le laboratoire. La boîte que l'on voit sur la photo contient seulement la partie avec les composants optiques, le laser et le spectromètre se trouvant en dessous.

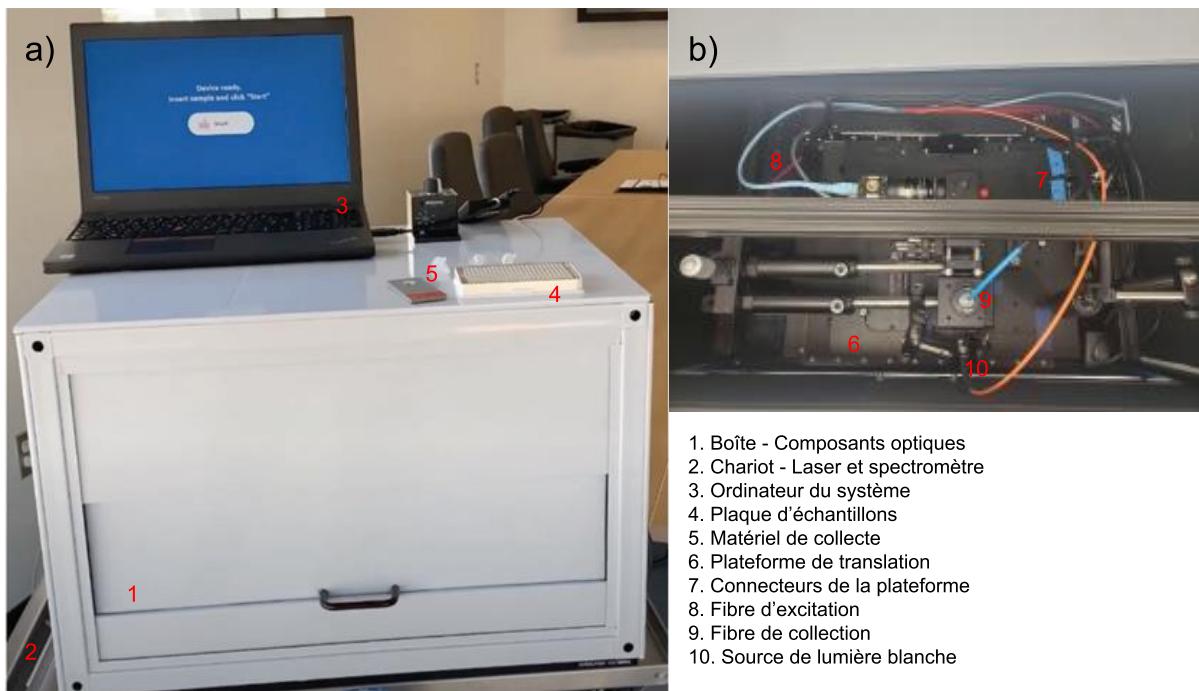


Figure 3.2. Premier système biofluides. a) Le système est composé d'une boîte pour l'acquisition des échantillons et d'un chariot mobile transportant le laser et le spectromètre. b) À l'intérieur de la boîte, les composants optiques, la plateforme de translation et l'échantillon étudié sont placés dans un environnement optique isolé.

Dans cette version du système, l'échantillon se composait d'une plaque industrielle ordonnée et pourvue de nombreux puits, puis recouverte d'une grande feuille d'aluminium. Cette feuille d'aluminium est utilisée comme support pour tous les échantillons liquides de ce rapport, car sa forme est malléable et surtout, l'aluminium ne possède pas de signal Raman significatif, ce qui réduit drastiquement le bruit au niveau du spectre Raman de l'échantillon d'intérêt. Cette plaque était fixée sur une plateforme de translation 3D, permettant de mettre l'échantillon exactement à la distance focale de la lentille selon un axe, mais aussi de prendre la mesure de plusieurs gouttes en une seule fois en déplaçant la plateforme selon les 2 autres axes. Cette configuration était pratique et fonctionnait très bien avec des gouttes séchées, cependant, elle ne fonctionne plus dès lors qu'il s'agit de prendre des mesures de gouttes liquides. En effet, ces dernières sèchent trop rapidement, ce qui fait que les mesures doivent être effectuées une par une pour conserver la répétabilité de l'expérience, et la plateforme perd alors tout son intérêt. Il y a également un défi de taille avec la plateforme de translation, qui occupe beaucoup d'espace. Une caméra et une source de lumière blanche situées en amont de l'échantillon permettaient de visualiser l'espace de travail.

Il a ensuite été décidé de réduire la taille du système et de se concentrer uniquement sur des échantillons de type liquide. L'objectif est d'arriver à un système, incluant laser et spectromètre, rentrant entièrement dans une valise de taille ordinaire, pouvant être transportée facilement. C'est ainsi qu'a été développée la nouvelle version du système, visible sur la figure 3.3. En plus de cette valise, une petite boîte isolée optiquement de l'extérieur, la boîte « lecteur », a été conçue. Elle y inclut une sonde optique, connectée à 2 fibres optiques : une d'excitation, reliée au laser, et l'autre de collection, reliée au spectromètre. Elle comporte également un tiroir imbriqué pour pouvoir facilement y insérer l'échantillon.

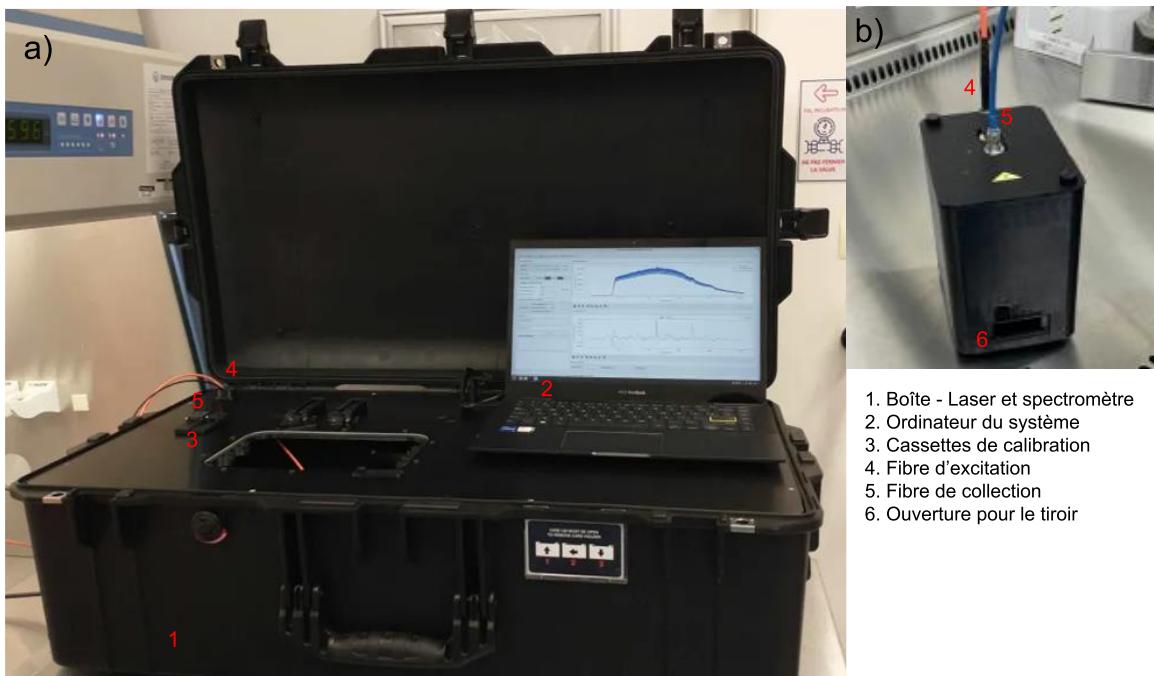


Figure 3.3. Second système biofluides – version actuelle. a) Ce système compact est logé dans une valise facilement transportable, intégrant le laser et le spectromètre. b) La boîte lecteur permet une acquisition simple de l'échantillon grâce à un système de tiroir. Elle est reliée à la valise par les fibres d'excitation et de collection, ainsi le trajet de la lumière est respecté entre le laser, l'échantillon et le spectromètre.

Dans cette version beaucoup plus épurée du système, la goutte n'est plus déposée sur la plaque industrielle constituée de plusieurs trous : il faut alors trouver le meilleur endroit où déposer l'échantillon. Les deux idées principales pour ce réceptacle sont une cassette comportant un seul trou et recouverte d'un morceau d'aluminium (figure 3.4 a), où l'avantage serait de contrôler la taille de la goutte de l'échantillon en contrôlant la taille du trou, et une cuvette creusée dans de l'aluminium solide (figure 3.4 b). Des tests de comparaison entre la cassette et la cuvette seront présentés dans un chapitre de ce rapport.

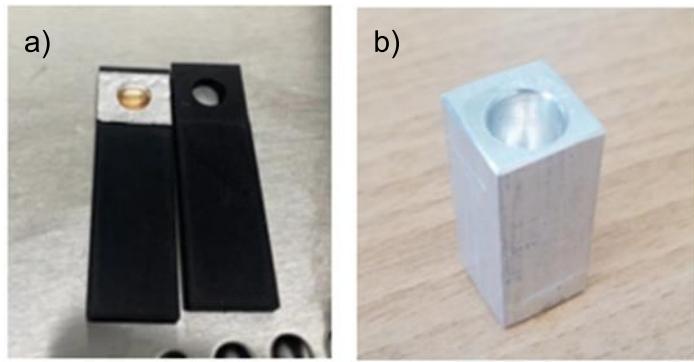


Figure 3.4. Dispositifs en aluminium conçus pour contenir l'échantillon de biofluide. a) La cassette, disponible avec différentes dimensions de puits, permet de mesurer de petits volumes de biofluide (jusqu'à $60\mu\text{L}$). b) La cuvette a été conçue pour accueillir des volumes d'échantillon plus importants.

Le système est constamment modifié, amélioré pour répondre aux besoins toujours changeants de l'équipe. Cependant, il est intéressant de noter quelque peu la temporalité du développement des systèmes. Le premier système a été développé par le LUMEDLAB et était opérationnel avant mon arrivée, alors que le second a été conçu durant ma maîtrise, avec la participation de l'ensemble de l'équipe Biofluides du laboratoire. Les cassettes ont été conçues puis fabriquées par impression 3D dans le cadre de ce projet, alors que la cuvette en aluminium a été découpée avec l'aide du département de génie physique de l'École Polytechnique de Montréal.

L'ordinateur du système vient avec une interface intuitive permettant à la fois de contrôler le laser et de récupérer les données du spectromètre. A partir de cette interface, il est possible de gérer les paramètres d'acquisition (puissance laser, temps d'exposition, nombre de spectres/accumulations par acquisition), de visualiser si le laser est allumé ou éteint, de noter des commentaires et même d'avoir un premier aperçu du spectre brut de l'échantillon que l'on vient tout juste de mesurer. C'est également ici que l'on observe le temps d'exposition automatique : en effet, pour chaque échantillon, il est possible de régler le temps d'exposition en mode automatique, qui se calibrera en fonction d'un certain pourcentage de la plage dynamique de la caméra. En général, on veut que le spectre brut soit à 85% de ce que la caméra est capable de détecter, donc cela peut prendre 400ms (pour le sang par exemple) ou 2000ms (avec la salive) pour chaque accumulation. Tout ceci est visible dans la capture d'écran de la figure 3.5.

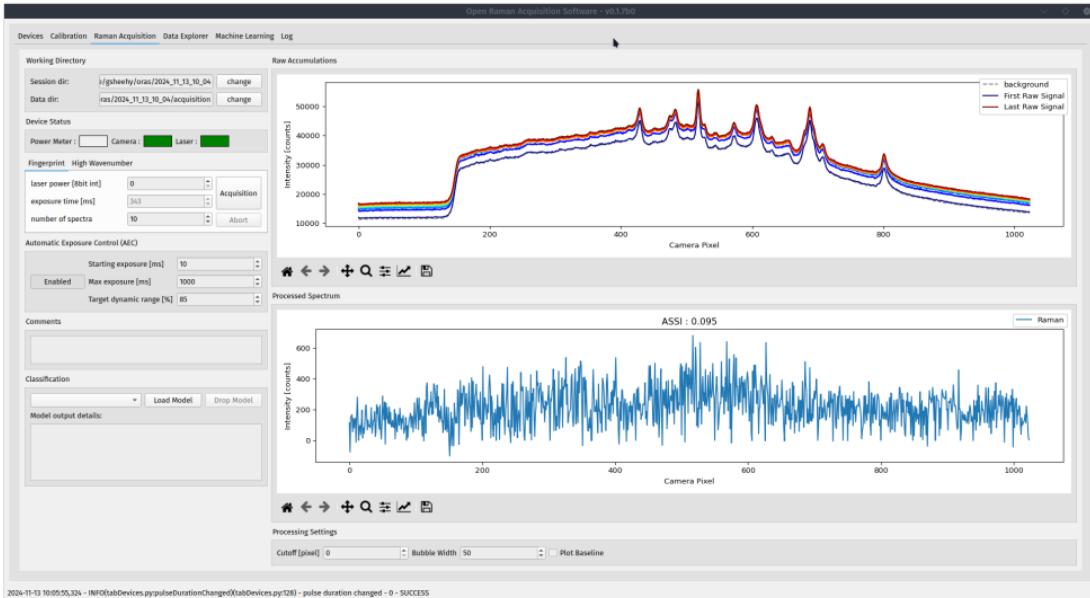


Figure 3.5. Capture d'écran de l'interface de contrôle du système. La partie gauche de l'écran permet de choisir les paramètres d'acquisition, de suivre l'état d'activation du laser, ou encore de récupérer les fichiers en sortie du spectromètre. La partie droite apporte une première visualisation du spectre brut de l'échantillon, permettant de vérifier en temps réel la bonne qualité de la mesure.

Finalement, tous les éléments qui composent le système de spectroscopie ont été présentés et discutés. Pour utiliser correctement le système et avoir des résultats viables et reproductibles entre toutes les mesures, il est important de suivre un protocole détaillé.

3.1.2 Mise en fonctionnement du système et traitement spectral

Avant chaque mesure, le capteur CCD du système a été refroidi à -65°C, pour réduire le bruit et augmenter la sensibilité du capteur notamment. Pour l'étape de calibration, l'étalonnage de l'axe des abscisses (décalage Raman) a été déterminé à partir d'un spectre acquis avec de la poudre d'acétaminophène (Tylenol®), donc les pics Raman sont connus et bien documentés, avant chaque mesure. Pour l'axe des ordonnées, la réponse du système a été caractérisée à l'aide du spectre d'un matériau de référence standard sensible à la longueur d'onde d'excitation (SRM 2214, *National Institute of Standards and Technology*, NIST, USA). Une étape de saturation peut être ajoutée à ce moment pour s'assurer d'occuper l'entièreté de la plage dynamique de la caméra, en mesurant la

poudre d'acétaminophène à une grande puissance. Les échantillons peuvent ensuite être mesurés, en précisant les caractéristiques de mesure sur l'interface présentée auparavant, dépendamment du type de biofluide. Cependant, après la prise de mesure, il y a de nombreuses opérations à effectuer avant d'arriver à un spectre Raman analysable et cohérent. Cette partie relate des principales étapes de la méthode de traitement de spectres développée par l'équipe du LUMEDLAB [52].

La première étape du traitement consiste à éliminer la portion initiale du spectre Raman, qui ne correspond pas à une réalité physique mesurable en raison de l'utilisation d'un filtre passe-haut destiné à supprimer la diffusion Rayleigh. Ensuite, les artefacts aléatoires induits par les rayons cosmiques doivent être identifiés et supprimés pour chaque acquisition spectrale. En effet, avec des temps d'exposition pouvant atteindre jusqu'à 2 minutes pour les biofluides, la probabilité d'impacts cosmiques devient non négligeable. Ces événements, produisant des pics de haute intensité non corrélés avec le signal Raman attendu, doivent être filtrés pour garantir l'intégrité des données. Une fois les rayons cosmiques retirés, il est impératif de soustraire le bruit de fond dû aux sources de perturbation externes, telles que la lumière ambiante et les interférences thermiques. Cette étape est suivie de la moyenne des 50 acquisitions individuelles, afin d'obtenir un spectre représentatif de l'échantillon et de réduire les contributions aléatoires (figure 3.6 a). Finalement, l'étape la plus importante du traitement spectral est la suppression de la ligne de base, réalisée à l'aide de l'algorithme *Bubblefill*, spécialement développé pour cette application. Cet algorithme permet de retirer le signal de fluorescence, pour obtenir un spectre Raman brut. La correction de la ligne de base est primordiale, car elle permet de révéler la structure vibratoire proprement dite, en supprimant l'effet parasite de la fluorescence, très présente dans les mesures sur des échantillons biologiques (figure 3.6 b). Enfin, bien que facultative, une étape de filtrage peut être appliquée pour lisser le spectre et atténuer les fluctuations aléatoires de faible amplitude. Une normalisation, par exemple en utilisant l'intensité maximale ou la zone sous le spectre, peut également être nécessaire pour permettre des comparaisons quantitatives entre différents spectres acquis dans des conditions expérimentales distinctes (figure 3.6 c).

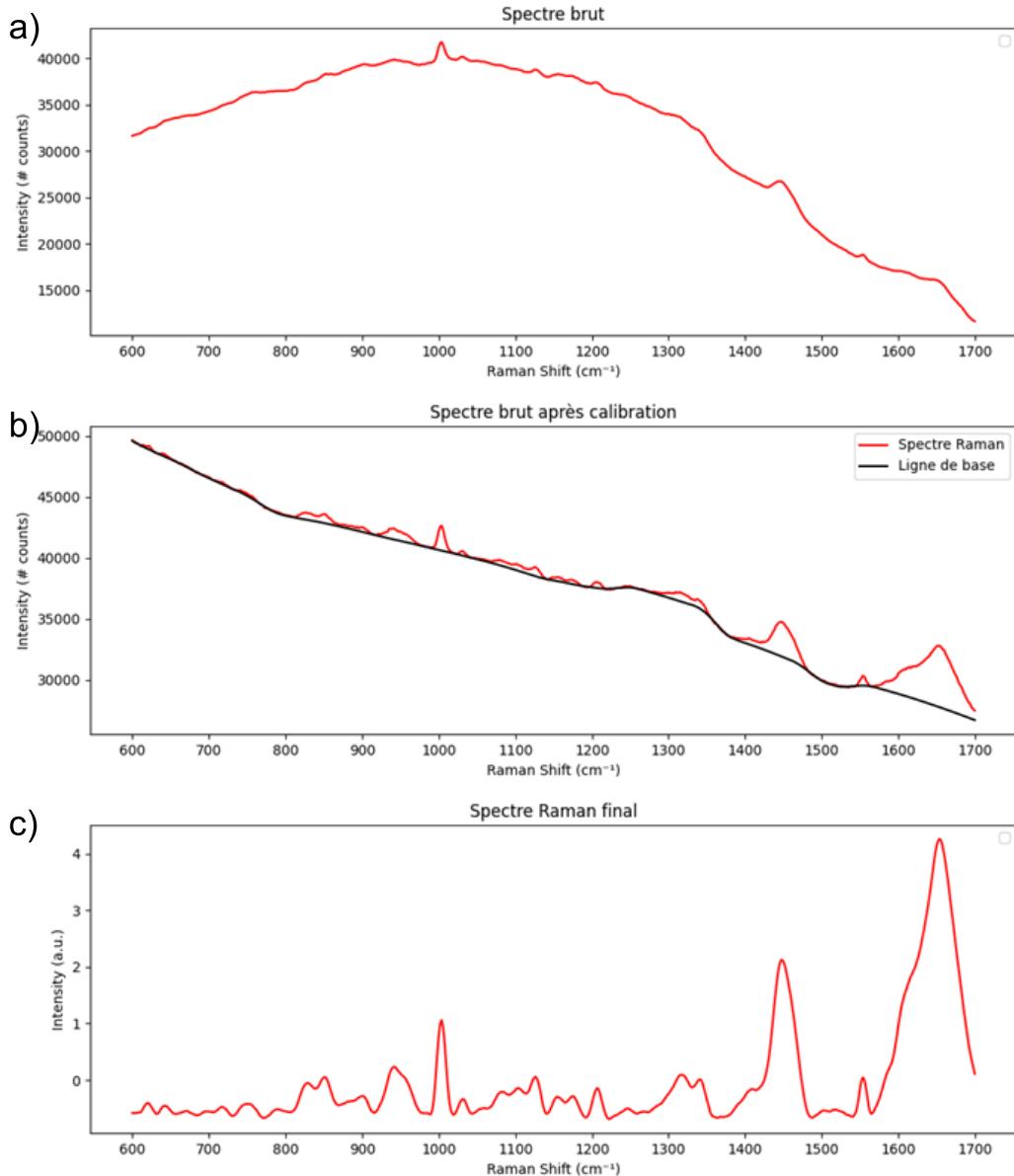


Figure 3.6. Spectres d'un échantillon de sérum sanguin à différentes étapes du traitement spectral. a) Spectre brut moyen, représentant les données initiales après acquisition. b) Spectre après calibration des axes et ligne de base, permettant de retirer le signal de fluorescence et de standardiser les données. c) Spectre Raman normalisé final, pour les études de comparaison et l'analyse et la discussion des résultats.

A la suite de ce traitement spectral, une métrique a été développé pour évaluer la qualité du signal Raman mesuré et répondre au caractère quantitatif de comparaisons spectrales. Premièrement, il s'agit du rapport signal sur bruit (SNR), défini comme le rapport entre le nombre total de photons du spectre Raman et la racine carrée du signal total (Raman + ligne de base + signal de fond), c'est-

à-dire tous les photons qui arrivent sur la caméra du spectromètre. Deuxièmement, on définit le rapport signal sur signal de fond (SBR), à savoir le rapport entre le nombre de photons dans le spectre Raman et dans la ligne de base. Un facteur de qualité (QF) a ensuite été calculé sur le SNV normalisé : c'est un nombre variant entre 0 et 1. Un spectre Raman avec une valeur de QF faible signifie que le signal possède un bruit élevé, une intensité de photons mesurés faible et des pics Raman finaux mal définis. Finalement, un signal aléatoire aurait un QF proche de 0,28. Cette métrique est utilisée pour faire un tri entre les spectres performants, utiles dans les comparaisons et la classification, et les autres, trop bruités pour être pertinents.

3.2 Étude cassette/cuvette sans échantillons

On va revenir en détails ici sur l'étude de comparaison entre la cassette et la cuvette comme support de biofluide, les raisons qui nous ont poussé à mener cette étude, les résultats observés puis le choix final, lequel des deux éléments nous avons conservé.

3.2.1 Enjeux et méthode

Premièrement, le support original était une cassette avec un puits de faible diamètre, pouvant contenir une goutte de l'ordre d'une dizaine de microlitres. Mais il faut garder en tête l'objectif initial : avoir le meilleur signal Raman possible. Or, en augmentant le volume de l'échantillon, on favorise davantage la diffusion Raman, ce qui améliore la qualité du spectre. Donc pour avoir un meilleur signal Raman, on peut tout simplement augmenter la taille de l'échantillon. On part alors de ce constat et on réfléchit à des supports offrant un volume plus important pour l'échantillon.

La première idée envisagée consiste à conserver la cassette tout en augmentant la taille du puits, bien que cette option soulève plusieurs aspects à examiner. Une des préoccupations majeures réside dans la limite de taille du trou, ainsi que dans la résistance de la feuille d'aluminium, qui pourrait se perforer lors de la création du puits destiné à recevoir la goutte. De plus, un risque de débordement de la goutte existe si le volume déposé dépasse la capacité du puits. Après plusieurs tests avec des cassettes de tailles de puits variées, une taille maximale raisonnable a été établie autour de 9 mm de rayon, correspondant à un volume théorique de 60 μL . En pratique, un volume supérieur peut être déposé, mais cela augmente le risque de débordement. Finalement, cette idée d'augmenter la taille du puits se confronte à une limite de volume maximal d'échantillon.

La seconde idée est de créer un puits directement dans un bloc d'aluminium solide, soit la cuvette. La seule contrainte est de concevoir la cuvette pour que sa hauteur soit assez petite pour lui permettre de rentrer dans la boîte « lecteur », mais le défi ici ne concerne pas le volume d'échantillon disponible. En effet, il est possible de créer un trou aussi profond que désiré, tant qu'il reste inférieur à la hauteur totale de la cuvette. Cependant, avec le matériel disponible dans le département de génie physique, il est difficile de contrôler précisément la taille du trou. De plus, si le trou est trop important par rapport à la quantité de liquide que l'on souhaite déposer, alors on va rencontrer des problèmes de tensions de surface et la goutte ne sera pas déposée de façon homogène au fond du trou. Il y a aussi le risque de nettoyage de la cuvette à prendre en compte lorsqu'on aura accès à des échantillons d'une biobanque sensible, mais un protocole rigoureux, comme explicité par la suite, réglera le problème. Ce problème est moins critique dans le cas de la cassette, parce que la feuille d'aluminium est retirée et une nouvelle y est apposée à chaque nouvel échantillon.

Il faut maintenant vérifier que la cuvette ne possède pas de signal Raman qui viendrait perturber nos mesures d'échantillons biologiques par la suite. Pour quantifier cela, on décide de prendre des mesures de cassette et cuvette à vide. Les paramètres suivants sont choisis pour cette expérience : puissance laser à l'échantillon d'environ 1W, pourcentage de la plage dynamique désiré de 85%, temps d'exposition automatique compris entre 100 et 10000ms par acquisition, puis 50 acquisitions par mesure. Bien que cette expérience ne comporte pas d'échantillons, le protocole de nettoyage, tel qu'il sera détaillé dans la méthode de l'étude principale, est respecté pour assurer la répétabilité des mesures, et on prend plusieurs fois chaque configuration dans un ordre aléatoire.

3.2.2 Résultats et choix final

Les spectres bruts présentent un signal très faible, ce qui semble concordant en l'absence d'échantillon ; il est donc préférable d'analyser ici les spectres normalisés. Pour chaque étude de comparaison et pour mettre en perspective un paramètre du protocole expérimental (influence de la cassette et de la cuvette dans cet exemple), il est décidé de regarder les spectres avec une normalisation SNV (*Standard Normal Variate*), qui permet de standardiser les données entre elles en fonction de leur moyenne et leur écart-type, plutôt que les spectres bruts. La normalisation SNV est calculée dans le traitement des données selon la formule :

$$SNV(i) = \frac{Raman(i) - \mu}{\sigma} \quad (3.1)$$

Ici, μ représente la moyenne du spectre Raman et σ son écart-type, indiquant la dispersion des valeurs autour de cette moyenne. Chaque spectre avec la normalisation SNV possède une valeur moyenne nulle et un écart-type égal à 1.

Pour chaque étude menée en spectroscopie Raman, il faut premièrement s'assurer que les mesures effectuées dans la même configuration possèdent le même signal, pour confirmer la stabilité du système et la fiabilité du protocole. Dans un second temps, la comparaison sera faite en prenant la moyenne des spectres Raman de chaque configuration et leurs écart-types : de cette manière, des différences spectrales pourront être observées selon les configurations. Sur la figure 3.7 a), quelques spectres Raman de cassette vide sont alors tracés, tandis que la sous-figure b) présente des spectres avec la cuvette.

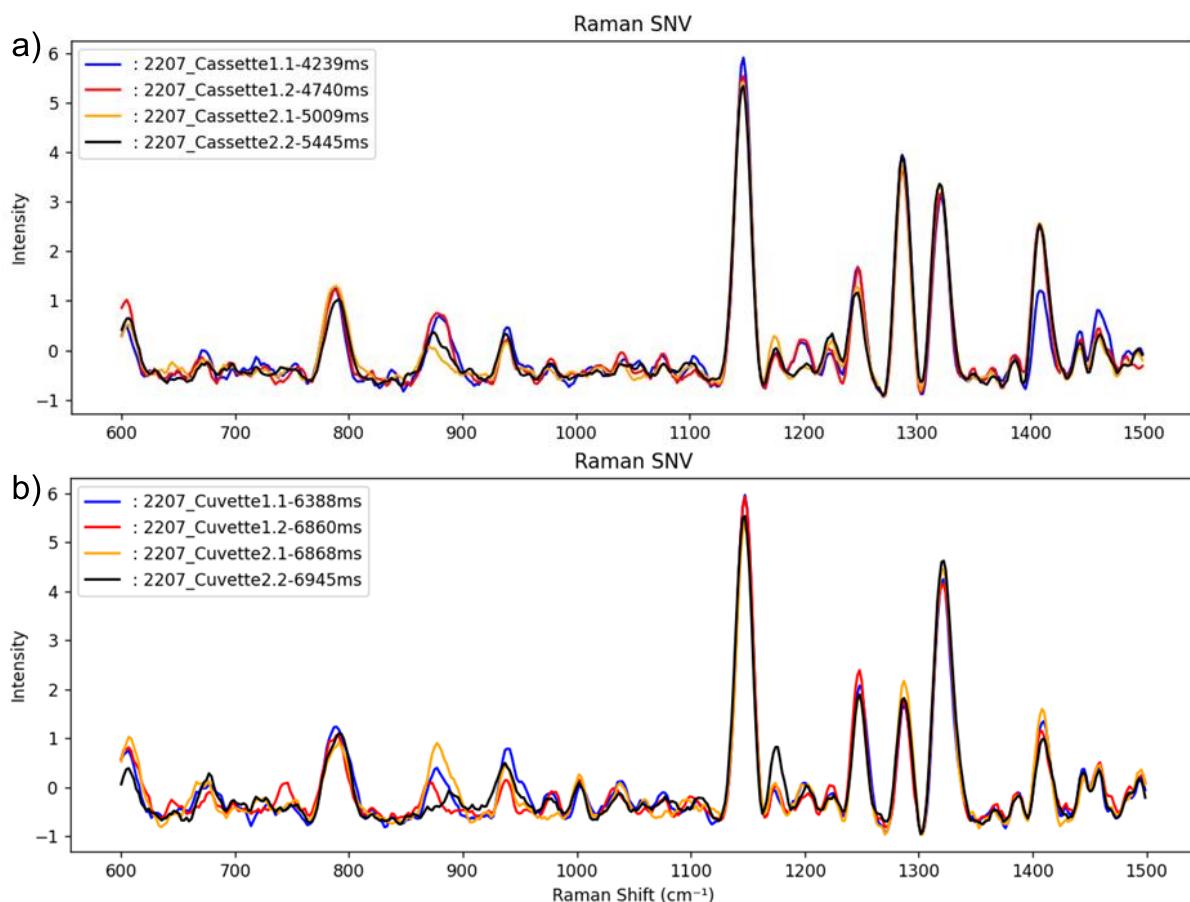


Figure 3.7. Spectres SNV d'aluminium, des dispositifs sans échantillons, de cassette a) et de cuvette b). La légende des figures inclut la date de mesure, le numéro d'identification et le temps d'exposition pour chaque échantillon mesuré et présenté.

Les premières observations apportées par les deux figures précédentes montrent des spectres cohérents, organisés autour de mêmes pics pour chaque figure. Cependant, les quatre spectres mesurés avec la cuvette à vide présentent plus de variations entre eux, constat surtout visible dans la zone autour de 900 cm^{-1} . Cette observation et ces quelques variations peuvent être expliquées par le fait qu'aucun échantillon n'est véritablement mesuré, ce qui rend la mesure naturellement instable. La dernière remarque à notifier ici concerne le temps d'exposition : il est globalement inférieur pour les mesures de cassette par rapport aux mesures de cuvette. Finalement, pour comparer véritablement ces deux dispositions entre elles, on décide maintenant de faire la moyenne pour chaque configuration et de tracer ces deux moyennes sur la même figure :

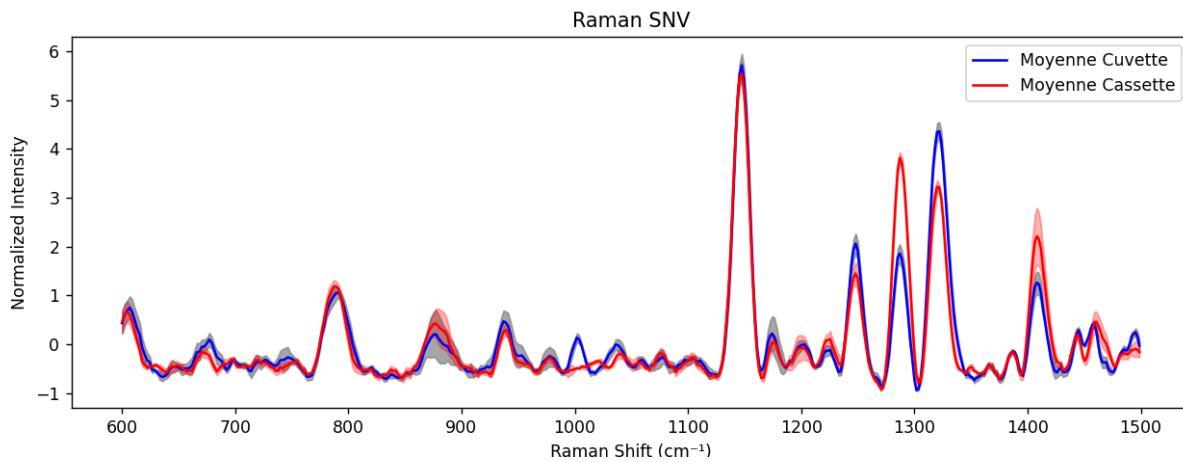


Figure 3.8. Spectres SNV d'aluminium, moyennés sur les données cassette et cuvette. Cette étude de comparaison a pour but de rendre compte des différences entre les deux techniques et de choisir le dispositif final pour les mesures de l'étude « Multi-Virus ».

Cette figure présente en rouge la moyenne des quatre spectres de la figure 3.7 et en bleu la moyenne des quatre spectres de la figure 3.8, ainsi que leurs écart-types respectifs représentés autour de leur valeur moyenne. Visuellement, on observe un pic à 1000 cm^{-1} présent seulement dans le spectre moyen de la cuvette et des intensités différentes en plusieurs endroits du spectre : 1250, 1290, 1320 et 1410 cm^{-1} . Le pic présent à 1000 cm^{-1} est très problématique, car il témoigne d'un matériau sensible à la diffusion Raman. La feuille d'aluminium utilisée sur la cassette est composée uniquement d'aluminium, car l'aluminium ne présente pas de signal Raman. Ce constat a été observé d'après les premières acquisitions du laboratoire, en s'assurant que la feuille d'aluminium présentait un signal neutre. La raison probable pour laquelle la cuvette présente un pic dans le

spectre Raman en plus que la cassette est le matériau dans lequel est creusé le puits : il n'est pas seulement de l'aluminium mais probablement un alliage avec un autre métal.

Ces observations ont permis de conclure que la configuration avec la cassette, composée d'aluminium pur, minimise les interférences dans les spectres Raman des échantillons biologiques, assurant ainsi une meilleure fiabilité des mesures, notamment pour les échantillons de faible intensité spectrale. C'est pourquoi le choix final vers lequel on s'est tourné est la configuration de la cassette. Ce choix est nécessaire pour que nos spectres d'échantillons biologiques ne soient pas pollués par un signal Raman extérieur, bien que faible. C'est d'ailleurs ce qui a été observé par la suite : pour une grosse quantité d'échantillon avec un fort signal Raman comme le sang par exemple, les spectres d'échantillons contenus dans la cassette et la cuvette ne présentent aucune différence. Mais lorsqu'il s'agit d'un spectre de salive, ou d'eau saturée avec un polluant, on remarque des différences spectrales pour un même échantillon.

Plusieurs cassettes avec des trous de différentes tailles ont ainsi été imprimées pour que des puits de différentes tailles soient disponibles en fonction de la quantité de liquide à analyser disponible. Toutes les futures mesures présentes dans cet article seront donc faites avec la cassette.

3.3 Développement des cassettes de calibration

En plus de la conception des cassettes servant à la mesure d'échantillons, il semble intéressant de développer des cassettes propres à l'étape de calibration. Ces cassettes sont différentes car elles ne sont plus valables pour une goutte et doivent prendre en compte la taille et la forme des éléments mesurés lors de cette étape. Comme dit précédemment, la calibration consiste à prendre la mesure de poudre d'acétaminophène, que l'on retrouve couramment dans un cachet de Tylenol et d'un matériau standard, surnommé le NIST. Les principales contraintes imposées comme cahier des charges concernant la conception de ces cassettes sont la robustesse, le système de spectroscopie étant souvent transporté, ces cassettes doivent être solides, et la facilité d'utilisation : en effet, à long terme, le système devrait être utilisable par une personne extérieure au laboratoire et au projet. Pour cette contrainte, il suffit de concevoir une cassette de même forme que les cassettes actuelles, le mécanisme de tiroir se chargeant de rendre ça intuitif. Finalement, pour les dimensions de ces cassettes, on choisit donc de conserver la dimension des cassettes initiales et de former un trou respectant les dimensions du Tylenol et du NIST. La figure 3.9 présente une vue de face de ces cassettes : la cassette pour le Tylenol sur la gauche et la cassette pour le NIST sur la droite.

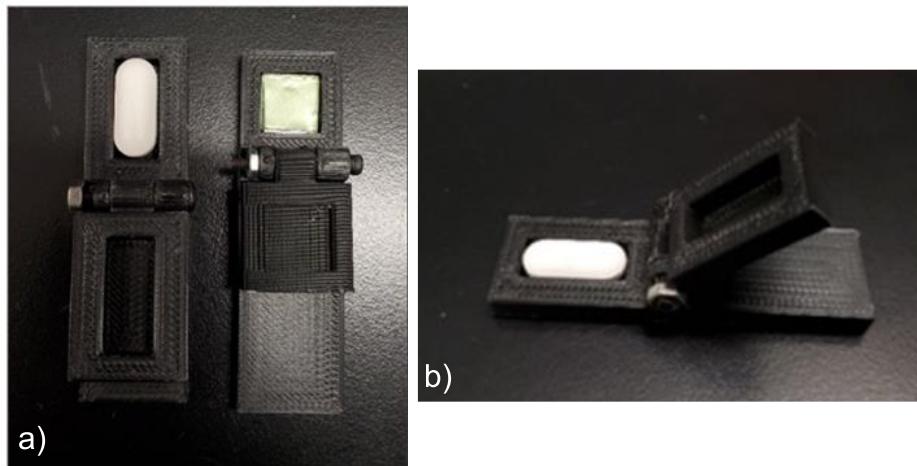


Figure 3.9. Cassettes conçues pour l'étape de calibration. a) Vue de face des cassettes pour la mesure de calibration du Tylenol et du NIST. b) Vue de côté de la cassette du Tylenol, pour illustrer le dispositif de fermeture.

Pour garantir la robustesse de ces cassettes, un système de fermeture a été imaginé avec un couvercle sécuritaire, visible sur la sous-figure b). Ce couvercle est conçu pour être solidement fixé à la cassette à l'aide d'un système vis/écrou, qui permet une ouverture et une fermeture par une simple rotation. Ce choix de mécanisme assure que les cassettes restent solidement fermées lorsqu'elles ne sont pas utilisées, protégeant ainsi les éléments de calibration à l'intérieur (principalement le NIST, qui est un matériau très contrôlé et assez difficile à obtenir) contre les dommages potentiels, comme la poussière, les chocs, ou les manipulations accidentelles. En gardant les cassettes fermées en tout temps, sauf pendant l'étape de calibration, on évite ainsi au maximum les risques de contamination ou de dégradation, pour garantir au mieux la fiabilité et la précision des mesures de calibration, qui déterminent à elles seules le reste des mesures de la journée.

Le processus de conception des cassettes a été réalisé à l'aide du logiciel *Freecad*, permettant de créer facilement et assez rapidement des dessins personnalisés. Une fois les modèles finalisés, les cassettes ont été imprimées à l'aide d'une imprimante 3D disponible dans le laboratoire, très pratique car cela permettait de passer rapidement de la phase de conception sur ordinateur à la phase de test des prototypes, en un jour ou occasionnellement plusieurs jours si le système était

achalandé. Plusieurs essais et configurations différentes ont été nécessaires pour arriver à ce résultat, notamment en raison du choix du mécanisme de fermeture. Finalement, cette version finale des cassettes de calibration semble optimale pour répondre aux besoins formulés initialement. En effet, elles sont maintenant en service depuis un certain temps, et, jusqu'à présent, elles fonctionnent sans qu'il ait été nécessaire d'apporter des modifications, tant bien à leur forme qu'à leur robustesse.

3.4 Étude avec des échantillons de salive

Bien que l'étude principale de ce mémoire porte sur des échantillons de sang, le système a été appréhendé dans un premier temps avec des échantillons de salive. Dans ce contexte, le développement de l'étape de collection de la salive semble pertinent pour faciliter et optimiser le travail préparatoire. Pour ce faire, de nombreux paramètres ont été discutés et remis en question : l'influence de la congélation sur le signal Raman, l'utilisation d'entonnoirs dans l'étape de collecte, le temps d'attente entre l'étape de rinçage et le prélèvement, ou encore une méthode de filtrage en remplacement de l'étape de centrifugation, longue et fastidieuse.

3.4.1 Technique de filtres : méthodes

Pour bien comprendre de quoi il s'agit dans ce chapitre, il convient de détailler maintenant le protocole de collecte et de mesure d'un échantillon de salive.

- L'expérience doit être effectué au moins 30 minutes après avoir mangé ou bu pour la dernière fois
- Retirer tout baume à lèvres ou rouge à lèvres à l'aide de lingettes démaquillantes
- Préparation du système : mesures de saturation et de calibration
- Déposer une feuille d'aluminium sur la cassette et former un puits
- Se rincer la bouche 3 fois avec de l'eau en bouteille et attendre de 5 à 10 minutes sans manger ni boire.
- Placer un entonnoir en amont de la seringue
- Recueillir la salive dans l'entonnoir. Une quantité d'au moins 0,2 ml doit être recueillie.
- Fixer le filtre en aval de la seringue.
- Presser sur le filtre et déposer la première goutte filtrée dans la cassette, directement à partir de la seringue.

- Prendre la mesure avec le système Raman - paramètres : puissance 1W / Gamme dynamique 90% / Temps d'exposition automatique 100-3000ms / 50 acquisitions pour chaque échantillon.
- Notez l'heure de la mesure, et des commentaires sur la mesure
- Changer la feuille d'aluminium pour une nouvelle mesure.

L'expérience explicitée dans cette partie concerne la nouvelle méthode de filtrage. Il a été prouvé précédemment dans le laboratoire que remplacer l'étape de centrifugation par une étape de filtration donnait des résultats satisfaisants, mais l'équipe cherche à obtenir des détails et optimiser cette étape. L'objectif de cette expérience est alors de déterminer le meilleur filtre, quantitativement (avec le signal Raman) et qualitativement (difficulté à faire tomber une goutte), parmi une liste de quatre filtres trouvés dans le commerce, présentant la même taille de pores, soit $0.45\mu\text{m}$, mais fabriqués avec des matériaux différents. La liste des filtres disponibles est disponible dans le tableau 3.1. Le protocole suivi pour cette expérience présentera des spécificités propres aux filtres. En effet, la collecte de salive fonctionne pour seulement deux filtres, ensuite le participant n'a en général plus assez de salive. Il faut alors rincer de nouveau à l'eau et attendre au moins 5 minutes pour laisser le temps à la salive de se reformer. De plus, il est décidé d'alterner l'ordre d'utilisation des filtres (ne pas commencer toujours par le même), pour réduire cette incidence sur les résultats. L'étude implique sept participants, recrutés sur la base du volontariat parmi l'équipe du LUMEDLAB, et s'est tenue sur une seule journée pour minimiser les variables externes et garantir une plus grande stabilité dans les conditions expérimentales.

Tableau 3.1. Liste des filtres utilisés dans le cadre de l'étude d'optimisation du protocole de mesure d'échantillons de salive. Les filtres possèdent une taille de pores identique, mais sont fabriqués à partir de matériaux différents.

Numéro	Type de filtre
Filtre n°1	Cytiva Whatman™ Uniflo Syringe Filters, $0.45\mu\text{m}$ PES
Filtre n°2	Cytiva Whatman™ Uniflo Syringe Filters, $0.45\mu\text{m}$ PVDF
Filtre n°3	Cytiva Whatman™ Uniflo Syringe Filters, $0.45\mu\text{m}$ PTFE
Filtre n°4	Millex® Nylon syringe filter, $0.45\mu\text{m}$

Enfin, beaucoup moins contraignant que la salive réelle, le modèle de salive permet de mener une approche initiale de cette expérience pour appréhender les premiers résultats sans difficulté particulière. Le modèle de salive utilisé dans le laboratoire se compose de divers ingrédients pour imiter la composition naturelle de la salive, notamment de l'eau en importante quantité mais aussi du chlorure de sodium, du phosphate de potassium, de l'albumine sérique humaine, et de petites quantités de glucose, d'urée, et de cholestérol. De ce fait, le modèle, particulièrement liquide, se trouve extrêmement facile à manipuler à travers les filtres, ne représentant finalement pas fidèlement la réalité, la partie qualitative de l'expérience est alors mise de côté. Finalement, aucune différence spectrale n'est observée en comparant les spectres normalisés de modèle de salive à travers les différents filtres, ce qui permet de vérifier que les différents matériaux des filtres ne présentent pas de signal Raman supplémentaire : il faut maintenant confirmer ce résultat dans des conditions réelles et discuter du critère qualitatif.

3.4.2 Technique de filtres : résultats

Tout d'abord, il est facile d'évaluer qualitativement les filtres utilisés en fonction des impressions de l'expérience, en prenant en note de façon détaillée l'étape de poussée sur le filtre pour obtenir une goutte mesurable. Il en ressort des différences notables en termes de praticité et de force de poussée, les filtres n°3 et n°4 montrant des performances insuffisantes. Lors de la collecte, il a été nécessaire d'exercer une pression trop importante sur la seringue pour obtenir une goutte de salive : cette contrainte rend l'utilisation de ces filtres moins pratique et peu adaptée à un protocole où la fluidité de la manipulation est essentielle.

Ensuite, pour évaluer le critère quantitatif, les moyennes des spectres SNV pour chaque filtre sur un échantillon de sept participants ont été tracées sur une même figure. Les résultats, visibles sur la figure 3.10, ont été jugés satisfaisants, ne montrant pas de pics supplémentaires et donc pas de différences majeures entre les filtres. En accord avec le modèle de salive, cela implique que les filtres utilisés sont en mesure de fournir des spectres répétables, indépendamment des participants.

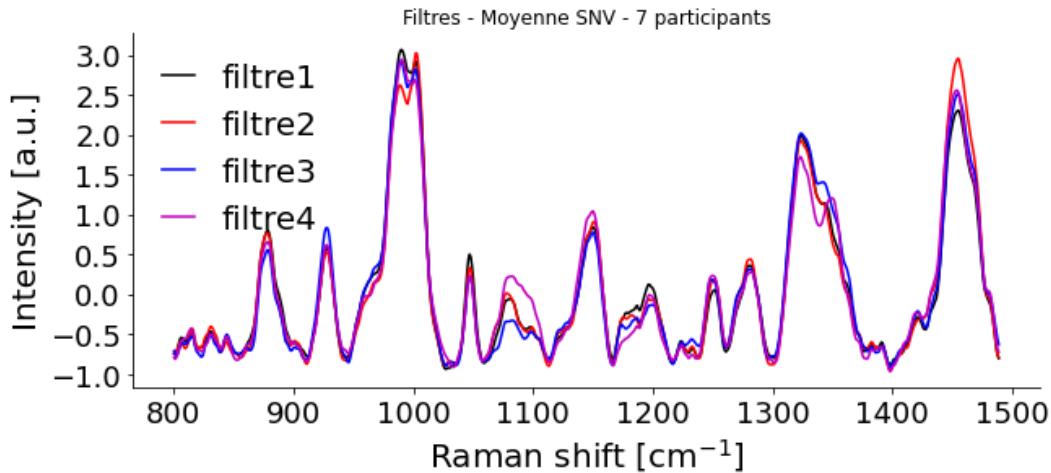


Figure 3.10. Spectres SNV de salive, moyennés sur 7 participants. Les quatre filtres étudiés ont été utilisés par tous les participants à la suite d'une collection d'échantillons salivaires.

Cependant, une analyse plus approfondie des spectres de façon individuelle a révélé des variations importantes. En effet, pour 3 des 7 participants, les spectres obtenus montraient des différences spectrales, alors même que la salive provenait du même échantillon, de gouttes distinctes néanmoins. Par exemple, la figure 3.11 présente deux spectres Raman du même échantillon : de vraies différences sont visibles sur toute l'étendue spectrale, mais notamment autour de 970 cm^{-1} et 1100 cm^{-1} , mettant en lumière la variabilité importante, même pour un même individu, des échantillons de salive.

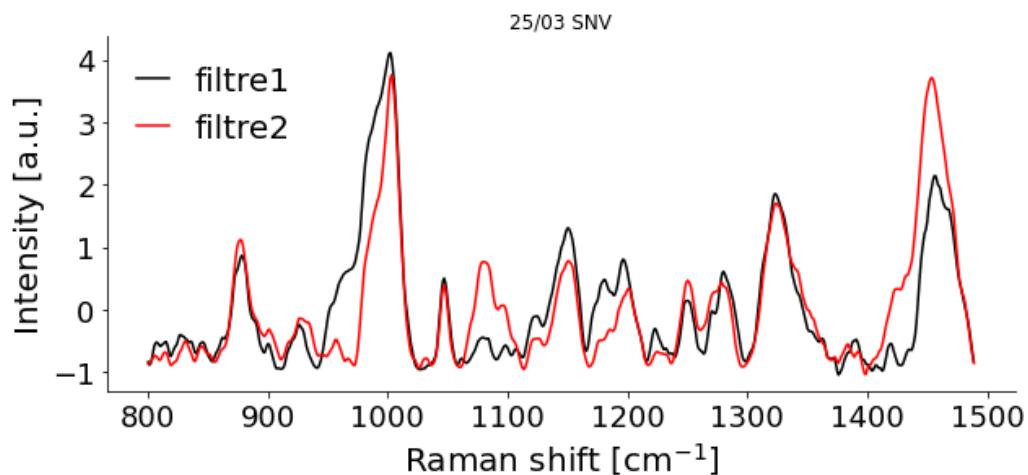


Figure 3.11. Spectres SNV de salive, d'un seul participant, avec deux filtres différents. Ces mesures illustrent l'impact des variations inter-individuelles sur les échantillons de salive.

Ces observations soulignent l'importance de la précision et la rigueur du protocole dans la manipulation d'échantillons de type salivaire, et bien que la comparaison de spectres individuels puisse amener des considérations pertinentes, une moyenne sur plusieurs spectres sera toujours privilégiée pour réduire au maximum ces variabilités intra-individuelles. Pour remettre en cause le protocole utilisé, il est important de noter que le temps d'attente entre le rinçage et la collecte doit impérativement dépasser 5 minutes, sans quoi la salive ne peut se trouver sous sa forme optimale.

En conclusion, l'étude qualitative au niveau protocolaire ayant éliminé les filtres n°3 et n°4, l'analyse comparative des filtres n°1 et n°2 n'a pas montré de différences quantitatives significatives entre les deux. Ainsi, les deux filtres peuvent être utilisés indifféremment pour la suite des expériences. Toutefois, en prenant en compte le coût légèrement inférieur du filtre n°1, ce dernier a été choisi pour les expérimentations futures.

3.4.3 Expériences préparatoires : influence de la congélation, entonnoirs

Les résultats pertinents obtenus à partir des expériences parallèles réalisées avec de la salive sont discutés expressément dans cette partie, et les avancées concernant les protocoles de mesure ou de collection sont mises en avant.

Premièrement, il a été donné d'évaluer l'influence de la congélation sur le spectre Raman : en effet, pour la bonne conservation des échantillons, il faut être certain que la congélation à -80°C pendant un certain temps n'exerce aucune influence sur le spectre Raman. Pour cela, l'expérience, incluant douze participants, a demandé de la salive neutre et de la salive saturée en chips de maïs pour discuter de la différence entre de faibles et forts signaux Raman. Il en ressort que les spectres de salive neutre sont trop bruités et inexploitables : les expériences suivantes auront alors pour principal objectif d'acquérir des spectres Raman de salive exploitables. Cependant, l'expérience est judicieuse avec les spectres de salivée en chips de maïs : aucune différence remarquable n'est remarquée sous l'influence de congélation, ce qui très encourageant pour la conservation facile d'échantillons salivaires, et donc dans un cadre de dépistage à grande échelle pour de futures évolutions du projet.

En remettant en perspective l'expérience précédente, le premier élément à remettre en cause est le temps d'attente entre le rinçage et le moment de la collection de salive. L'objectif est de déterminer ce temps d'attente nécessaire pour arriver à un niveau de bruit satisfaisant sur le signal Raman. En prenant différents temps d'attente au fur et à mesure, la conclusion est que les meilleurs résultats sont obtenus après 10 minutes d'attente, avec une stagnation de la qualité du signal ensuite, mais des résultats satisfaisants fonctionnent à partir de 5 minutes. Ce critère est alors intégré au protocole de collecte de salive. On remarque également avec cette expérience que les pics Raman sont plus prononcés lorsque la mesure est effectuée sur la première goutte filtrée, plutôt que les gouttes suivantes, ce qui semble cohérent, la première goutte étant sensiblement plus concentrée en termes de biomarqueurs. Finalement, les paramètres du système ont également été remis en question, il en ressort simplement qu'une puissance de 1W au niveau de l'échantillon, plutôt qu'1.5W auparavant n'entraîne pas d'effets négatifs, seulement une légère augmentation du temps d'exposition pour un signal de qualité équivalente donc cette puissance est maintenant utilisée.

Ensuite, en revenant sur les ajustements du protocole, un point d'interrogation portait sur l'impact potentiel de l'utilisation d'un entonnoir sur le signal Raman. L'idée était d'examiner si l'ajout de cet élément, destiné à faciliter la collecte de la salive, pourrait altérer la qualité des spectres obtenus. Plusieurs essais ont été menés, avec des échantillons de salive filtrée à travers l'entonnoir comparés à ceux sans. Les spectres obtenus ont été comparés à chaque fois pour chaque participant, comme dans la figure 3.12, ne révélant aucune différence notable, quelle que soit la configuration utilisée.

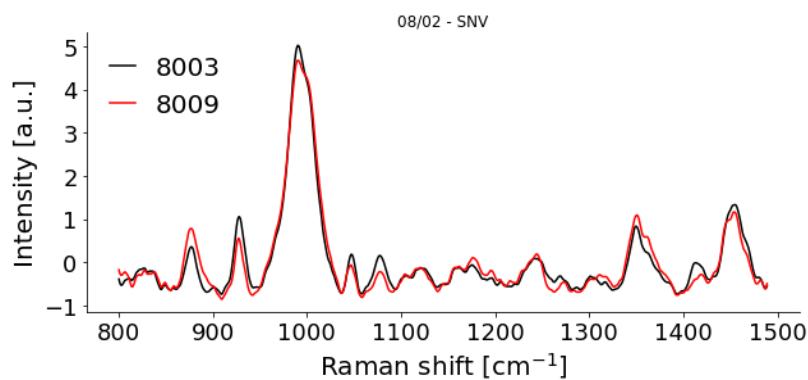


Figure 3.12. Spectres SNV de salive, avec et sans entonnoir, pour un participant. La courbe noire issue de l'échantillon 8003, représente un échantillon collecté sans entonnoir, tandis que la courbe rouge correspond à l'échantillon 8009, collecté à l'aide d'un entonnoir. Ce résultat démontre l'absence de signal Raman supplémentaire à l'ajout d'un entonnoir dans le protocole de collecte d'échantillons de salive.

De plus, en complément des résultats techniques, l'utilisation de l'entonnoir dans le protocole de collecte a également suscité des retours positifs de la part des participants, qui ont noté que l'entonnoir rend la collecte de salive beaucoup plus simple. Cet effort de commodité durant la collecte est particulièrement précieux dans de potentiels futures campagnes de dépistage, où le confort des participants et l'efficacité du processus jouent un rôle clé, notamment dans des environnements à grande échelle ou dans des situations où la collecte d'échantillons doit être réalisée rapidement. L'entonnoir, très pratique pour les participants et sans effet mesurable sur le signal Raman, a donc été intégré définitivement dans le protocole de collecte d'échantillons salivaires.

CHAPITRE 4 ÉTUDE PRINCIPALE « MULTI-VIRUS »

Ce chapitre est le corps principal de ce mémoire, qui va rendre compte de l'étude principale de la maîtrise. Cette étude s'intéresse à la détection de maladies virales (COVID-19, RSV, Influenza) par spectroscopie Raman dans des échantillons sanguins sur une population pédiatrique.

L'hypothèse de travail est définie comme telle : la spectroscopie Raman permet de détecter des différences distinctives dans les spectres d'échantillons sanguins entre des patients atteints de maladies virales spécifiques (COVID-19, RSV, Influenza) et des individus sains. Ces différences spectrales reflètent de variations dans les biomarqueurs moléculaires associés à chaque maladie virale.

L'objectif est d'étudier les spectres Raman d'une population pédiatrique positive à certaines maladies virales. Cet objectif principal est découpé en objectifs intermédiaires, qui vont servir de jalons dans le bon déroulement de l'étude :

- Mener des expériences préliminaires pour évaluer la réponse du système Raman à des échantillons de sang.
- Acquérir les spectres Raman associés à des échantillons d'une biobanque conséquente.
- Identifier et quantifier des différences spectrales entre les échantillons issus de patients positifs à des maladies virales et les échantillons de contrôle.

4.1 Étude préliminaire – Répétabilité des mesures d'échantillons sanguins

Tout d'abord, pour obtenir des spectres cohérents entre les nombreux échantillons disponibles, il est important de vérifier que les mesures faites avec le système sont bien répétables avec des échantillons de sang. En effet, il s'agit de la première étude, effectuée avec une grande quantité d'échantillons sanguins provenant d'une biobanque, du système développé par le LUMEDLAB. L'objectif de cette étude préliminaire est donc simple, il s'agit de vérifier qu'il est possible de mesurer des échantillons de sang de manière reproductible avec ce système.

4.1.1 Méthodes

Au niveau du protocole, les échantillons utilisés pour cette étude sont des échantillons sanguins de rat : on a accès à deux rats pour lesquels sont disponibles trois fioles de plasma et trois fioles de sérum chacun. Chacune des fioles possède un volume correspondant à environ une centaine de microlitres : en prenant un volume de $30\mu\text{L}$ pour chaque échantillon, on pourra faire plusieurs mesures, de 2 à 6, pour chaque fiole. En utilisant la cassette de volume théorique de $60\mu\text{L}$ et un volume de goutte de seulement $30\mu\text{L}$, la goutte ne remplit pas entièrement le puits donc il faut faire attention ici à bien positionner la goutte au centre du puits. La figure 4.1 montre une illustration de la goutte à l'intérieur du puits, représentant ce problème de façon précise. Le protocole de nettoyage, qui sera détaillé dans la partie méthodologique suivante, et l'étape de centrifugation sont bien respectés.



Figure 4.1. Illustration d'une goutte de sang déposée à l'intérieur du puits formé à partir de la cassette. Le volume de la goutte étant de $30\mu\text{L}$, et alors que ce puits peut contenir des gouttes allant jusqu'à $60\mu\text{L}$, le liquide doit se trouver au centre du puits pour obtenir une mesure fidèle à cet échantillon.

4.1.2 Résultats

Malgré l'utilisation d'une goutte de $30\mu\text{L}$, les mesures obtenues sont de bonne qualité avec un faible niveau de bruit. Les temps d'exposition restent constants et sont conformes aux valeurs habituellement observées avec ce système. Finalement, on obtient dix-huit mesures d'échantillons de plasma et treize mesures de sérum. On représente certains de ces spectres, ceux du deuxième rat, dans la figure 4.2 a) pour le plasma et dans la figure 4.2 b) pour le sérum.

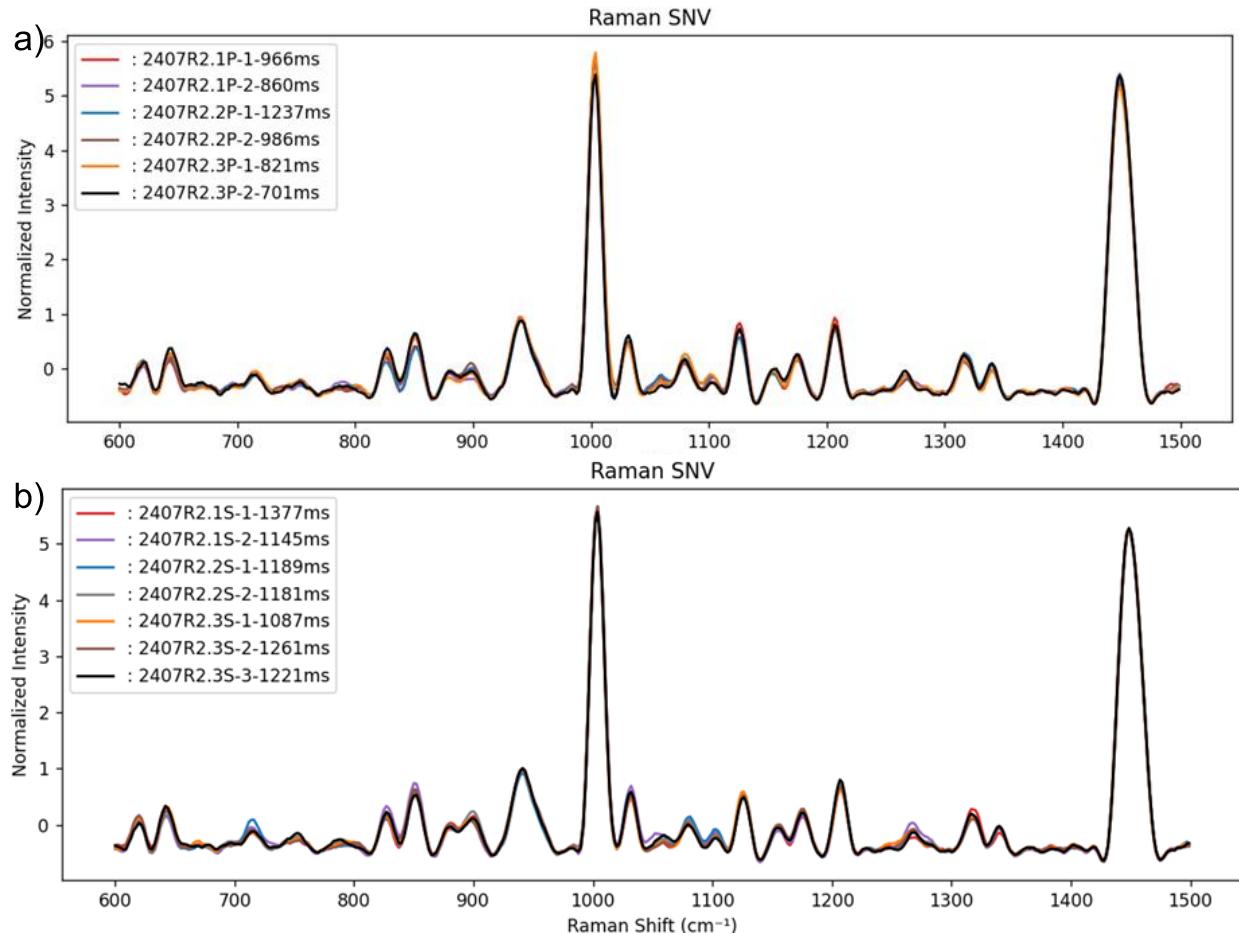


Figure 4.2. Spectres SNV de sang, provenant de différents échantillons de rat. La sous-figure a) représente les spectres associés aux échantillons de plasma alors que la sous-figure b) montre ceux des échantillons de sérum. La légende des figures inclut la date de mesure, le numéro d'identification et le temps d'exposition pour chaque échantillon mesuré et présenté.

Les graphes montrent que les spectres obtenus présentent une forte similarité, sans apparition de pics inhabituels ou incohérents : ces résultats indiquent que les spectres, bien que mesurés sur des volumes d'échantillon réduits à 30 µL, sont homogènes et exempts de signaux aberrants. La répétabilité des mesures est ainsi confirmée, tant pour les échantillons de plasma que pour ceux de sérum. Ces observations valident non seulement la méthode employée, mais aussi la performance du système et la pertinence du protocole de mesure mis en place. Cela ouvre ainsi la voie à des analyses plus approfondies sur des échantillons sanguins, avec la garantie d'une méthodologie robuste et reproductible.

4.1.3 Discussion

Pour obtenir des spectres Raman ressemblant à ceux pouvant être observés dans la littérature scientifique, on décide de modifier plusieurs éléments du code de pré-traitement des spectres. En particulier, il est préférable pour des spectres de sang de se rendre jusqu'à 1700 cm^{-1} , alors qu'on utilisait auparavant une plage allant de 600 à 1500 cm^{-1} pour l'écart en longueur d'onde, qui se trouvait être pertinente auparavant pour l'étude de spectres de salive, car cette plage permettait de supprimer le pic de l'eau (évidemment très présent dans la salive), présent dans la région spectrale autour de 1600 cm^{-1} . De plus, on va changer de valeur pour la taille de la *bubble_width*, paramètre nécessaire dans l'algorithme utilisé pour supprimer le signal de fluorescence (développé de façon plus détaillée dans la bibliographie), et prendre deux valeurs de *bubble_width* pour traiter le spectre différemment en fonction du pic de l'eau. Voilà l'impact de ces changements, avec pour le premier spectre, un *bubble_width* de taille 20, puis sur le second spectre, un *bubble_width* de taille 60 avant 1400 cm^{-1} et de taille 120 après cette longueur d'onde.

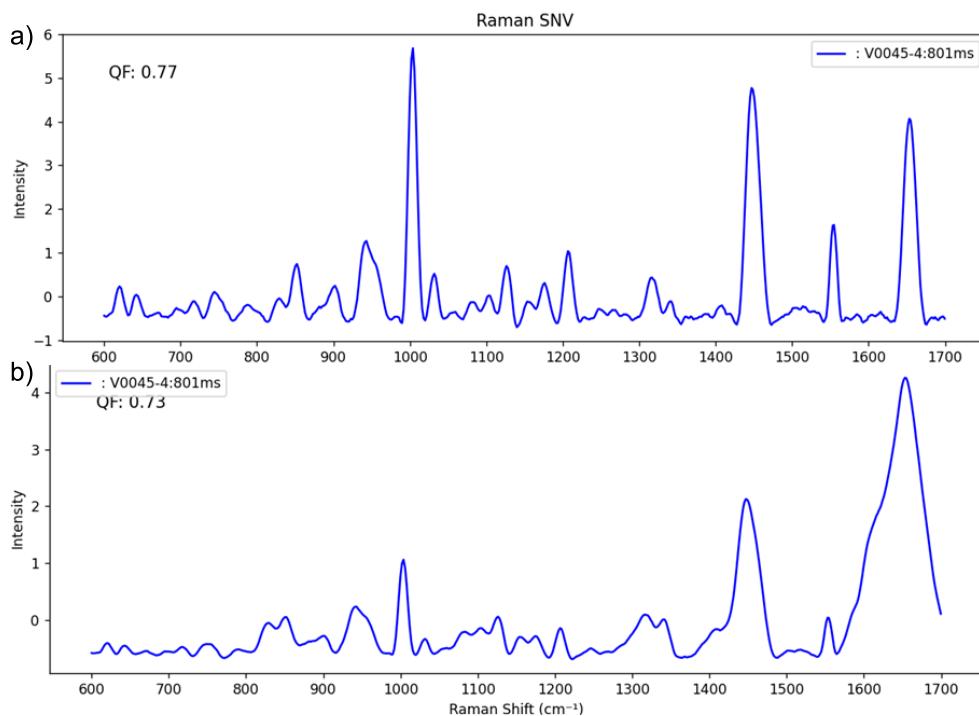


Figure 4.3. Spectres SNV de sang, provenant de la même mesure et utilisés pour discuter de l'influence d'un paramètre du code de traitement. a) Ce spectre Raman, obtenu avec les paramètres utilisés jusqu'à précédent, met l'accent sur l'intensité des pics. b) Ce spectre Raman met davantage l'accent sur la forme générale du spectre, correspondant ainsi de plus fidèlement aux spectres Raman déjà obtenus dans le laboratoire et dans la littérature.

Les spectres précédents sont tellement différents qu'il faut préciser qu'ils proviennent bien tous deux de la même mesure du même échantillon. Les pics du second spectre sont les mêmes que ceux du premier spectre, bien présents pour les mêmes longueurs d'onde, mais ils sont visuellement moins marqués et plus larges. Finalement ce processus de traitement spectral a été validé et sera systématiquement employé dorénavant, et ce jusqu'à la fin de mon travail. Maintenant, on décide de tracer sur une même figure la moyenne accompagnée de l'écart-type des échantillons de sérum et de plasma pour un unique rat, dans le but de valider le résultat de cette étude préliminaire, puis d'évaluer dans un second temps la différence spectrale entre ces deux biofluides.

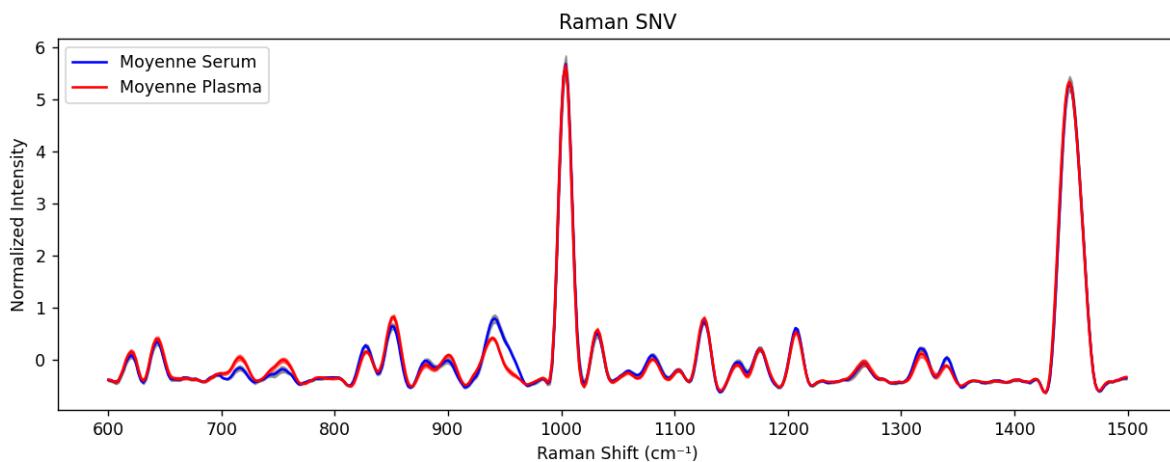


Figure 4.4. Spectres SNV de sang, moyennés sur des mesures d'échantillons provenant d'un seul rat. La courbe bleue représente la moyenne des échantillons de sérum et la courbe rouge montre la moyenne des échantillons de plasma. Ce résultat souligne le faible impact des variations inter-individuelles pour les échantillons sanguins, ainsi que l'absence de différences significatives entre des échantillons de plasma et de sérum.

En conclusion, cette étude préliminaire montre bien que les mesures effectuées sur des échantillons de sang sont correctement menées, et mènent à des résultats satisfaisants, malgré l'utilisation de petits volumes. En effet, sur la figure 4.4, les écart-types sont pratiquement indistinguables, ce qui prouve bien que le protocole et les mesures sont répétables, les méthodes pertinentes et précises, et les spectres prêts à être analysés. De plus, cette stabilité de résultats des échantillons de sang contraste avec ceux provenant des échantillons de salive, qui montraient bien davantage de variations inter-individuelles, déjà discutées à partir de la figure 3.11. Le dernier résultat pouvant être discuté ici concerne la différence entre plasma et sérum : spectralement, cette différence est

négligeable sur tout la région spectrale étudiée, excepté le phénomène observé autour de 950 cm^{-1} dans la figure précédente, mais ce dernier n'a pas été observé sur des échantillons provenant d'humains, ce qui permet de conclure à une similarité des spectres et indique que ces deux types d'échantillons pourront être utilisés de manière coordonnée dans nos études ultérieures.

4.2 Méthodes – Étude principale : « Multi-Virus »

Cette partie va détailler les méthodes choisies pour traiter ces échantillons, en particulier un peu de contexte, des détails sur l'étape de nettoyage puis bien sûr le protocole spécialement mis en place pour cette étude.

4.2.1 Contexte et données

Contextuellement, une étude a été réalisée en collaboration avec un laboratoire du centre de recherche du Centre Hospitalier Universitaire de Ste-Justine (CHUSJ), situé au cœur d'un des principaux hôpitaux pédiatriques de Montréal. Des discussions régulières ont eu lieu pour l'organisation de l'étude, il a fallu coordonner nos disponibilités avec des étudiants d'un autre laboratoire, et enfin, il a fallu déplacer le système et tout le matériel requis pour les mesures. Des approbations éthiques ont aussi été nécessaires pour la tenue de cette étude. Cette section de ma maîtrise a été réalisée sous la supervision de mon co-directeur, Mathieu Dehaes.

Les échantillons de plasma et sérum proviennent tous d'une population pédiatrique, pour des enfants allant de quelques jours à 17 ans. Ils ont été prélevés de janvier à juin 2024, lors de consultations aux urgences de l'hôpital, centrifugés puis stockés dans un congélateur à -80°C , température indispensable pour conserver correctement les échantillons biologiques. Lorsque le patient restait quelques jours, plusieurs échantillons ont pu être prélevés, ce qui est intéressant pour mettre en place des études longitudinales, c'est-à-dire des études permettant d'analyser l'évolution de la maladie au fil du temps, à l'inverse des études transverses, qui examinent la maladie à un moment précis. Les échantillons de contrôle, c'est-à-dire ceux notés comme « Négatif ciblé, ou *Healthy* » dans ce rapport, sont issus de patients testés négativement aux trois maladies de l'étude mais ces patients ne souffrent pas pour autant forcément d'aucune maladie ; cette situation sera discutée prochainement dans une autre partie.

Les mesures de ces échantillons ont été prises dans une salle de laboratoire de niveau de confinement 2 du centre de recherche du CHUSJ. Par mesure de sécurité, les étudiants devaient

porter une blouse, un masque ainsi que des gants lors des prises de mesures. Les échantillons avaient comme contrainte de devoir être manipulés en tout temps sous une hotte. Or le système de spectroscopie ne rentre pas en entier sous la hotte, il a donc fallu l'installer sur un chariot près de la hotte et installer seulement la boîte « lecteur » sous la hotte. Ce dispositif est représenté sur la photographie de la sous-figure 4.5 a).

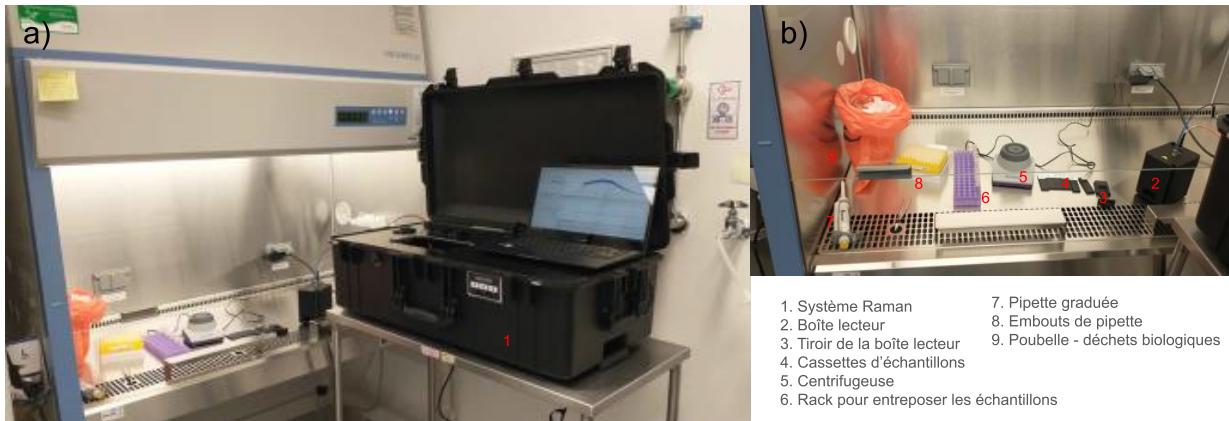


Figure 4.5. Système Raman dans le cas de mesures en conditions réelles. a) Le système est situé sur un chariot mobile, tandis que la boîte lecteur se trouve sous une hotte, essentielle dans le cadre de manipulation d'échantillons biologiques. b) Les éléments nécessaires aux manipulations que l'on retrouve sous la hotte sont illustrés ici. De plus, les éléments importants sont indiqués et détaillés à l'aide d'une numérotation.

Sous la hotte, on retrouve tout le matériel nécessaire pour manipuler des échantillons biologiques, bien présent sur l'illustration de la sous-figure b) précédente : un vortex utilisé pour centrifuger les échantillons juste avant la mesure, une micropipette et ses embouts, un sac de déchets biologiques, un porte-échantillons, les cassettes et enfin la boîte « lecteur ».

Les données sont présentées dans le tableau 4.1, incluant 284 échantillons provenant de 168 participants différents. On choisit plusieurs catégories pour classer ces données. Premièrement, on note la maladie à laquelle chaque participant est associé et pour laquelle il a été testé positivement, puis on classe les données en fonction du type d'échantillon sanguin, plasma ou sérum, la quantité disponible pour la mesure ou encore le sexe du participant. Enfin, on ajoute deux autres catégories pertinentes pour de prochaines comparaisons : un classement par groupe d'âge, et la raison de l'admission du participant aux urgences de l'hôpital.

Tableau 4.1. Données de l'étude « Multi-Virus ».

	Covid	RSV	Influenza	Négatif ciblé	Total
Nombre total de participants	9	21	53	85	168
Groupes d'âge, n (%)					
0-1 mois	3 (34)	3 (14)	3 (6)	9 (11)	18 (10)
2-6 mois	1 (11)	7 (33)	2 (4)	9 (11)	19 (11)
7-12 mois	1 (11)	3 (14)	-	7 (8)	11 (7)
1-3 ans	1 (11)	6 (29)	13 (24)	22 (26)	42 (25)
4-6 ans	2 (22)	1 (5)	17 (32)	11 (13)	31 (18)
7-10 ans	1 (11)	1 (5)	11 (21)	12 (15)	25 (15)
11-13 ans	-	-	4 (7)	7 (8)	11 (7)
14-17 ans	-	-	3 (6)	8 (9)	11 (7)
Type de sang, n (%)					
Sérum	9	19 (90)	47 (89)	77 (91)	152 (91)
Plasma	-	2 (10)	4 (7)	8 (9)	14 (8)
Inconnu	-	-	2 (4)	-	2 (1)
Quantité, n (%)					
60µL	8 (89)	20 (95)	53	76 (89)	157 (93)
Moins de 60µL	1 (11)	1 (5)	-	9 (11)	11 (7)
Sexe, n (%)					
F	5 (56)	10 (48)	22 (42)	37 (44)	74 (44)
M	4 (44)	11 (52)	31 (58)	48 (56)	94 (56)
Raison de l'admission					
Fièvre	3	1	12	18	34
Pneumonie	1	3	6	29	39
Bronchiolite	1	9	-	9	19
Toux	-	1	-	1	2
Asthme	-	2	-	1	3
IVRS	-	1	1	4	6
AFF	-	1	3	7	11
Bronchopneumonie	-	1	2	3	6
Neutropénie (fébrile)	-	1	2	2	5
Influenza	-	-	16	-	16
Température	-	-	1	1	2
Status asthmaticus	-	-	1	1	2
Fatigue	-	-	2	1	3
Myosite	-	-	5	-	5
Virémie	-	-	3	5	8
Laryngite	-	-	1	2	3
Bronchospasme	-	-	-	3	3
Virose	-	-	-	2	2
Symptômes respiratoires	-	-	-	2	2
Autres	5	2	11	15	33

4.2.2 Protocole expérimental

Pour mener à bien cette étude, les mesures ont été effectués par trois étudiants en recherche : une personne responsable de la manipulation des échantillons sous la hotte, une personne responsable du système et de la prise de mesure, moi-même, et une dernière personne responsable de la phase de nettoyage. Ces trois étudiants sont une doctorante de l'équipe du laboratoire du centre de recherche du CHUSJ, une stagiaire du LUMEDLAB en deuxième année de ses études d'ingénierie et moi-même.

Pour cette étude, l'étape de nettoyage de la cassette avant chaque mesure s'est trouvée être cruciale et très chronophage, si bien qu'une personne était spécialement assignée à cette tâche. Pour obtenir des résultats identiques à cette étude, il convient de détailler le protocole de nettoyage. Il est disponible au complet en annexe mais on va rappeler les grandes étapes ici : une nouvelle feuille d'aluminium est déposée sur la cassette et un puits est formé délicatement pour ne pas trouer l'aluminium. Ensuite, la feuille d'aluminium est nettoyée avec de l'éthanol, de l'eau puis séchée à l'aide d'air comprimé, plusieurs fois.

Enfin, il est temps de s'intéresser en détails au protocole expérimental ayant permis d'acquérir les spectres Raman de cette étude. Premièrement, il faut commencer par l'étape de préparation du système de spectroscopie (mise en route, saturation, calibration), qui peut être faite en parallèle de l'étape de décongélation progressive des échantillons sanguins. Cette étape est cruciale pour préserver l'intégrité des échantillons et garantir la qualité des analyses ultérieures. En évitant une décongélation trop rapide, on évite la formation de cristaux de glace qui pourraient endommager les cellules sanguines et altérer leurs liaisons moléculaires, compromettant ainsi la fiabilité des résultats. De plus, cette décongélation progressive permet de maintenir la stabilité des biomarqueurs, qui peuvent être sensibles aux variations de température soudaines. Enfin, elle favorise une homogénéisation liquide de l'échantillon, qui facilite l'étape suivante de centrifugation. Cette décongélation est réalisée d'abord en laissant les échantillons un certain temps dans de la glace sèche, puis dans de la glace pilée et enfin à l'air libre.

Puis s'en suit l'étape de nettoyage de la cassette et celle de préparation de l'échantillon sous la hotte, constituée d'un moment de centrifugation, nécessaire pour homogénéiser le liquide à analyser. C'est le moment de déposer la goutte de $60\mu\text{L}$ de sang dans la cassette, le plus délicatement possible. Ensuite, la mesure est prise avec les paramètres suivants : puissance laser

1.31W à l'échantillon (soit un *Power Digit* de 80), pourcentage de la plage dynamique désiré 85%, temps d'exposition automatique compris entre 100 et 3000 ms, 50 acquisitions par échantillon. De nombreuses informations sont ensuite collectées sur l'échantillon, telles que l'heure de l'acquisition, la couleur prédominante, la quantité de liquide restant, et autres commentaires qualitatifs. L'étape de nettoyage et le protocole expérimental détaillés sont disponibles en annexe.

Voici le protocole détaillé de l'étude, comme elle a été réalisée, les personnes responsables de chaque étape sont précisées, pour que l'enchaînement des tâches puisse être reproduit si besoin :

- Mettre une nouvelle feuille d'aluminium sur la cassette et utiliser délicatement le tournevis pour créer un puits au-dessus du trou, puis suivre le protocole de nettoyage et déposer la cassette sous la hotte quand elle est propre. (Coline)
- Mixer l'échantillon de sang avec un vortex pendant 10 à 20 secondes. (Maryam)
- Utiliser la pipette pour placer une goutte de 60 μ L (si disponible) dans le puits. (Maryam)
- Mettre la cassette dans la boîte « lecteur ». (Maryam)
- Prendre la mesure de l'échantillon à partir du système. (Antoine)
- Remplir le fichier avec les informations de l'échantillon. (Antoine)
- Retirer la feuille d'aluminium, la jeter dans la poubelle réservée aux déchets biologiques puis déposer la cassette vide pour l'étape de nettoyage. (Maryam)

Un total de 284 échantillons a été mesuré en suivant cette méthode, puis ces données ont été analysées et seront présentées dans la partie des résultats qui suit.

4.3 Résultats – Étude principale : « Multi-Virus »

Les résultats de l'étude sont présentés en détails dans ce chapitre. Nous y exposerons les données obtenues, mais aussi les décisions choisies et les réflexions ayant conduit à ces choix.

4.3.1 Nature exploratoire de la recherche

Pour commencer, il est important de souligner que l'ensemble des 284 échantillons mesurés ne reflète pas fidèlement la réalité : en effet, les données longitudinales sont incluses dans cet ensemble de données, c'est-à-dire que plusieurs échantillons provenant d'une même personne, prélevés à différents moments, sont pris en compte. Cela peut introduire des biais, puisque ces échantillons répétés ne sont pas indépendants les uns des autres et pourraient fausser les conclusions. En ne prenant en compte uniquement le premier échantillon prélevé pour chaque participant, il en résulte la table 4.1 discutée préalablement dans une partie antérieure. Les premières observations apportées à cette table conduisent à un bilan initial sur le nombre d'échantillons disponibles : il y a relativement peu de données concernant les maladies telles que le COVID-19 et le RSV mais il existe une importante quantité d'échantillons disponibles pour l'Influenza, ainsi que de nombreux échantillons de contrôle, qui semblent représenter de manière adéquate la situation.

Cependant, et afin de représenter la diversité des conditions de collecte, tous les échantillons mesurés sont retenus, ce qui inclut à la fois les échantillons de sérum et ceux de plasma, mais aussi tous les échantillons avec une goutte de volume standard de 60 μ L et de volumes inférieurs. Cette situation est survenue en de rares cas, en raison de la faible quantité disponible suite à la collection de l'échantillon auprès du patient. Ce choix de sélectionner le maximum d'échantillons repose sur plusieurs considérations : d'une part, l'utilisation de tous les types d'échantillons permet de simuler les conditions réelles d'une campagne de dépistage où la collecte peut être influencée par des contraintes logistiques ou bien par les spécificités des participants, comme l'âge ou l'état de santé par exemple. De plus, la contrainte de limite de disponibilité pour le volume d'échantillons est représentative de réalités cliniques et de dépistage de terrain, où des volumes de collecte limités peuvent être fréquents, notamment dans les cas de prélèvements pédiatriques ou lorsque l'échantillonnage doit être réalisé rapidement et sans gêne pour le participant. D'autre part, la majorité des échantillons collectés dans la biobanque étant de type sérum et de volume égal à 60 μ L, une première analyse avec ces échantillons garantit une bonne représentativité de la population

étudiée et permet de bénéficier de la plus grande homogénéité possible pour les premières comparaisons entre groupes. Les échantillons de plasma et de quantité inférieure à $60\mu\text{L}$ représentent respectivement moins de 10% de l'ensemble des échantillons : ajouter alors ces quelques échantillons différents permet de conserver des résultats fiables tout en diversifiant les données, ce qui semble mener à une situation puis des résultats concrets et robustes. En somme, la décision de conserver tous les types d'échantillons et volumes répond à un double objectif : optimiser la représentativité des données vis-à-vis de la population cible et anticiper les conditions variées d'utilisation de la méthode en contexte réel, tout en assurant une base de données suffisamment robuste pour évaluer la performance des analyses spectroscopiques dans des situations cliniques diverses.

4.3.2 Premier résultat – Comparaison entre échantillons infectés et échantillons de contrôle

Tout d'abord, il a été décidé de comparer les résultats de façon générale. C'est pourquoi la première figure représente la différence entre les échantillons des patients testés positivement et ceux testés négativement aux maladies virales concernées dans cette étude. Concrètement, les échantillons de la première catégorie proviennent de participants testés positivement à une des trois maladies : COVID-19, RSV ou Influenza. Dans notre jeu de données, seules deux personnes étaient à la fois positives à deux maladies. Ces échantillons spéciaux sont inclus ici dans la comparaison générale mais seront exclus ensuite des comparaisons par maladie. Les échantillons propres à la seconde catégorie quant à eux sont issus de personnes testés négativement à ces maladies virales mais présentant néanmoins des symptômes similaires à des infections virales : cette population d'échantillons servira de contrôle pour tous les résultats de cette étude, permettant de ce fait de conserver cette idée de proposer une étude proche d'une réalité clinique, plus représentative que d'utiliser des échantillons contrôles sans aucune maladie ni symptômes. L'objectif de cette comparaison initiale est de repérer les signatures spectrales communes aux infections, sans chercher à distinguer les spécificités propres à chaque maladie, et de dégager des caractéristiques générales associées aux états infectieux.

Tous les résultats propres à cette partie seront représentés sous la forme de quatre sous-figures, permettant d'approfondir l'analyse de ces comparaisons. La sous-figure a) présente les spectres moyens et écart-types de chacune des catégories, et montre visuellement les tendances générales

des intensités entre les groupes, avec certaines différences apparentes autour de longueurs d'onde spécifiques, sans pour autant révéler les variations significatives. Ensuite, on trace une autre sous-figure afin de mieux visualiser les différences entre les deux spectres, souvent trop subtiles à l'œil nu sur la première figure. La différence entre les moyennes des spectres est exprimée en pourcentage de l'intensité maximale. Cette représentation permet de mettre en évidence les régions spectrales où l'écart entre les groupes est le plus marqué. Enfin, les zones grisées, présentent communément sur la deuxième et la troisième sous-figures, correspondent aux longueurs d'onde pour lesquelles la p-value calculée est inférieure à 0.05, indiquant des différences statistiquement significatives entre les groupes.

La sous-figure c) présente justement les p-values inférieurs à 0.05 et les longueurs d'onde qui leur sont propres. En effet, les p-values sont calculées en utilisant un test de Student, adapté dans ce cas car utile pour déterminer si les moyennes de deux groupes sont significativement différentes l'une de l'autre. Dans ce contexte, l'hypothèse nulle H_0 du test de Student est établie de cette manière : il n'existe pas de différence significative entre les intensités spectrales des deux groupes pour une longueur d'onde donnée. Pour chaque longueur d'onde, la p-value est calculée en fonction de l'écart entre les moyennes des intensités, tout en prenant en compte la variance et la taille des groupes d'échantillons. Concrètement, elle indique la probabilité de respecter l'hypothèse nulle, c'est-à-dire s'il n'y a réellement aucune différence entre les groupes. Pour cette analyse, un seuil de $p < 0.05$ a été retenu comme critère de significativité. Autrement dit, si la p-value est inférieure à 0.05 pour une longueur d'onde donnée, cela signifie qu'il existe moins de 5 % de chances que la différence observée soit due au hasard. Une p-value inférieure à ce seuil permet donc de rejeter l'hypothèse nulle, indiquant qu'il y a une différence statistiquement significative entre les groupes pour cette longueur d'onde. Enfin, les zones grisées permettent d'identifier visuellement des régions spectrales présentant des différences statistiquement significatives entre les groupes, qui pourraient correspondre à des biomarqueurs d'intérêt pour détecter la présence d'une infection virale. Le recours au test de Student propose une importante plus-value à cet ensemble de figures, car il permet de poser un critère quantitatif aux différences observées qualitativement auparavant. Cela garantit que les variations observées dans les spectres Raman ne sont pas simplement le fruit du hasard, mais bien le reflet de différences biologiques significatives associées aux états infectieux : ce test statistique renforce ainsi la robustesse et la fiabilité des conclusions tirées des données spectrales en permettant l'identification de signatures potentielles.

Finalement, la dernière sous-figure est un *boxplot*, composé de plusieurs boîtes à moustaches, représentant l'intensité spectrale et sélectionnées chacune pour correspondre à une des longueurs d'onde pour lesquelles la p-value est inférieure à 0.05 et présente dans une des zones grisées : des étoiles sont ajoutées pour indiquer les seuils de significativité statistique : * pour $0.01 < p < 0.05$, ** pour $0.005 < p < 0.01$, et *** pour $p < 0.005$. Chaque boîte met en évidence les variations d'intensité pour les échantillons des deux groupes, en comparant les moyennes ainsi que l'étendue des valeurs d'intensité observées. Les moustaches indiquent la dispersion des données, tandis que les points situés au-delà représentent les valeurs aberrantes. Cela permet d'illustrer visuellement les différences potentielles entre les groupes, en montrant les tendances centrales et la variabilité autour de celles-ci.

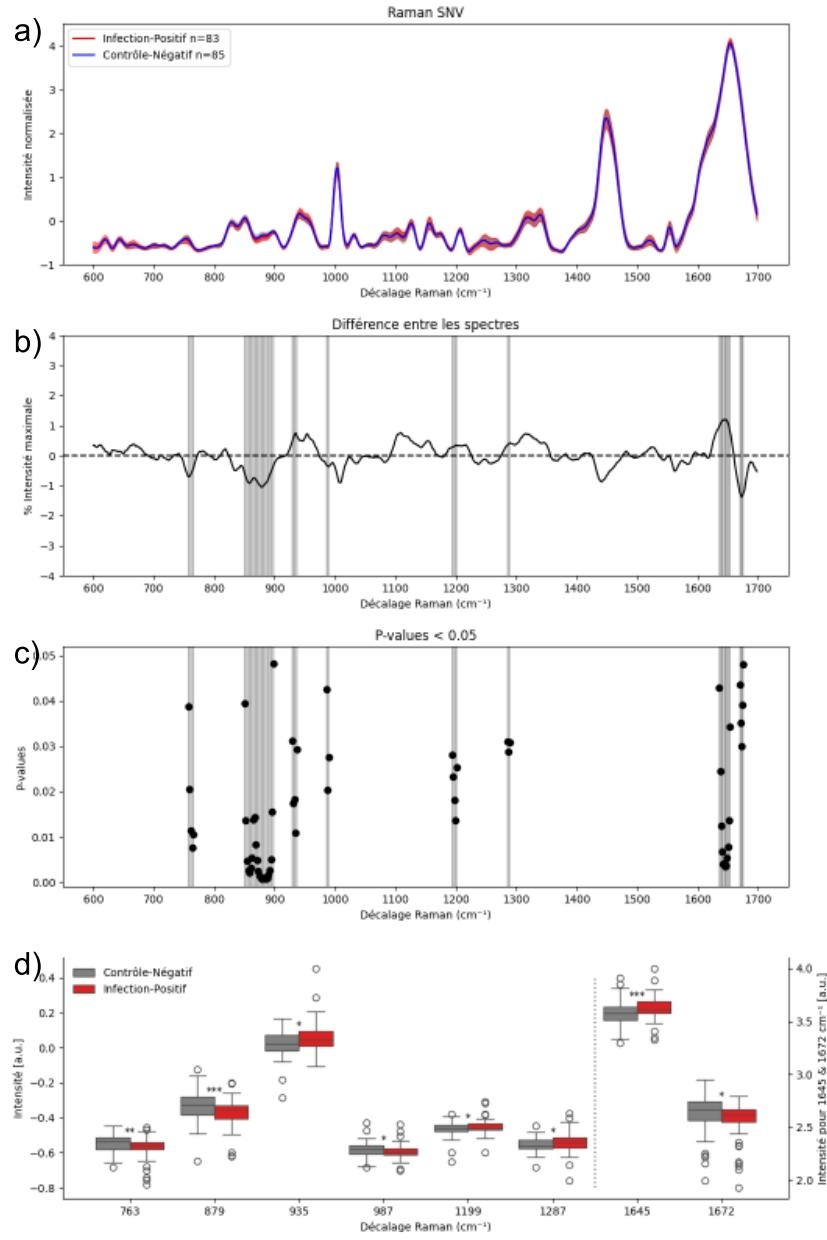


Figure 4.6. Résultats de l'étude de comparaison des échantillons testés positivement à une des infections virales, soit COVID-19, RSV et/ou Influenza, et des échantillons testés négativement à ces trois maladies. a) Comparaison des spectres SNV de cette étude. Les moyennes et écart-types sont affichés respectivement en rouge et en bleu. b) Différence entre les moyennes des spectres précédents, en pourcentage du maximum de l'intensité. Les régions grisées correspondent aux longueurs d'onde pour lesquelles la p-value calculée est inférieure à 0.05. c) P-values inférieures à 0.05 pour chaque longueur d'onde, illustrant les régions spectrales significativement différentes entre les groupes d'échantillons. d) Diagrammes en boîtes associés à des régions spectrales où plusieurs longueurs d'onde partagent des p-values < 0.05. Les seuils de significativité statistique sont indiqués par des étoiles *: 0.01 < p < 0.05, **: 0.005 < p < 0.01, ***: p < 0.005

L'analyse de cette première figure révèle plusieurs éléments clés quant aux différences spectrales entre les échantillons de ces deux groupes de patients. Dans la sous-figure a), les profils spectraux montrent globalement des tendances similaires, mais certaines longueurs d'onde présentent des variations distinctes et ce constat pose les bases pour des analyses plus approfondies afin de déterminer les longueurs d'onde où les écarts sont statistiquement significatifs, et les biomarqueurs qui leur sont associés. La sous-figure b) met en évidence ces quelques différences observées initialement, avec une représentation de la différence des deux spectres, en pourcentage de l'intensité maximale, ce qui permet de mieux visualiser les écarts, en soulignant également les régions où l'intensité des spectres diverge le plus fortement à l'aide des zones grisées. La sous-figure c) vient confirmer ces observations en montrant les p-values associées aux tests de Student appliqués à chaque longueur d'onde, qui indiquent des différences statistiquement significatives. On observe des clusters de points en dessous de ce seuil dans certaines régions spectrales, notamment autour de 900 cm^{-1} et 1650 cm^{-1} ce qui souligne l'existence de différences cohérentes et significatives entre les groupes pour ces régions spectrales. Ces résultats soutiennent l'hypothèse selon laquelle les différences observées ne sont pas dues au hasard, mais bien à des variations biochimiques associées aux infections virales. Enfin, les boîtes présentes dans la dernière sous-figure permettent une analyse plus fine des différences d'intensité entre les groupes. Les échantillons de patients positifs montrent des variations d'intensité spectrale significatives par rapport aux échantillons de contrôle, particulièrement marquées à certaines longueurs d'onde comme 763 , 879 , et 1645 cm^{-1} , avec des niveaux de significativité indiqués par des étoiles. Par exemple, la différence d'intensité pour la longueur d'onde 1645 cm^{-1} est particulièrement prononcée, avec un niveau de significativité $p < 0.005$ (indiqué par trois étoiles), ce qui souligne une distinction nette entre les groupes pour cette région. Cette sous-figure montre également une certaine dispersion des valeurs, reflétant la variabilité individuelle au sein de chaque groupe, mais les différences de médiane entre les groupes restent bien marquées pour les longueurs d'onde les plus significatives. En conclusion, cette première figure révèle que des signatures spectrales liées aux infections virales peuvent être identifiées, ce qui renforce l'intérêt de cette approche pour le diagnostic. Ces résultats initiaux sont très encourageants pour commencer l'étude et ils permettent d'envisager d'autres analyses plus spécifiques, en comparant les échantillons de chaque maladie aux échantillons de contrôle, dans l'objectif d'identifier des biomarqueurs propres à la maladie étudiée.

4.3.3 Résultats intermédiaires – Comparaisons entre chaque maladie

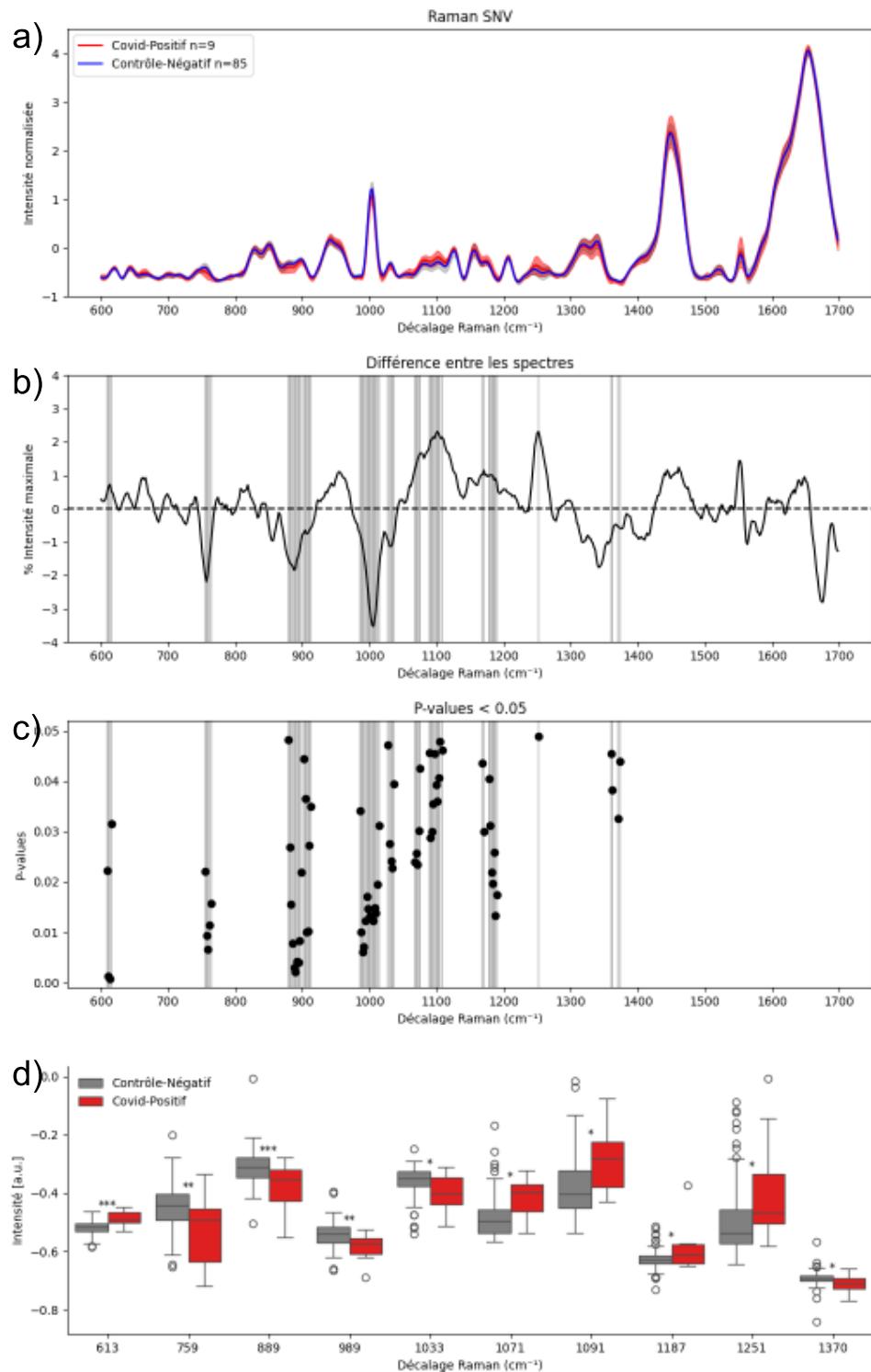


Figure 4.7. Résultats de l'étude de comparaison des échantillons testés positivement au COVID-19 et des échantillons testés négativement aux trois maladies virales étudiées.

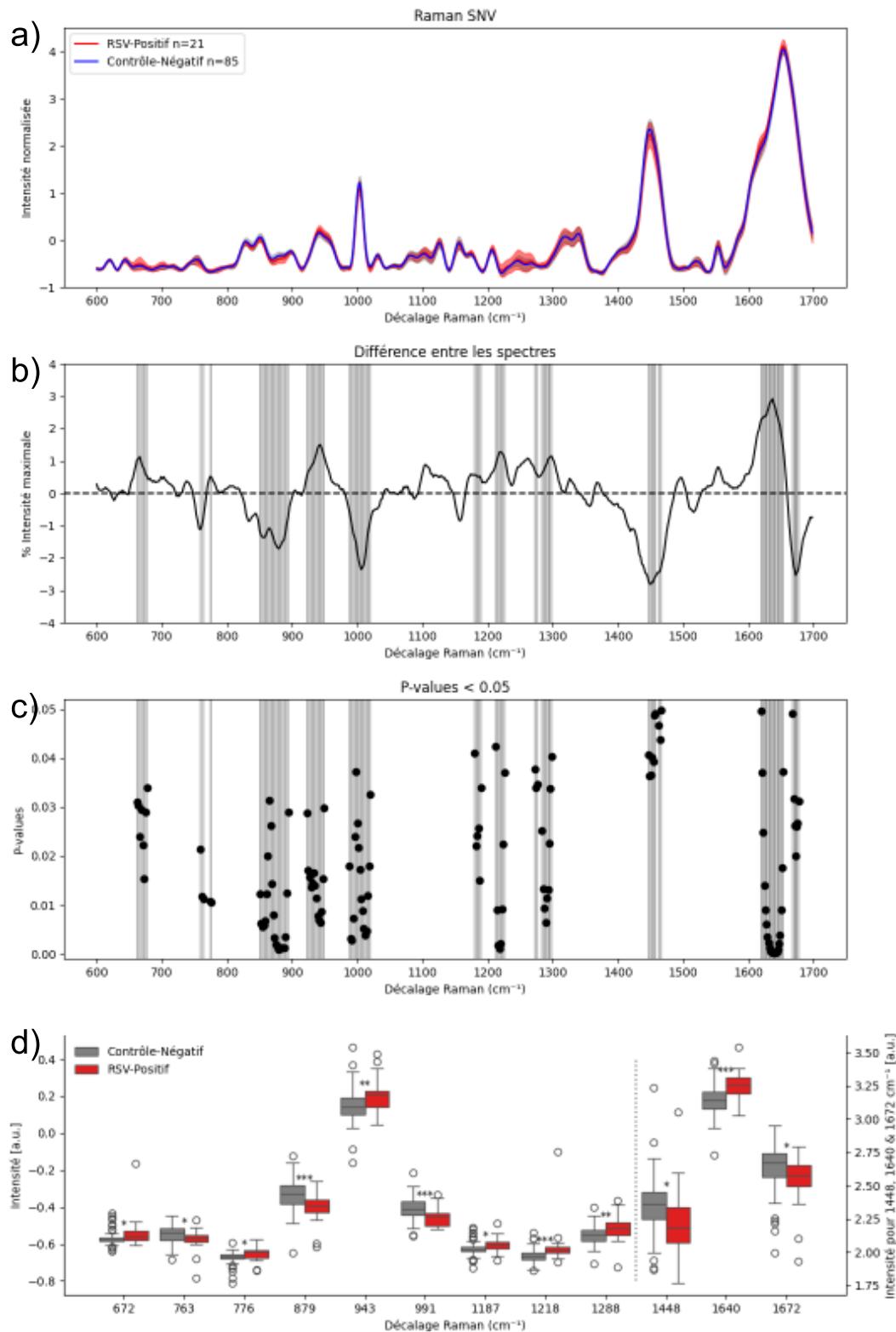


Figure 4.8. Résultats de l'étude de comparaison des échantillons testés positivement au RSV et des échantillons testés négativement aux trois maladies virales étudiées.

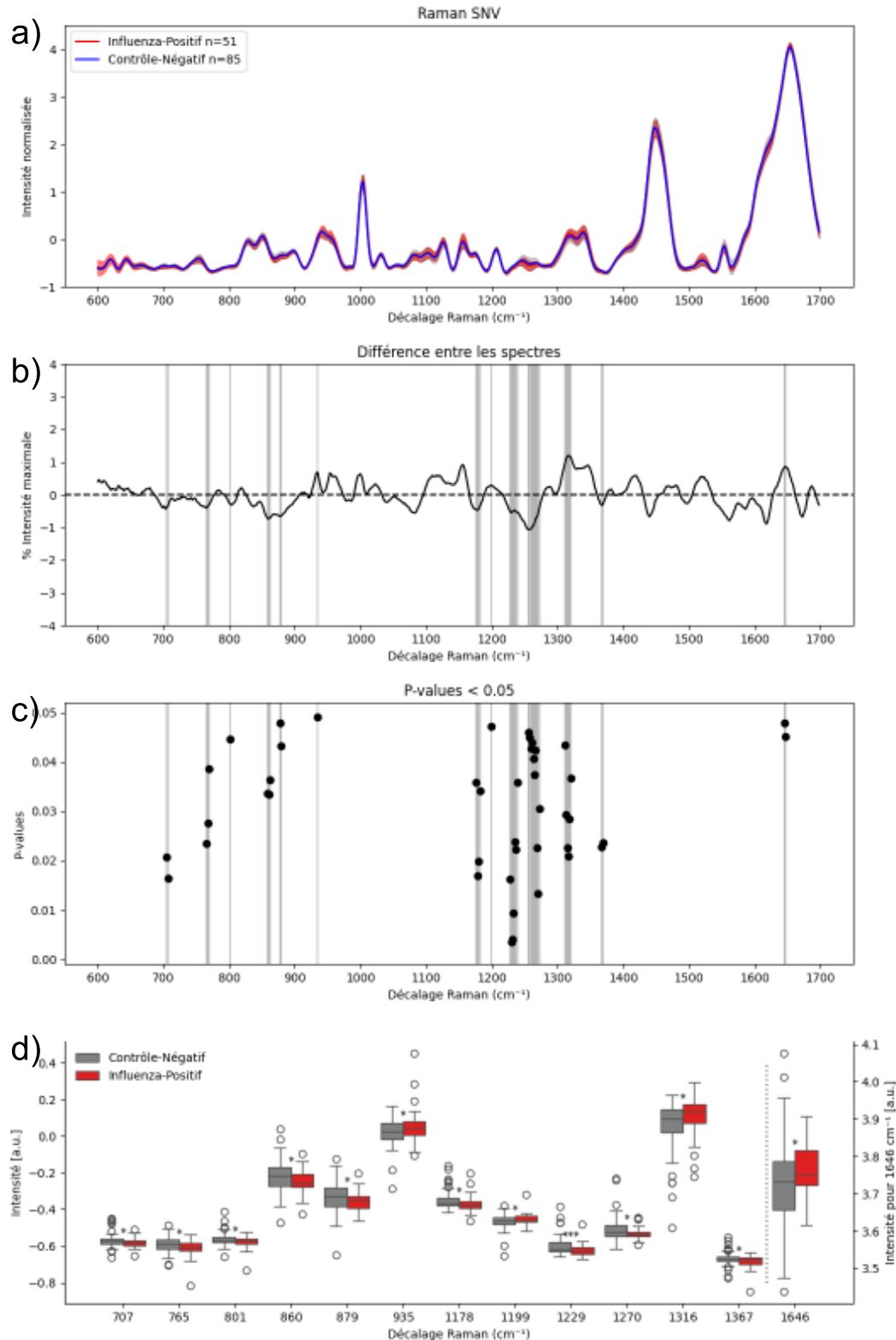


Figure 4.9. Résultats de l'étude de comparaison des échantillons testés positivement à Influenza et des échantillons testés négativement aux trois maladies virales étudiées.

L'analyse de ces figures permet de mettre en évidence des différences spectrales entre les groupes des trois maladies et de contrôle, bien que les différences de population entre les groupes demandent une certaine prudence quant à la conclusion des résultats. Il y a beaucoup moins d'échantillons notés comme COVID-19 ou même RSV comparativement aux échantillons de contrôle, ce qui fait que le calcul des p-values avec le test de Student est moins fiable qu'avec deux groupes de population équivalente. Les échelles des sous-figures b) et c) ont été maintenues entre les figures pour faciliter la comparaison des données et des variations observées. De même que pour la figure 4.6, les sous-figures a) présentent les profils spectraux entre les groupes, tandis que les sous-figures b) confirment les tendances observées et appuient sur les écarts en représentant la différence relative des spectres en pourcentage, soulignant les régions de divergence identifiées par les zones grisées associées aux p-values inférieures à 0.05, que l'on retrouve sur la sous-figure c), en ajoutant les valeurs de toutes les p-values inférieures à 0.05. Enfin, la sous-figure d) compare les intensités de longueurs d'onde spécifiques, où des variations significatives apparaissent, avec des boîtes à moustache.

Les analyses révèlent que les différences les plus marquées se situent entre les échantillons positifs au RSV et les échantillons de contrôle, avec de nombreuses p-values inférieures à 0.05, dont certaines très proches de 0. Cela indique des écarts significatifs et cohérents dans les profils spectraux, suggérant des variations biochimiques importantes et spécifiques au RSV. À l'inverse, les échantillons positifs à l'Influenza présentent peu de différences par rapport au groupe contrôle, ce qui est particulièrement visible sur la sous-figure b) où la différence des spectres reste minime par rapport aux autres figures et échantillons positifs au COVID-19 et à Influenza. Concernant les échantillons notés COVID-19 justement, l'interprétation est plus délicate en raison du faible nombre d'échantillons disponibles, mais certaines variations significatives se détachent malgré tout, ce qui pourrait indiquer des signatures spécifiques bien que moins marquées. Pour approfondir ces observations, nous allons procéder à une comparaison globale de tous ces groupes en sélectionnant les longueurs d'onde communes aux trois groupes et associées à des p-values significatives. Ces longueurs d'onde seront représentées à l'aide de boîtes à moustaches, ce qui permettra de visualiser les différences entre les groupes de manière plus intégrée et d'identifier des signatures spectrales potentielles partagées ou spécifiques.

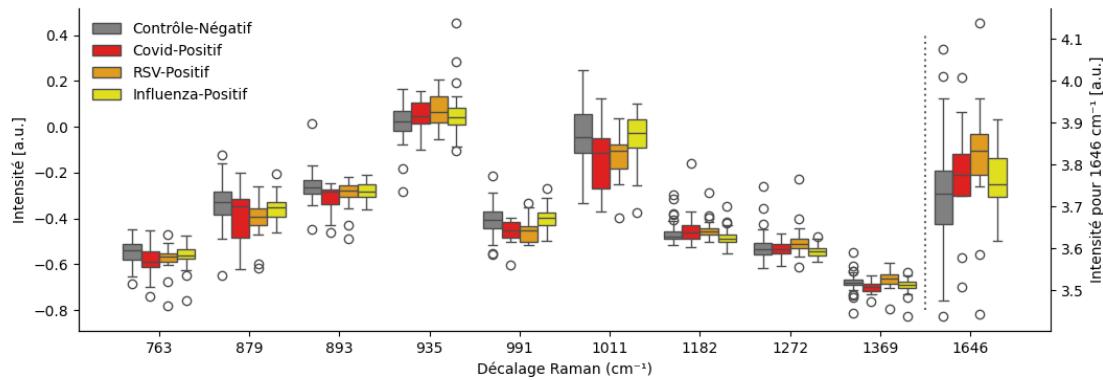


Figure 4.10. Diagrammes en boîtes illustrant les différences entre tous les groupes d'échantillons. Les diagrammes sont associés à des régions spectrales où plusieurs longueurs d'onde provenant de différents groupes partagent des p-values < 0.05.

Les 10 longueurs d'onde choisies pour représenter les différences spectrales entre les groupes de maladies couvrent diverses régions, allant de 763 cm^{-1} à 1646 cm^{-1} , ce qui permet de mieux visualiser les variations d'intensité entre les groupes. En effet, on retrouve dans cette figure des résultats observés auparavant : les données provenant des échantillons positifs au RSV sont indubitablement différentes, que ce soit en termes de médianes ou de premiers et troisièmes quartiles notamment. Parmi les différentes longueurs d'onde choisies, ces différences liées à la maladie RSV sont particulièrement présentes pour 879 cm^{-1} , 935 cm^{-1} , 1272 cm^{-1} ou encore 1369 cm^{-1} . De plus, cette figure permet de notifier les différences spectrales entre les maladies infectieuses. Il en ressort sur plusieurs aspects que le virus Influenza est à part par rapport au COVID-19 et au RSV, et semble plus similaire aux échantillons de contrôle. Ce constat est visible à partir de plusieurs régions spectrales : 879 cm^{-1} , 991 cm^{-1} , 1011 cm^{-1} ou bien même 1182 cm^{-1} . Il s'agira d'identifier les liaisons moléculaires responsables de ces pics Raman pour tenter d'attribuer les variations observées aux caractéristiques biochimiques propres à chaque pathologie, ou communs à certains groupes de cette étude.

4.3.4 Dernier résultat – Comparaison des échantillons en fonction de l'âge des patients

Ensuite, en accord avec les données présentées dans la Table 4.1, il semble intéressant de considérer dans l'analyse l'âge des participants. En effet, l'âge peut influencer les résultats spectroscopiques, et il est donc pertinent d'explorer s'il existe des tendances particulières ou des résultats remarquables à retirer de cette analyse par catégorie d'âge. La revue de littérature établie à ce sujet fait état de différences spectrales entre adultes et enfants [58], mais pas à différents stades de l'enfance, ce qui en fait un domaine d'investigation particulièrement intéressant. Pour évaluer cela, nous avons choisi de diviser les participants en huit catégories d'âge, couvrant la période allant de la naissance à 17 ans. Cette subdivision se base sur les changements biologiques et physiologiques qui surviennent à différents âges de l'enfance, et ces catégories sont clairement détaillées dans la Table 4.1.

Afin d'explorer cette influence de l'âge sur les résultats spectroscopiques, nous avons tracé les spectres par catégorie d'âge pour les trois maladies étudiées, à savoir le COVID-19, le RSV et l'Influenza. Pour chaque maladie, deux types de figures ont été créés : la première montre les spectres distincts pour chaque groupe d'âge, et la seconde présente un spectre moyen pour l'ensemble des échantillons, dans le but d'identifier des tendances globales. Ces figures sont disponibles dans le matériel supplémentaire, mais brièvement, les résultats obtenus montrent des différences spectrales marquées, particulièrement pour les échantillons RSV, mais aussi visibles dans les spectres des maladies COVID-19 et Influenza. Notamment, les deux premiers groupes d'âge, correspondant à des enfants de moins de six mois et représentés en noir et en rouge dans les figures, affichent des spectres distincts par rapport aux autres catégories : ces spectres se démarquent notamment par des pics supplémentaires ou absents, ou encore un écart-type à part. Ces variations soulèvent des questions importantes quant à l'influence de l'âge sur les résultats. Il est possible que ces premières phases de la vie, marquées par des changements biologiques rapides, aient un impact significatif sur les signatures spectrales. Pour vérifier ces résultats, il est décidé de tracer la figure de comparaison entre les populations d'échantillons d'individus supérieurs à six mois et ceux plus jeunes.

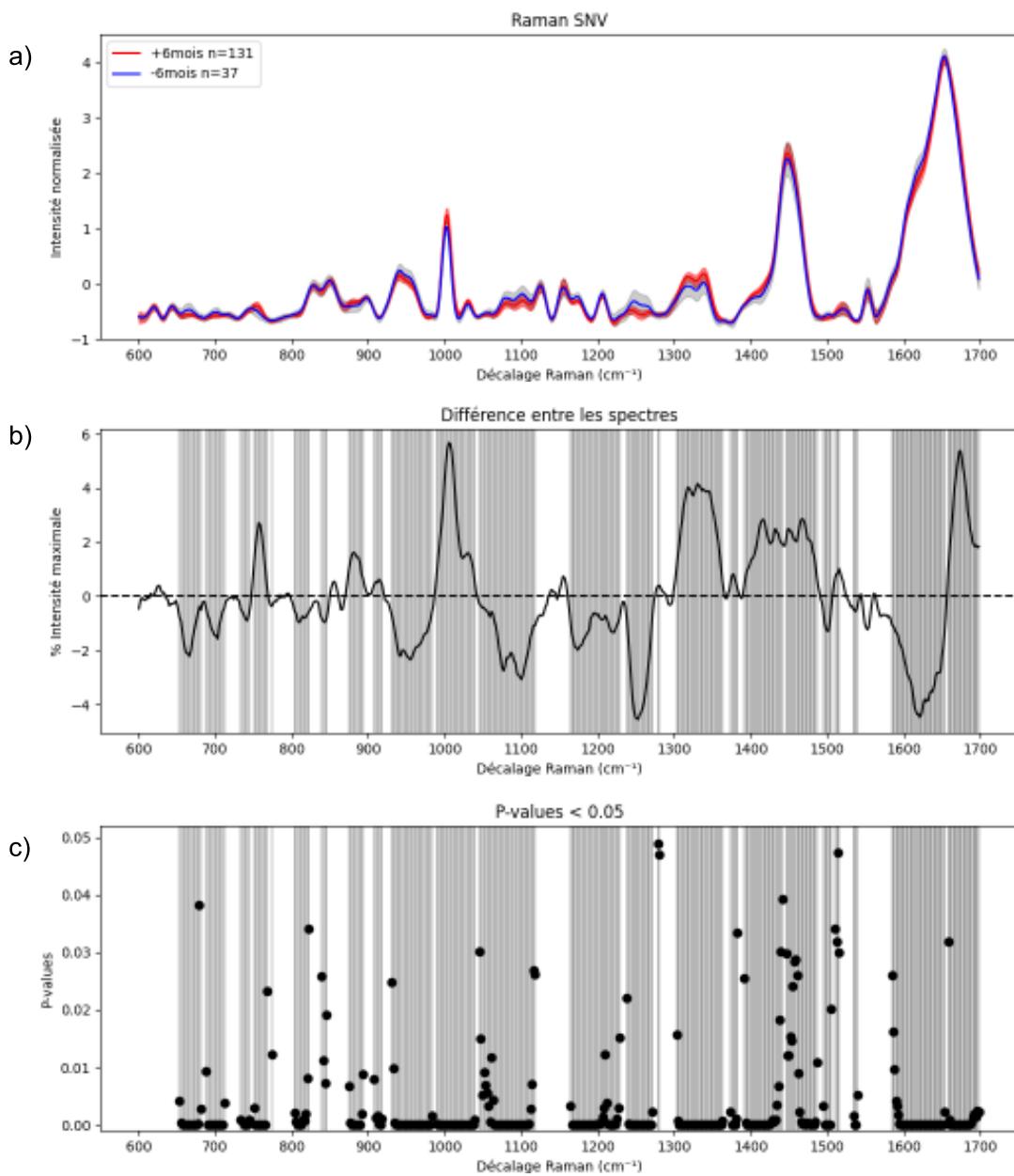


Figure 4.11. Résultats de l'étude de comparaison des échantillons provenant d'enfants âgés de moins de six mois et ceux provenant d'enfants de plus de six mois.

Finalement, les différences spectrales entre les échantillons sanguins d'enfants âgés de plus de six mois et ceux de moins de six mois sont ici confirmées de manière frappante. Les p-values, très proches de 0 sur la majeure partie de l'étendue spectrale étudiée, indiquent des écarts hautement significatifs entre les deux groupes d'âge. Cela suggère que les profils spectroscopiques des enfants plus âgés possèdent des caractéristiques biochimiques nettement distinctes, potentiellement liées aux changements physiologiques qui surviennent après cet âge, comme des évolutions dans le métabolisme, le développement immunitaire ou d'autres processus biologiques. Ces résultats mettent en avant les variations de spectres Raman associées à l'âge chez les enfants, qui pourrait servir d'indicateur dans des analyses plus poussées pour comprendre comment le vieillissement ou le développement affecte la composition biochimique du sang. En particulier, les longueurs d'onde présentant les différences les plus significatives pourraient correspondre à des biomarqueurs spécifiques, apportant des pistes intéressantes pour la recherche en pédiatrie.

4.4 Discussion – Étude principale : « Multi-Virus »

Le premier objectif de l'étude, visant à évaluer la réponse du système à des échantillons sanguins, a été atteint avec succès. Les mesures effectuées sur plusieurs échantillons de rat ont montré une excellente répétabilité, ce qui confirme la fiabilité du système pour cette application. L'utilisation d'une cassette pour certaines quantités inférieures au volume recherché a également permis de mieux comprendre les conditions de mesure optimales, et les premiers résultats obtenus sont encourageants quant à la robustesse du dispositif. De plus, une modification du code de traitement spectral a été apportée pour améliorer l'analyse des données, contribuant par la suite à une lecture plus précise et rapide des spectres, ce qui pourrait s'avérer bénéfique pour des études ultérieures. En somme, les expériences préliminaires ont bien démontré le potentiel du système avec des échantillons de sang et offert une base solide en introduction de l'étude principale.

Le deuxième objectif de l'étude, l'acquisition des données, a été validé sans encombre. Tous les échantillons ont pu être mesurés, il n'y a eu aucune mauvaise manipulation et aucun incident à rapporter. Cependant, l'étape de nettoyage s'est trouvée être très contraignante et pas forcément utile en tout point : le principal points remis en question ici concerne le nettoyage avec l'éthanol, qui semble superflu en prenant une nouvelle feuille d'aluminium à chaque fois, le nettoyage à l'eau distillée semblant suffisant dans ce cas. Une étude de comparaison simple pourra ainsi être

effectuée pour répondre à cette question. Malgré cette contrainte de nettoyage, les acquisitions se sont bien déroulées, le protocole était satisfaisant et pourra être réutilisé par la suite dans le laboratoire. L'organisation entre les étudiants a aussi été optimale : la mesure d'un échantillon, avec les tâches environnantes, prenait en moyenne deux minutes, ce qui est un résultat extrêmement encourageant dans l'optique de développer le système comme outil de dépistage rapide.

Concernant le troisième objectif, des différences spectrales ont en effet été observées, permettant de valider l'hypothèse définie au préalable : la spectroscopie Raman permet effectivement de détecter des différences dans les spectres d'échantillons sanguins entre des patients atteints de maladies virales spécifiques (COVID-19, RSV, Influenza) et des individus sains. Ces observations qualitatives ont été corroborées par des analyses quantitatives, effectuées à l'aide d'un test de Student, permettant le calcul de p-values pour évaluer la signification statistique des résultats obtenus sur l'ensemble du spectre étudié. Cependant, dans l'analyse de spectres Raman, les différences peuvent être observées soit sur l'ensemble du spectre, comme discuté déjà en détails, ou alors soit à partir de pics spécifiques. Les pics dans un spectre Raman correspondent aux vibrations des liaisons moléculaires, et leur position, leur intensité et leur largeur peuvent varier en fonction de la composition moléculaire ou de l'environnement chimique. Généralement, les deux approches, analyse de l'ensemble du spectre et analyse de pics spécifiques, sont souvent combinées pour maximiser la précision de l'analyse, les pics spécifiques servant à une première interprétation rapide, tandis que l'analyse globale du spectre apporte des informations complémentaires et permet une classification plus robuste dans des études comparatives.

En conséquence, pour approfondir l'analyse, nous avons comparé les pics Raman associés à des différences significatives entre des groupes de maladie et observés dans nos échantillons de sang avec ceux décrits dans la littérature scientifique. Cette comparaison permet de vérifier la concordance de nos résultats avec les données établies et d'interpréter biologiquement les différences spectrales significatives. Certains pics correspondent à des liaisons moléculaires bien connues, comme les vibrations de groupes de protéines, lipides ou autres.

Tableau 4.2. Identification de certains pics Raman observés dans nos résultats. Ces pics sont valables pour des échantillons sanguins et correspondent à des longueurs d'onde responsables de différences significatives entre deux groupes de notre étude. Ils sont tirés de la littérature scientifique (1 : [53], 2 : [54], 3: [55], 4: [56], 5: [57])

Pic Raman (cm ⁻¹)	Intervalle (cm ⁻¹)	Attribution moléculaire	Références associées
755	741-769	Protéine (tryptophane, phénylalanine, aspartate, isoleucine)	1, 2, 3, 5
849	835-863	Protéine (alanine, tyrosine, leucine, lysine, proline), glucose	1, 2, 5
870	863-877	Protéine (threonine, glutamate, arginine, aspartate, cystine), lipide (C-O-O)	1, 2, 3, 4, 5
903	896-910	Protéine (C-H), lipide (C-H)	1, 2, 3, 4
952	941-963	Protéine (tryptophane, valine)	3
1005	998-1012	Protéine (phénylalanine, tryptophane), caroténoides	1, 2, 3, 5
1034	1027-1041	Protéine (phénylalanine, arginine, glycine, lysine, isoleucine)	1, 2, 3, 5
1085	1073-1097	Acide gras (C-C)	4
1206	1197-1215	Protéine (tyrosine)	1, 2, 5
1255	1253-1265	Tryglycerol, protéine (amide III)	3, 4
1318	1312-1324	Protéine (C-H, arginine, glutamate, phénylalanine, tryptophane)	1, 2, 3, 5
1451	1436-1466	Protéine (C-H, isoleucine, valine), glucose, lipide (CH ₂ /CH ₃)	1, 2, 3, 4, 5
1654	1631-1677	Protéine (amide I), lipide (C=C)	1, 2, 5, 6

Pour parvenir à identifier les pics liés aux différences significatives et associés à nos résultats, il suffit de superposer les sous-figures a) et c) pour sélectionner les longueurs d'onde correspondantes : par exemple, avec la figure 4.8 présentant les résultats de la comparaison entre les échantillons de patients testés positivement au RSV et les échantillons contrôle, la superposition de ces sous-figures montre que la zone spectrale autour de 1000 cm⁻¹ inclut de nombreuses p-values inférieures à 0.05. D'après le tableau 4.2 précédent, cette région spectrale est associée à des acides aminés essentiels, à la phénylalanine ou au tryptophane notamment dans ce cas. D'après les spectres Raman de notre étude, le pic de la phénylalanine étant plus intense en moyenne pour des échantillons sanguins de personnes saines que d'échantillons porteurs du RSV, cela suggère une concentration plus élevée de phénylalanine chez une personne saine que chez une personne positive au RSV, ce qui pourrait indiquer un déficit de cet acide aminé induit par l'infection virale. Cette approche comparative nous permet donc non seulement de valider la pertinence des pics observés par rapport à la littérature, mais aussi d'interpréter les différences spectrales de manière plus fonctionnelle. En identifiant les liaisons moléculaires sous-jacentes à ces variations, nous obtenons des indices sur les changements biochimiques associés à chaque infection virale. Cette méthode

pourrait ainsi fournir des bases solides pour établir des marqueurs diagnostiques potentiels spécifiques à chaque pathologie et contribuer à la compréhension des effets des virus sur la composition sanguine des patients.

Finalement, pour discuter de la limite de nos résultats, le développement d'un classificateur binaire à partir de ces données, n'a pas été envisagé en raison principalement de l'écart de population entre les différents groupes de participants. Cependant, les mesures effectuées dans le cadre de cette étude seront un point de départ pour le développement futur d'un modèle de classification capable de différencier efficacement les groupes étudiés. En effet, ces données pourront être intégrées dans un processus d'apprentissage supervisé, en tenant compte des caractéristiques spectrales identifiées comme significatives, et, à terme, un tel classificateur pourrait offrir des applications concrètes, notamment dans le domaine du diagnostic rapide ou du suivi personnalisé des patients, renforçant ainsi l'impact clinique et scientifique des résultats obtenus.

Pour clore cette discussion, il a été remarqué que les résultats les plus prometteurs ne concernent pas forcément ce pour quoi l'étude a été imaginée à ses débuts : la comparaison d'échantillons d'enfants de plus ou moins de six mois a été significative et semble de première importance pour la suite du projet. La prochaine étape à envisager semble de recruter plus de participants durant la prochaine vague de virus hivernaux, en se concentrant sur une population pédiatrique positive au COVID-19 ou au RSV, mais l'objectif semble surtout de sélectionner des échantillons sanguins d'enfants d'âge inférieur à six mois. Ensuite, il s'agira de confirmer les résultats obtenus dans l'étude présentée dans ce rapport sur des populations d'échantillons équivalentes, puis de développer un classificateur binaire en priorisant les catégories de plus ou moins de six mois.

CHAPITRE 5 DISCUSSION GÉNÉRALE

On va revenir maintenant sur les objectifs et questions abordés en introduction, qui découlaient de l'objectif principal du projet : optimiser le signal de détection Raman pour détecter des biomarqueurs dans des échantillons sanguins.

Pour cela, il a fallu opérer des modifications et tests sur le système de spectroscopie, qui ont néanmoins permis d'améliorer certaines caractéristiques tout en conservant une bonne qualité de signal. De plus, les avancées sont bénéfiques pour le futur du laboratoire car ce sont des solutions viables et durables dans le temps : le choix de la cassette plutôt que celui de la cuvette offre des résultats plus répétables, la miniaturisation du système n'affecte pas la qualité des spectres Raman, et les cassettes de calibration imprimées améliorent la facilité d'utilisation du système. On peut affirmer aisément que l'objectif n°1 est rempli, le système de spectroscopie avec ses composants externes a été développé pour qu'il corresponde au mieux aux attentes de l'équipe. Pour les futures évolutions du système, deux axes d'amélioration clés sont envisagés pour répondre aux besoins du laboratoire en termes de taille du système et de performance. La première amélioration consiste à réduire la taille du système, en se concentrant en particulier sur la miniaturisation du laser et du spectromètre. Actuellement, bien que le système soit déjà relativement compact, l'objectif est de concevoir une version encore plus petite, pour rendre le système plus facilement transportable et utilisable sur le terrain ou dans des environnements cliniques où l'espace est limité. Pour cela, il faudrait commencer par sélectionner un spectromètre plus petit mais, en accord avec des tests déjà effectués, cela a un impact négatif sur la qualité des spectres, notamment à cause de la caméra du spectromètre donc la résolution deviendrait trop faible pour nos applications. Finalement, le principal défi de la miniaturisation du système est de conserver le niveau de qualité des spectres Raman déjà disponible. En effet, le second axe d'amélioration porte sur l'optimisation du signal de diffusion Raman. Actuellement, une partie de la lumière diffusée est perdue en raison de la qualité des fibres optiques utilisées et de la configuration de la sonde optique. Pour remédier à cela, il est envisagé de remplacer les fibres optiques, ce qui permettrait d'augmenter la quantité de lumière recueillie par le spectromètre et, par conséquent, d'améliorer le rapport SNR. Les autres composantes du système qui peuvent être modifiées pour rendre le système plus performant et améliorer la qualité des spectres sont la sonde optique, essentielle pour converger le faisceau

lumineux en un point précieux et pour collecter les photons diffusés, le design de la boîte-lecteur permettant d'isoler l'échantillon lors de la mesure ou encore une nouvelle façon innovatrice de contenir l'échantillon étudié.

Ensuite, pour parvenir à mesurer une biobanque aussi conséquente, il a fallu mettre en place un protocole solide. Celui imaginé et réalisé pour la mesure de l'étude principale a donné entière satisfaction. Cependant, et malgré l'étude de répétabilité du sang qui a confirmé nos attentes dans le protocole, il aurait été judicieux de passer plus de temps à développer ce protocole pour l'optimiser. Il est pertinent de remettre en question le protocole de nettoyage, la quantité de liquide mesuré, la durée de l'étape de centrifugation... Par manque de temps, ces étapes d'essais/erreurs n'ont pas été approfondies. L'objectif n°2, mettre en place un protocole expérimental robuste garantissant la précision et la reproductibilité des mesures d'échantillons sanguins, n'est pas tout à fait accompli.

Enfin, pour mener à un bilan qui répond à l'introduction de ce rapport, le travail fourni au cours de la recherche a permis de confirmer l'hypothèse de travail, le système de spectroscopie du LUMEDLAB permet d'obtenir un signal Raman cohérent et d'observer des différences spectrales entre des échantillons malades et sains, ce qui valide définitivement le dernier objectif, mais ce travail a également permis de dépasser cette hypothèse et remplir l'objectif général fixé : le signal Raman a bien été optimisé à partir d'un nouveau système de spectroscopie et pour des échantillons de sang.

CHAPITRE 6 CONCLUSION

Pour conclure, il s'agit ici de faire le point sur deux aspects du travail effectué jusqu'à présent. Premièrement, la recherche conduite autour du développement du système s'est trouvée être positive : les résultats obtenus avec la technique de la cassette recouverte d'une feuille d'aluminium sont fiables et permettent d'obtenir des spectres Raman pertinents. Le cahier des charges imposé par le laboratoire, en termes de robustesse ou de qualité de signal, a été respecté. Au niveau robustesse, la cassette donne une mesure fiable et permet de prendre en compte la quantité d'échantillon que l'on souhaite mesurer en changeant la taille du puits, donc cette solution semble partie pour durer. Au niveau qualité du signal maintenant, tous les spectres obtenus après modifications du système présentaient un QF de qualité ainsi que des spectres très précis pour des données semblables, validant de ce fait ces modifications ainsi que le protocole expérimental. Deuxièmement, la recherche menée sur l'étude de la biobanque « Multi-virus », pour la première fois avec des échantillons sanguins pour ce système, est encourageante, l'étape de méthodologie et le protocole s'étant déroulés sans encombre et avec efficacité. En termes de résultats, il a été décidé de mener une recherche de nature exploratoire, en conservant la totalité des échantillons pour établir une base préliminaire de résultats. Par la suite, il s'agira de confirmer ces résultats, en équilibrant par d'autres prises de données les groupes de maladies étudiées, et d'approfondir la réflexion des différences spectrales entre les échantillons sanguins d'enfants âgés de plus ou moins de six mois.

RÉFÉRENCES

- [1] K. Ember *et al.*, « Saliva-based detection of COVID-19 infection in a real-world setting using reagent-free Raman spectroscopy and machine learning », *J. Biomed. Opt.*, vol. 27, n° 2, Art. n° 2, févr. 2022, doi: 10.1117/1.jbo.27.2.025002.
- [2] M. Jermyn *et al.*, « Intraoperative brain cancer detection with Raman spectroscopy in humans », *Sci. Transl. Med.*, vol. 7, n° 274, p. 274ra19-274ra19, févr. 2015, doi: 10.1126/scitranslmed.aaa2384.
- [3] K. Rebrosova, O. Samek, M. Kizovsky, S. Bernatova, V. Hola, et F. Ruzicka, « Raman Spectroscopy—A Novel Method for Identification and Characterization of Microbes on a Single-Cell Level in Clinical Settings », *Front. Cell. Infect. Microbiol.*, vol. 12, avr. 2022, doi: 10.3389/fcimb.2022.866463.
- [4] J. R. Ferraro, *Introductory Raman Spectroscopy*. Elsevier, 2003.
- [5] E. Smith et G. Dent, *Modern Raman Spectroscopy: A Practical Approach*. John Wiley & Sons, 2019.
- [6] C. Boudoux, *Fundamentals of Biomedical Optics*. 2016.
- [7] L. Lin *et al.*, « Direct and quantitative assessments of near-infrared light attenuation and spectroscopic detection depth in biological tissues using surface-enhanced Raman scattering », *Med-X*, vol. 1, n° 1, p. 9, août 2023, doi: 10.1007/s44258-023-00010-2.
- [8] B. E. A. Saleh et M. C. Teich, *Fundamentals of photonics*, Third edition. in Wiley series in pure and applied optics. Hoboken, NJ: Wiley, 2019.
- [9] G. A. Wagnières, W. M. Star, et B. C. Wilson, « In vivo fluorescence spectroscopy and imaging for oncological applications », *Photochem. Photobiol.*, vol. 68, n° 5, p. 603-632, nov. 1998.
- [10] P. Giamougiannis *et al.*, « A comparative analysis of different biofluids towards ovarian cancer diagnosis using Raman microspectroscopy », *Anal. Bioanal. Chem.*, vol. 413, n° 3, p. 911-922, janv. 2021, doi: 10.1007/s00216-020-03045-1.
- [11] F. Tian, L. F. das C. e S. de Carvalho, A. Casey, M. S. Nogueira, et H. J. Byrne, « Surface-Enhanced Raman Analysis of Uric Acid and Hypoxanthine Analysis in Fractionated Bodily Fluids », *Nanomaterials*, vol. 13, n° 7, Art. n° 7, janv. 2023, doi: 10.3390/nano13071216.
- [12] A. Golparvar, L. Thenot, A. Boukhayma, et S. Carrara, « Soft Epidermal Paperfluidics for Sweat Analysis by Ratiometric Raman Spectroscopy », *Biosensors*, vol. 14, n° 1, Art. n° 1, janv. 2024, doi: 10.3390/bios14010012.
- [13] E. Zamani *et al.*, « Spectral effects and enhancement quantification in healthy human saliva with surface-enhanced Raman spectroscopy using silver nanopillar substrates », *Lasers Surg. Med.*, vol. 56, n° 2, Art. n° 2, févr. 2024, doi: 10.1002/lsm.23746.
- [14] I. Taleb *et al.*, « Diagnosis of hepatocellular carcinoma in cirrhotic patients: a proof-of-concept study using serum micro-Raman spectroscopy », *Analyst*, vol. 138, n° 14, p. 4006-4014, 2013, doi: 10.1039/C3AN00245D.

- [15] H. Hano *et al.*, « Power of Light: Raman Spectroscopy and Machine Learning for the Detection of Lung Cancer », *ACS Omega*, vol. 9, n° 12, p. 14084-14091, mars 2024, doi: 10.1021/acsomega.3c09537.
- [16] J. L. González-Solís, J. C. Martínez-Espinosa, L. A. Torres-González, A. Aguilar-Lemarroy, L. F. Jave-Suárez, et P. Palomares-Anda, « Cervical cancer detection based on serum sample Raman spectroscopy », *Lasers Med. Sci.*, vol. 29, n° 3, p. 979-985, mai 2014, doi: 10.1007/s10103-013-1447-6.
- [17] A. Sahu, S. Sawant, H. Mamgain, et C. Murali Krishna, « Raman spectroscopy of serum: an exploratory study for detection of oral cancers », *Analyst*, vol. 138, n° 14, p. 4161-4174, 2013, doi: 10.1039/C3AN00308F.
- [18] X. Li, T. Yang, et S. Li, « Discrimination of serum Raman spectroscopy between normal and colorectal cancer using selected parameters and regression-discriminant analysis », *Appl. Opt.*, vol. 51, n° 21, p. 5038-5043, juill. 2012, doi: 10.1364/AO.51.005038.
- [19] X. Li *et al.*, « Blood plasma resonance Raman spectroscopy combined with multivariate analysis for esophageal cancer detection », *J. Biophotonics*, vol. 14, n° 9, p. e202100010, sept. 2021, doi: 10.1002/jbio.202100010.
- [20] P. Rekha, P. Aruna, E. Brindha, D. Koteeswaran, M. Baludavid, et S. Ganesan, « Near-infrared Raman spectroscopic characterization of salivary metabolites in the discrimination of normal from oral premalignant and malignant conditions », *J. Raman Spectrosc.*, vol. 47, n° 7, p. 763-772, 2016, doi: 10.1002/jrs.4897.
- [21] R.-M. Borşa *et al.*, « Developing New Diagnostic Tools Based on SERS Analysis of Filtered Salivary Samples for Oral Cancer Detection », *Int. J. Mol. Sci.*, vol. 24, n° 15, Art. n° 15, janv. 2023, doi: 10.3390/ijms241512125.
- [22] A. Falamas *et al.*, « Rapid and noninvasive diagnosis of oral and oropharyngeal cancer based on micro-Raman and FT-IR spectra of saliva », *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*, vol. 252, p. 119477, mai 2021, doi: 10.1016/j.saa.2021.119477.
- [23] S. Feng *et al.*, « Surface-enhanced Raman spectroscopy of saliva proteins for the noninvasive differentiation of benign and malignant breast tumors », *Int. J. Nanomedicine*, vol. 10, p. 537-547, janv. 2015, doi: 10.2147/IJN.S71811.
- [24] W. Liu *et al.*, « Data mining methods of lung cancer diagnosis by saliva tests using surface enhanced Raman spectroscopy », in *2014 7th International Conference on Biomedical Engineering and Informatics*, oct. 2014, p. 623-627. doi: 10.1109/BMEI.2014.7002849.
- [25] J. de J. Zermeño-Nava *et al.*, « Determination of sialic acid in saliva by means of surface-enhanced Raman spectroscopy as a marker in adnexal mass patients: ovarian cancer vs benign cases », *J. Ovarian Res.*, vol. 11, n° 1, p. 61, juill. 2018, doi: 10.1186/s13048-018-0433-9.
- [26] Y. Chen *et al.*, « Salivary Analysis Based on Surface Enhanced Raman Scattering Sensors Distinguishes Early and Advanced Gastric Cancer Patients from Healthy Persons », *J. Biomed. Nanotechnol.*, vol. 14, n° 10, p. 1773-1784, oct. 2018, doi: 10.1166/jbn.2018.2621.
- [27] G. Del Mistro *et al.*, « Surface-enhanced Raman spectroscopy of urine for prostate cancer detection: a preliminary study », *Anal. Bioanal. Chem.*, vol. 407, n° 12, p. 3271-3275, mai 2015, doi: 10.1007/s00216-015-8610-9.

- [28] X. Cui *et al.*, « Label-free detection of multiple genitourinary cancers from urine by surface-enhanced Raman spectroscopy », *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*, vol. 240, p. 118543, oct. 2020, doi: 10.1016/j.saa.2020.118543.
- [29] J. B. Phyo *et al.*, « Label-Free SERS Analysis of Urine Using a 3D-Stacked AgNW-Glass Fiber Filter Sensor for the Diagnosis of Pancreatic Cancer and Prostate Cancer », *Anal. Chem.*, vol. 93, n° 8, p. 3778-3785, mars 2021, doi: 10.1021/acs.analchem.0c04200.
- [30] V. Moisoiu *et al.*, « Breast Cancer Diagnosis by Surface-Enhanced Raman Scattering (SERS) of Urine », *Appl. Sci.*, vol. 9, n° 4, Art. n° 4, janv. 2019, doi: 10.3390/app9040806.
- [31] S. Li *et al.*, « Noninvasive prostate cancer screening based on serum surface-enhanced Raman spectroscopy and support vector machine », *Appl. Phys. Lett.*, vol. 105, n° 9, p. 091104, sept. 2014, doi: 10.1063/1.4892667.
- [32] D. K. R. Medipally *et al.*, « Vibrational spectroscopy of liquid biopsies for prostate cancer diagnosis », *Ther. Adv. Med. Oncol.*, vol. 12, p. 1758835920918499, janv. 2020, doi: 10.1177/1758835920918499.
- [33] Y. Ma, J. Chi, Z. Zheng, A. Attygalle, I. Y. Kim, et H. Du, « Therapeutic prognosis of prostate cancer using surface-enhanced Raman scattering of patient urine and multivariate statistical analysis », *J. Biophotonics*, vol. 14, n° 1, Art. n° 1, janv. 2021, doi: 10.1002/jbio.202000275.
- [34] S. Chen *et al.*, « Raman Spectroscopy Reveals Abnormal Changes in the Urine Composition of Prostate Cancer: An Application of an Intelligent Diagnostic Model with a Deep Learning Algorithm », *Adv. Intell. Syst.*, vol. 3, n° 4, p. 2000090, 2021, doi: 10.1002/aisy.202000090.
- [35] N. Chen *et al.*, « Surface-enhanced Raman spectroscopy of serum accurately detects prostate cancer in patients with prostate-specific antigen levels of 4–10 ng/mL », *Int. J. Nanomedicine*, vol. 12, p. 5399-5407, juill. 2017, doi: 10.2147/IJN.S137756.
- [36] D. K. R. Medipally *et al.*, « Development of a high throughput (HT) Raman spectroscopy method for rapid screening of liquid blood plasma from prostate cancer patients », *Analyst*, vol. 142, n° 8, p. 1216-1226, 2017, doi: 10.1039/C6AN02100J.
- [37] X. Shao *et al.*, « Evaluation of expressed prostatic secretion and serum using surface-enhanced Raman spectroscopy for the noninvasive detection of prostate cancer, a preliminary study », *Nanomedicine Nanotechnol. Biol. Med.*, vol. 13, n° 3, p. 1051-1059, avr. 2017, doi: 10.1016/j.nano.2016.12.001.
- [38] A. Stefancu *et al.*, « Combining SERS analysis of serum with PSA levels for improving the detection of prostate cancer », *Nanomed.*, vol. 13, n° 19, p. 2455-2467, oct. 2018, doi: 10.2217/nnm-2018-0127.
- [39] V. C. Munteanu *et al.*, « New Insights into the Multivariate Analysis of SER Spectra Collected on Blood Samples for Prostate Cancer Detection: Towards a Better Understanding of the Role Played by Different Biomolecules on Cancer Screening: A Preliminary Study », *Cancers*, vol. 14, n° 13, Art. n° 13, juin 2022, doi: 10.3390/cancers14133227.
- [40] S. Gao *et al.*, « Label-free surface enhanced Raman spectroscopy analysis of blood serum via coffee ring effect for accurate diagnosis of cancers », *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*, vol. 267, p. 120605, févr. 2022, doi: 10.1016/j.saa.2021.120605.

- [41] V. Moisoiu *et al.*, « SERS-based differential diagnosis between multiple solid malignancies: breast, colorectal, lung, ovarian and oral cancer », *Int. J. Nanomedicine*, vol. 14, p. 6165-6178, août 2019, doi: 10.2147/IJN.S198684.
- [42] D. Hu *et al.*, « Detecting urine metabolites of bladder cancer by surface-enhanced Raman spectroscopy », *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*, vol. 247, p. 119108, févr. 2021, doi: 10.1016/j.saa.2020.119108.
- [43] S. Chen *et al.*, « Identifying non-muscle-invasive and muscle-invasive bladder cancer based on blood serum surface-enhanced Raman spectroscopy », *Biomed. Opt. Express*, vol. 10, n° 7, p. 3533-3544, juill. 2019, doi: 10.1364/BOE.10.003533.
- [44] S. Li *et al.*, « Characterization and noninvasive diagnosis of bladder cancer with serum surface enhanced Raman spectroscopy and genetic algorithms », *Sci. Rep.*, vol. 5, n° 1, p. 9582, mai 2015, doi: 10.1038/srep09582.
- [45] Y. Zhang *et al.*, « Classifying low-grade and high-grade bladder cancer using label-free serum surface-enhanced Raman spectroscopy and support vector machine », *Laser Phys.*, vol. 28, n° 3, p. 035603, févr. 2018, doi: 10.1088/1555-6611/aa9d6d.
- [46] N. Jackson, J. Hassan, et H. J. Byrne, « Raman spectroscopic analysis of human serum samples of convalescing COVID-19 positive patients », *Clin. Spectrosc.*, vol. 5, p. 100028, déc. 2023, doi: 10.1016/j.clispe.2023.100028.
- [47] G. Yin *et al.*, « An efficient primary screening of COVID-19 by serum Raman spectroscopy », *J. Raman Spectrosc.*, vol. 52, n° 5, p. 949-958, 2021, doi: 10.1002/jrs.6080.
- [48] C. Carluomagno *et al.*, « COVID-19 salivary Raman fingerprint: innovative approach for the detection of current and past SARS-CoV-2 infections », *Sci. Rep.*, vol. 11, n° 1, p. 4943, mars 2021, doi: 10.1038/s41598-021-84565-3.
- [49] S. Desai *et al.*, « Raman spectroscopy-based detection of RNA viruses in saliva: A preliminary report », *J. Biophotonics*, vol. 13, n° 10, p. e202000189, 2020, doi: 10.1002/jbio.202000189.
- [50] S. Shanmukh, L. Jones, Y.-P. Zhao, J. D. Driskell, R. A. Tripp, et R. A. Dluhy, « Identification and classification of respiratory syncytial virus (RSV) strains by surface-enhanced Raman spectroscopy and multivariate statistical techniques », *Anal. Bioanal. Chem.*, vol. 390, n° 6, p. 1551-1555, mars 2008, doi: 10.1007/s00216-008-1851-0.
- [51] K. J. Ember *et al.*, « Rapid detection of COVID-19 in liquid saliva using reagent-free Raman spectroscopy », in *Optical Diagnostics and Sensing XXIV: Toward Point-of-Care Diagnostics*, SPIE, mars 2024, p. PC1285005. doi: 10.1117/12.3001877.
- [52] G. Sheehy *et al.*, « Open-sourced Raman spectroscopy data processing package implementing a baseline removal algorithm validated from multiple datasets acquired in human tissue and biofluids », *J. Biomed. Opt.*, vol. 28, n° 2, p. 025002, févr. 2023, doi: 10.1117/1.JBO.28.2.025002.
- [53] C. G. Atkins, K. Buckley, M. W. Blades, et R. F. B. Turner, « Raman Spectroscopy of Blood and Blood Components », *Appl. Spectrosc.*, vol. 71, n° 5, p. 767-793, mai 2017, doi: 10.1177/0003702816686593.

- [54] J. Y. Qu, B. C. Wilson, et D. Suria, « Concentration measurements of multiple analytes in human sera by near-infrared laser Raman spectroscopy », *Appl. Opt.*, vol. 38, n° 25, p. 5491-5498, sept. 1999, doi: 10.1364/AO.38.005491.
- [55] A. Rygula, K. Majzner, K. M. Marzec, A. Kaczor, M. Pilarczyk, et M. Baranska, « Raman spectroscopy of proteins: a review », *J. Raman Spectrosc.*, vol. 44, n° 8, p. 1061-1076, 2013, doi: 10.1002/jrs.4335.
- [56] K. Czamara, K. Majzner, M. Z. Pacia, K. Kochan, A. Kaczor, et M. Baranska, « Raman spectroscopy of lipids: a review », *J. Raman Spectrosc.*, vol. 46, n° 1, p. 4-20, 2015, doi: 10.1002/jrs.4607.
- [57] G. Zhu, X. Zhu, Q. Fan, et X. Wan, « Raman spectra of amino acids and their aqueous solutions », *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*, vol. 78, n° 3, p. 1187-1195, mars 2011, doi: 10.1016/j.saa.2010.12.079.
- [58] A. A. Figaji, « Anatomical and Physiological Differences between Children and Adults Relevant to Traumatic Brain Injury and the Implications for Clinical Assessment and Care », *Front. Neurol.*, vol. 8, déc. 2017, doi: 10.3389/fneur.2017.00685.

ANNEXE

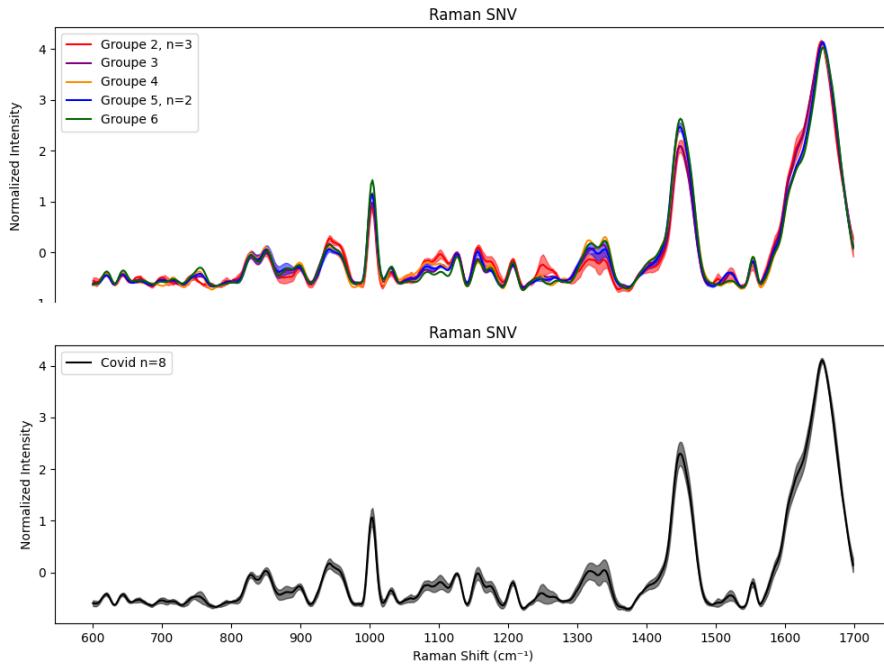


Figure supplémentaire n°1. Comparaison des échantillons positifs au COVID-19 par catégories d'âge des participants.

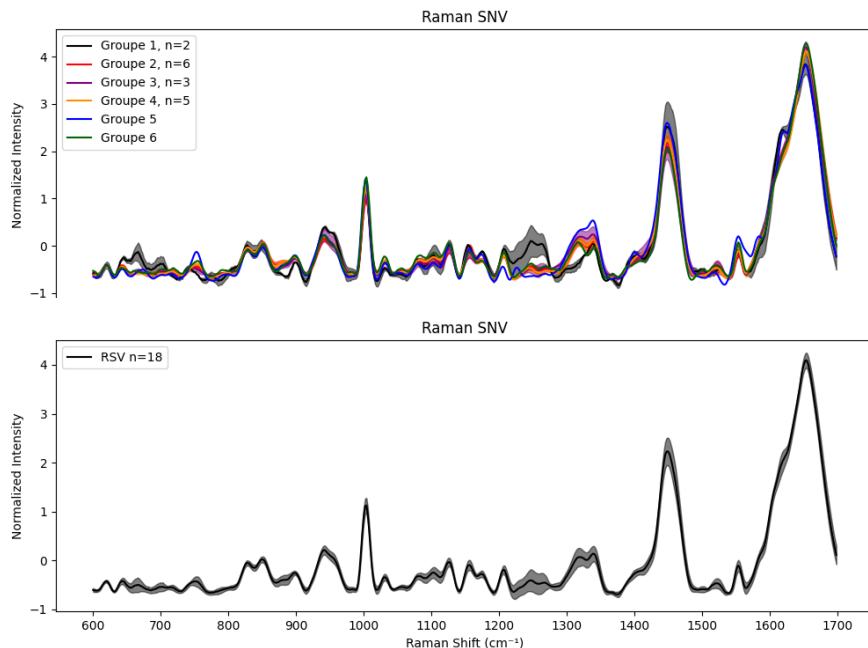


Figure supplémentaire n°2. Comparaison des échantillons positifs au RSV par catégories d'âge des participants.

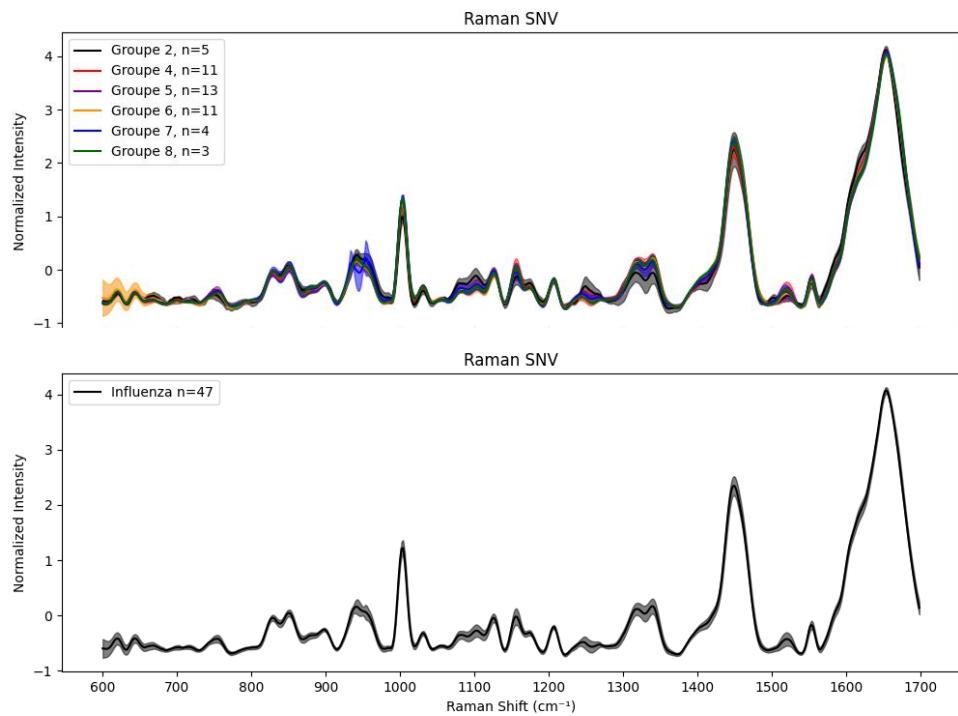


Figure supplémentaire n°3. Comparaison des échantillons positifs à Influenza par catégories d'âge des participants.