



**HAL**  
open science

# Protéomique unicellulaire pour accélérer la découverte de nouveaux médicaments

Wafa Hechiche

► **To cite this version:**

Wafa Hechiche. Protéomique unicellulaire pour accélérer la découverte de nouveaux médicaments. Biologie cellulaire. Université Paris sciences et lettres, 2023. Français. NNT : 2023UPSLS009 . tel-04187935

**HAL Id: tel-04187935**

**<https://pastel.hal.science/tel-04187935v1>**

Submitted on 25 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**THÈSE DE DOCTORAT**  
DE L'UNIVERSITÉ PSL

Préparée à l'École supérieure de physique et de chimie industrielles de la ville de Paris  
à l'UMR 8249 \_ Plasticité du cerveau

**Protéomique unicellulaire pour accélérer la découverte de nouveaux médicaments**

SINGLE CELL PROTEOMICS FOR ACCELERATING DRUG DISCOVERY PIPELINE

Soutenue par  
Wafa HECHICHE  
Le 25 janvier 2023

École doctorale n° 388  
**Chimie physique et  
chimie analytique de  
Paris centre**

Spécialité  
**Chimie Analytique**

Composition du jury :

**Anne, VARENNE**

Professeure, Chimie Paris Tech, PSL

*Présidente du jury*

**Claire, SMADJA**

Professeure, Paris Saclay

*Rapporteuse*

**Marie-Pierre, BOUSQUET**

Professeure, Paul Sabatier Toulouse

*Rapporteuse*

**François, FENAILLE**

Chercheur, CEA Saclay

*Examineur*

**Jacques, FATTACCIOLI**

MCU SU, IPGG, ENS Paris, PSL

*Co-directeur de thèse*

**Joëlle, VINH**

DR CNRS, SMBP ESPCI-PSL

*Directrice de thèse*

**Armelle, BUZY**

Cadre supérieur de laboratoire, Sanofi

*Encadrante Sanofi*

**Jean-Claude, GUILLEMOT**

Cadre supérieur de laboratoire, Sanofi

*Encadrant Sanofi*



À ma famille et mes amis,

À feu mon grand-père, Mennad Hellal, qui  
répétait souvent que « le travail paie toujours »

## REMERCIEMENTS

La réalisation de ce projet de thèse a été possible grâce au soutien de nombreuses personnes à qui je souhaite témoigner ma reconnaissance.

Je remercie tout d'abord Jean-Claude Guillemot qui a été à l'origine de ce projet et qui m'a donné la chance d'y participer et m'a accordé sa confiance et ses conseils. J'espère que tu profites d'une retraite bien méritée.

Je remercie également un million de fois Armelle Buzy qui a été mon encadrante de stage avant d'être l'épaule sur laquelle j'ai pu me reposer pendant ces trois années. Tes conseils, ton enthousiasme pour la science et nos discussions ont fait passer les bons et les moins bons moments de cette thèse avec plus de facilité. Je te souhaite le meilleur dans ta vie professionnelle et personnelle. Je remercie également les Dr. Joëlle Vinh et Jacques Fattaccioli. Ce travail a été une collaboration enrichissante entre trois équipes et environnements très différents dans lesquels j'ai pu apprendre de chacun.

Je souhaiterais remercier particulièrement Marie-Pierre Bousquet et Claire Smadja d'avoir accepté de lire mon travail et d'être mes rapporteuses. Vos remarques ont été très précieuses. J'exprime ma gratitude à Anne Varenne et François Fenaille de m'avoir fait l'honneur de faire partie de mon jury et d'avoir pris le temps de lire mon travail. La discussion avec tous les membres du jury a été très enrichissante et un excellent souvenir de la thèse pour ma part.

Merci à toutes les personnes que j'ai rencontrées et auprès de qui j'ai énormément appris et passé de super moments au labo en commençant par les membres de mes différentes équipes. A l'ESPCI, je voudrais remercier Iman Haddad, Emanuelle Demey, Alexandra Emmanuel, Yann Verdier, Giovanni Chiappetta, Nicolas Eskenazi et les doctorants : Jessica Limaj, Ranjith Kumar, Zeyuan Xu, Jean Bouvet. Merci également aux anciens : Huu Hien Huynh et Massamba N'diaye qui m'a énormément aidée. Aux membres de l'équipe micro-fluidique de l'IPGG, merci aux anciens doctorants Chaohe Hu et Pinon Léa. Aux membres de l'équipe de Sanofi, merci à Anne Remaury, Jean-Luc Zachayus, Sébastien Roudières, Catherine Déon, Corinne Célinain sans oublier ceux qui sont partis à la retraite (Pascal, Stéphanie) pour votre soutien et bon courage à tous pour la suite !

Merci aux membres du bureau du club jeune de la SFSM et aux membres du club jeune avec qui on a organisé plusieurs rencontres scientifiques et passé de super moments en congrès !

Enfin, je souhaite remercier ma famille et mes amis qui ont subi ces 3 années avec moi. Merci à ma sœur Asma et mes parents pour leur soutien moral, rien de tout cela n'aurait été possible sans vous ! Merci à Nicolas, Paul et Chaf d'avoir suivi mes épopées... Merci à la famille Battut et Salsa pour le soutien et les bons moments. Et enfin, merci à mon soutien le plus régulier qui a été là à chaque instant et qui m'a jamais cessé de m'encourager, (bientôt Docteur) Alexandre Battut.



# Sommaire

REMERCIEMENTS.....	3
Sommaire.....	5
Table des figures.....	7
Table des tableaux.....	14
Table des équations.....	15
Liste des abréviations.....	16
I. INTRODUCTION.....	1
Chapitre 1 L'analyse de cellules uniques.....	2
1.A L'hétérogénéité cellulaire.....	3
1.B L'apport des disciplines omiques.....	5
1.C Les principales étapes en protéomique <i>bottom-up</i> classique.....	10
1.D L'analyse des peptides par spectrométrie de masse.....	18
Chapitre 2 Défis et problèmes de la protéomique en cellule unique.....	29
2.A Un problème de taille... ..	30
2.B Les défis de la protéomique classique.....	31
2.C La différence avec l'analyse des ARNs.....	32
2.D Les propriétés des analytes.....	33
Chapitre 3 Les avancées en protéomique de la cellule unique.....	36
3.A Les solutions proposées.....	37
3.B Les débuts de la protéomique en cellule unique.....	47
II. OBJECTIFS DU PROJET DE RECHERCHE.....	61
Chapitre 4 Objectifs.....	61
III. ETUDES ET RESULTATS.....	63
Chapitre 5 Adaptation du protocole à l'analyse de petites quantités d'échantillons.....	64
5.A Manipulation de petites quantités de cellules.....	65
5.B Étude de l'étape de réduction et d'alkylation.....	77
5.C Étude de l'étape de digestion.....	80
5.D Optimisation des conditions d'analyses en LC-MS.....	85
Chapitre 6 Importance du support.....	95
6.A Influence de la géométrie et du matériau des supports.....	96
6.B Comparaison du support micro-puits en PDMS avec les plaques commerciales.....	106
6.C Compatibilité de la plaque micro-puits avec le trieur de cellules.....	108
6.D Apport d'un traitement de surface.....	109

Chapitre 7 La lyse des cellules.....	135
7.A Cas de la lyse cellulaire par traitement thermique .....	136
7.B Cas de la lyse cellulaire par sonication.....	140
7.C La méthode de fabrication influence l'état de surface du support .....	143
IV. CONCLUSION GENERALE.....	145
Chapitre 8 Bilan des résultats et discussion .....	146
8.A Des protéines aux peptides : une série d'étapes remaniées .....	147
8.B Conditions de séparation chromatographique.....	148
8.C Les surfaces de contact : un vrai piège à peptides ? .....	150
8.D La manipulation des cellules : une étape très en amont à fort impact .....	153
8.E Tentatives d'automatisation .....	154
Chapitre 9 Conclusion et perspectives.....	156
V. BIBLIOGRAPHIE .....	157
VI. ANNEXES .....	169
Chapitre 10 Matériels et méthodes.....	169
10.A MICRO-FABRICATION POUR LES ESSAIS DE TRAITEMENTS DE SURFACE.....	169
10.B MICRO-FABRICATION DES SUPPORTS.....	172
10.C PREPARATION DES ECHANTILLONS .....	173
10.D FABRICATION DE LA COLONNE <i>HOMEMADE</i> .....	174
10.E ANALYSE LC-MS.....	174
10.F ANALYSE DES DONNEES.....	175
Chapitre 11 Figures annexes .....	177



## Table des figures

Figure 1 : Schéma d'une tumeur et de son environnement tumoral composé de la matrice extracellulaire, des cellules endothéliales, stromales et immunitaires vasculaires <sup>14</sup> .....	4
Figure 2 : Schéma du principe de l'isolement des cellules par un système de création de gouttelettes par DMF ou « <i>Digital MicroFluidics</i> » <sup>27</sup> .....	5
Figure 3 : Protocole de séquençage des ARNs en cellule unique (sc RNA-Seq). Les différentes étapes sont : 1) isolement de cellules uniques, 2) lyse cellulaire préservant l'ARNm, 3) capture de l'ARNm, 4) transcription inverse de l'ARNm en ADN complémentaire (ADNc), 5) amplification de l'ADNc, 6) préparation des bibliothèques d'ADNc, 7) regroupement des bibliothèques, 8) utilisation d'outils bio-informatiques pour vérifier la qualité et la variabilité des données, 9) utilisation d'outils pour analyser et visualiser les datas. <i>t-SNE t-distributed stochastic neighbour embedding</i> . <sup>29</sup> .....	7
Figure 4 : (a) Protocole d'isolation cellulaire par microfluidique utilisé en séquençage de l'ARN en cellules uniques avec un encodage des ARNm de différentes cellules. (b) Stratégie de comparaison de 3 méthodes de séquençage de l'ARN en fonction des questions biologiques posées (inDrop, 10X ou Drop-seq) <sup>31</sup> .....	7
Figure 5 : 20 populations distinctes observées par regroupement des t-SNE obtenus par le séquençage de l'ARN des cellules uniques de foies humains sains. ( <i>MACs: macrophages, t-SNE: t-distributed stochastic neighbour embedding</i> ) <sup>34</sup> .....	8
Figure 6 : Structure primaire d'un résidu protéique : chaîne polypeptidique.....	9
Figure 7 : Les bases du succès en protéomique sont un équilibre entre l'amélioration de l'instrumentation, de la préparation d'échantillons et des logiciels de traitement de données <sup>42</sup> .....	10
Figure 8 : Exemple des étapes détaillées d'un protocole en protéomique classique.....	10
Figure 9 : (a) Cellule animale et ses organites (b) membrane plasmique formée d'une bicouche lipidique. ....	11
Figure 10 : Classification des différents types de lyses cellulaires.....	11
Figure 11 : Lyse cellulaire avec utilisation de détergent pour détruire la membrane plasmique et libérer les composants intracellulaires.....	12
Figure 12 : Schéma des réactions de réduction (en haut) et d'alkylation (en bas) des protéines.....	13
Figure 13 : Combinaison des enzymes trypsine et Lys-C pour la digestion enzymatique d'une protéine.....	14
Figure 14 : Schéma d'un exemple de montage en HPLC U3000 (Thermo Fisher) (en haut) avec une préconcentration, (en bas) en injection directe.....	17
Figure 15 : Schéma des composants d'un spectromètre de masse.....	18
Figure 16 : Comparaison des méthodes de quantification des protéines dans les cellules uniques <sup>65</sup> .....	19
Figure 17 : Schéma explicatif de la séparation des peptides par chromatographie liquide à leur analyse en spectrométrie de masse en tandem.....	20
Figure 18 : Schéma de fragmentation original de Roepstorff et Fohlmann de 1984 <sup>68</sup> .....	21
Figure 19 : Exemple de DDA versus DIA sur un instrument Orbitrap.....	22
Figure 20 : Schéma d'un Orbitrap HF (ThermoFisher Scientific). ....	23
Figure 21 : L'analyseur Orbitrap composé d'une électrode centrale et d'une électrode axiale. ....	23
Figure 22 : Schéma d'un Orbitrap Tribrid Eclipse (ThermoFisher Scientific). ....	24
Figure 23 : Schéma d'un TIMS TOF Pro (Bruker).....	25

Figure 24 : Représentation de deux stratégies de quantification et leur application dans deux conditions A et B. Les conditions A et B sont comparées et les deux résultats de quantification indiquent une plus grande abondance des peptides dans la condition A <sup>85</sup> .	26
Figure 25 : Stratégie pour les analyses quantitatives par l'utilisation d'isotopes lourds (A) et de marqueurs de masse en tandem (B). Dans la figure, un exemple de 4 canaux est montré avec des isotopes aux m/z égales à 114, 115, 116 et 117.	27
Figure 26 : Structure des marqueurs TMT pour « Tandem mass tags ».	27
Figure 27 : Analyse des données expérimentales des données obtenues par LC-MS/MS.	28
Figure 28 : Conflits en protéomique pour l'obtention d'une analyse à la fois sensible, à haut débit et exhaustive.	31
Figure 29 : Schéma représentant la fraction des protéines quantifiées ou identifiées sur un protéome total par spectrométrie de masse en protéomique.	32
Figure 30 : Représentation de la complexité du protéome comparé au génome ou au transcriptome <sup>98</sup> .	32
Figure 31 : Exemples de surfaces sur lesquelles les protéines et peptides peuvent s'adsorber pendant la préparation des échantillons en protéomique.	35
Figure 32 : Le smoothie représente un protéome total dont l'analyse donne une idée des ingrédients mais sans la profondeur d'analyse, on ne connaîtrait jamais la quantité ou la représentation de chaque fruit du smoothie (métaphore utilisée par Ryan Kelly en conférence).	37
Figure 33 : Mise en place de l'étape de lyse dans des dispositifs microfluidiques montrée dans deux exemples (à gauche) Irimia <i>et al.</i> (à droite) Jen <i>et al.</i> <sup>110,111</sup> .	38
Figure 34 : Puces microfluidiques servant d'interface avec (à gauche) l'ESI par Ramsey <i>et al.</i> (à droite) avec la HPLC par Ji <i>et al.</i>	39
Figure 35 : iProChip permettant un protocole complet de protéomique dans une puce microfluidique (à gauche) photographie (à droite) schéma des étapes du protocole.	40
Figure 36 : Structure du polypropylène.	41
Figure 37 : Structure du PDMS, n représente un monomère.	42
Figure 38 : Schéma de la fabrication d'un dispositif en PDMS en 2 étapes par photolithographie (à gauche) et moulage par réplique (à droite) <sup>122</sup> .	43
Figure 39 : Traitement au plasma à oxygène de la surface du PDMS qui devient oxydée.	45
Figure 40 : Schéma du revêtement d'une surface par une fine couche de polymères pour mieux contrôler les interactions de la surface avec son environnement <sup>148</sup> .	45
Figure 41 : Stratégies de création des brosses de polymères (a) méthodologie « grafting to » (b) méthodologie « grafting from » <sup>156</sup> .	47
Figure 42 : Principe de la cytométrie de masse. Les cellules sont marquées par des anticorps conjugués à des isotopes métalliques de masses différentes. Les cellules sont ensuite séparées en gouttelettes monocellulaires et nébulisées dans un plasma puis analysées par spectrométrie de masse <sup>164</sup> .	49
Figure 43 : Principe du Western Blot en cellule unique a) schéma du gel c) schéma du protocole.	49
Figure 44 : Schéma du protocole dans le dispositif iPAD-1.	51
Figure 45 : Schéma du protocole de protéomique en cellule unique dans le dispositif OAD <sup>171</sup> .	52
Figure 46 : Schéma des méthodologies Scope MS (a), Scope 2 (b) et nPOP (c).	54
Figure 47 : Première génération de puces NanoPOTS (en haut). Schéma du protocole de préparation dans les NanoPOTS (en bas).	55
Figure 48 : Schéma de la puce N2 (à gauche) et du nouveau protocole amélioré (à droite).	55

Figure 49 : Schéma de la proteoCHIP et du protocole appliqué. (a) 16 cellules uniques sont préparées dans le CellenOne®, (b) sont automatiquement mélangées par centrifugation et (c) sont directement mis à l'interface d'un échantillonneur automatique pour l'acquisition. ....	56
Figure 50 : (a) Photographie du Smart Aliquotor (TebuBio). (b) Zoom sur le puits central et les 100 canaux qui l'entourent. (c) Zoom sur le canal qui arrive au puits récepteur n°61. ....	66
Figure 51 : (a) Photographie du marqueur de 1mm de taille à mesurer sous le microscope (grossissement 100x) (1mm taille totale, 100µm taille entre deux graduations longues, 10µm taille entre deux graduations courtes) (b) courbe des correspondances de la taille (en cm) sur la photographie du marqueur à 1mm en fonction du grossissement choisi. ....	67
Figure 52 : Schéma du smart Aliquotor avec le puits central dont le diamètre externe est de 2.77mm et le diamètre interne est de 0.7mm puis un canal de 0.06mm de large et un puits d'arrivée de 1.8mm de diamètre. ....	67
Figure 53 : Différentes étapes de création d'une puce de séparation microfluidique pour l'isolement des cellules. (a) dessin du modèle et choix des dimensions. (b) micro-fabrication commence par l'impression d'un masque pour la photolithographie et processus de photolithographie sur un wafer. (c) coulage du PDMS pour former la puce et poinçonnage des puits puis collage sur une lame de verre pour l'étanchéité. ....	68
Figure 54 : Designs sur AutoCAD de différentes puces (de A à F) avec différentes formes et dimensions pour l'isolement cellulaire. ....	69
Figure 55 : Ouverture du fichier avec le logiciel Klayout et délimitation des designs de puces. ....	70
Figure 56 : (a) Wafer avec les différents modèles. (b) Différentes puces en PDMS fabriquées à partir du wafer. ....	70
Figure 57 : Modèles A, F et C où le liquide (colorant rouge) s'écoule par capillarité du puits central vers tous les puits récepteurs. ....	71
Figure 58 : Modèles D, D2 et E où le liquide (colorant bleu) ne s'écoule pas dans tous les canaux et n'atteint pas tous les puits récepteurs. ....	71
Figure 59 : Test sur deux modèles avec des billes fluorescentes. (a) Zoom sur le puits central dans le modèle A. (b) Zoom sur un puits récepteur dans le modèle F. La barre en rouge indique un marqueur de taille de 100µm. ....	72
Figure 60 : Photographies de l'isolement cellulaire à différentes concentrations par les puces de différents modèles. (a) puits central du modèle B à 10 <sup>5</sup> cellules/ml. (b) canal du modèle A à 10 <sup>5</sup> cellules/ml. (c) puits récepteur du modèle A à 10 <sup>4</sup> cellules /ml. (d) canal du modèle F à 2000 cellules/ml. (e) intersection des canaux du modèle F à 2000 cellules/ml. (f) avancée du liquide dans un canal après injection dans modèle F à 2000 cellules/ml. Les flèches rouges indiquent la présence de cellules. ....	74
Figure 61 : Photographies de puits récepteurs dans deux modèles (C et F). (a) puits récepteur apparaît noir au microscope optique. (b) déchirure de la surface du puits due à la manipulation. (c) imperfections dans le poinçonnage des trous rend difficile la visualisation. ....	74
Figure 62 : Principe du tri cellulaire activé par la fluorescence (FACS). ....	75
Figure 63 : Photographie du trieur de cellule SH-800 (Sony) (a) l'appareil (b) zoom sur l'emplacement où se fait la collection des cellules avec les plaques de diffraction (1), la poubelle (2), l'emplacement de la plaque (3), un porte-plaques (4) et un adaptateur (6). ....	76
Figure 64 : La puce microfluidique et le mode de tri. (a) la puce microfluidique utilisée pour créer (b) un flux de gouttelettes permanent. (c) mode de tri en cellule unique. Seules les gouttelettes contenant des cellules cibles en leur centre sont triées (en gris), les autres sont jetées à la poubelle (en blanc). ....	76

Figure 65 : (a) Stratégie de tri en fonction des marqueurs de taille des cellules (sélection A) et élimination des doublets en faisant une seconde sélection (sélection C). (b) Tri sur une plaque 384 standard avec un tri en mode cellule unique de 200, 50, 10 et 1 cellules. ....	76
Figure 66 : Nombre de protéines et de peptides identifiés à partir de 10ng d'extraits de levure <i>Saccharomyces cerevisiae</i> dans les puces micro-puits PDMS 10µl, avec deux protocoles de préparation des échantillons. ERAD : extrait réduit alkylé et digéré, ED : extrait digéré. Les médianes des 3 réplicats est annoté au-dessus des barres d'intervalles de confiance à 95%. Les * indiquent une p-value < 0.05 obtenue par un t-test non apparié (avec *<0.05, **<0.005, ***<0.0005). ....	77
Figure 67 : Nombre de protéines et peptides identifiés dans des échantillons de 100 ou 1000 cellules Jurkat après une lyse thermique selon 3 protocoles : A = dénaturation + réduction + alkylation + digestion, B = dénaturation + digestion, C = digestion. Les médianes des 2 réplicats sont annotées au-dessus des barres d'intervalles de confiance à 95%. Les * indiquent une p-value < 0.05 obtenue par un Anova à un facteur (avec *<0.05, **<0.005, ***<0.0005). ....	79
Figure 68 : Nombre de protéines et peptides identifiés à partir d'extraits humains ou de levure à différents temps de digestion en minutes (avec t15 = 15minutes, t30 = 30minutes, etc.). ....	81
Figure 69 : Nombre de protéines et peptides identifiés dans (a) 10ng et (b) 2ng analysés d'extraits protéiques de levure digérés par deux ratios de trypsine – Lys C/substrat. La médiane est annotée au-dessus des barres d'intervalles de confiance à 95%. Les * indiquent une p-value < 0.05 selon un t-test non apparié (avec *<0.05, **<0.005, ***<0.0005). ....	82
Figure 70 : Formule brute du DMSO. ....	83
Figure 71 : Nombre de protéines et peptides identifiés (a) depuis un digest de levures, ou (b) après digestion d'un extrait de levure, tous deux dilués dans différentes conditions (avec dilution dans le DMSO). Les médianes des 3 réplicats sont annotées au-dessus des barres d'intervalles de confiance à 95%. Les * indiquent une p-value < 0.05 obtenue par un ANOVA à un facteur (avec *<0.05, **<0.005, ***<0.0005). ....	84
Figure 72 : Représentation graphique du nombre d'identifications protéiques à différentes concentrations de digests de HeLa avec différentes colonnes et configurations. ....	86
Figure 73 : Comparaison de la largeur des pics après extraction de l'ion « 588.3190 » de la BSA en pic de base après séparation sur une colonne de 75µm de diamètre interne (noir) ou sur une colonne de 50µm de diamètre interne (rouge). ....	86
Figure 74 : Représentation du changement de pression de la pompe NC reflétant la pression dans la colonne lors de l'augmentation croissante du débit dans deux réplicats de blancs d'échantillons (a et b). Les marques en pointillés indiquent les changements de débit décrits dans le tableau précédent. ....	88
Figure 75 : Problèmes rencontrés avec les colonnes <i>homemade</i> (a) phase s'échappant d'une colonne fabriquée en laboratoire. (b) observation de « trous » dans la phase créant (c) des épaulements dans la pression lors de la montée de débit qu'il est important de surveiller. Les flèches rouges indiquent les problèmes discutés. ....	88
Figure 76 : Représentation du profil de pression de colonne de séparation lors d'un gradient de 112 minutes dans deux réplicats d'échantillons témoins (a et b). Les marques en pointillés indiquent les changements de gradient décrits dans le tableau précédent. ....	89
Figure 77 : Représentation graphique du nombre d'identifications protéiques à différentes concentrations de digests de HeLa avec une colonne « <i>homemade</i> » et une colonne commerciale en différentes configurations avec deux appareils. ....	91
Figure 78 : Photographie d'une colonne Aurora (Ion Opticks) de 75µm d.i. et 25cm de long, avec à gauche la tête de la colonne (adaptée aux vis nano Viper) et à droite le tip intégré protégé par une capuche qui s'intègre directement à la source nano ESI du Tims TOF Pro (Bruker). ....	91

Figure 79 : Représentation graphique du nombre d'identifications protéiques pour deux quantités d'un digestat peptidique issu de cellules HeLa analysé en LC-MS/MS sur un Tims TOF Pro avec une colonne Aurora 75µm d.i. x25cm par 2 méthodes MS qui diffèrent par leur temps d'accumulation en TIMS (std = 100ms, low = 188ms).....	94
Figure 80 : Gamme de digestats extraits de cellules HeLa analysée en LC-MS/MS sur le Tims-TOF Pro avec une colonne Aurora de 75µm d.i. x 25cm, avec un gradient de séparation de 30min sur plusieurs mois.....	94
Figure 81 : Couverture de séquence pour chaque sous unité de l'hémoglobine après dépôt de 200 ng ou 50 ng sur une plaque MALDI.....	97
Figure 82 : Nombre de protéines et peptides d'extraits de levure <i>Saccharomyces cerevisiae</i> à 50ng/µl digérés pendant 2h (a) à 37°C dans un tube ou après dépôt sur une plaque MALDI OptiTOF (b) à 20°C dans un tube ou après dépôt sur une plaque <i>AnchorChip</i> . Les médianes des réplicats sont annotées au-dessus des barres d'intervalles de confiance à 95%.....	100
Figure 83 : 1 <sup>er</sup> design du support miniaturisé (a) schéma du design (b) photographies du design poinçonné ..	101
Figure 84 : (a) Dessin sur AutoCAD du 1 <sup>er</sup> design de puce selon les dimensions d'un <i>wafer</i> (b) dimensions des puits et espacements entre les puits en µm.....	102
Figure 85 : 2 <sup>ème</sup> design du support miniaturisé (à gauche) photographie de l'usineuse en plein fabrication du moule en laiton (à droite en haut) photographie du moule en laiton et d'une puce en PDMS fabriquée à partir du moule (à droite en bas) photographie du moule en laiton vu de haut et dimensions d'un puits.....	103
Figure 86 : 3 <sup>ème</sup> design de support miniaturisé (a) Dessin du design de moule en laiton à fond V (b) zoom sur la forme des puits (reflet négatif de la plaque PDMS).....	103
Figure 87 : 3 <sup>ème</sup> design du support miniaturisé (en haut) photographie du moule en laiton et d'une plaque micro-puits en PDMS fabriquée à partir de ce moule (en bas) dimensions d'un puits de la plaque.....	104
Figure 88 : (a) Dimensions du puits dans la plaque micro-puits en PP (b) design de la plaque sur le logiciel Fusion (Autodesk) (c) photographie d'une plaque micro-puits en PP à la fin de l'impression 3D (d) zoom sur une plaque micro-puits en PP (PP 10µl) et sa jupe d'ancrage.....	105
Figure 89 : Nombre de protéines et peptides identifiés à partir d'extraits de protéines humaines ou de levure, digérées en plaque commerciale 384-puits standard (PP120) ou dans une plaque micro-puits en PDMS (PDMS10). Les valeurs au-dessus des barres d'intervalles de confiance à 95% représentent les médianes des 12 réplicats. Les * représentent une p-value <0.05 obtenue par un t-test non apparié (avec *<0.05, **<0.005, ***<0.0005).....	107
Figure 90 : Nombre de protéines et peptides obtenus à partir de cellules Jurkat triées sur une plaque micro-puits (PDMS 10µl). Les médianes des 13 réplicats de préparation sont annotées au-dessus des barres d'intervalles de confiance à 95%.....	108
Figure 91 : Structure du polyvinylpyrrolidone (PVP).....	109
Figure 92 : Dessin sur le logiciel AutoCAD des puces créées pour le test du traitement au PVP sur un <i>wafer</i> .....	110
Figure 93 : Design de la puce pour les traitements de surface préliminaires. (a) Design sur le logiciel AutoCAD avec les dimensions des puits créés par photolithographie. (b) Photographie de 3 puces attachées sur une lame de verre. Le cercle rouge indique 1 puce avec son entrée et sa sortie branchées à des capillaires.....	111
Figure 94 : Nombre de protéines, peptides identifiés à partir de (a) 50ng (b) 1.6ng de digestat de HeLa ayant traversé des puces en PDMS traitées (pvp) ou non (ctrl) au PVP comparée au nombre identifiés à partir de digestat de HeLa n'ayant pas traversé de puce (ref). La médiane des réplicats de préparation est indiquée au dessus de chaque barre d'intervalles de confiance à 95%.....	114

Figure 95 : Chromatogramme en pic de base d'un blanc dans une plaque micro-puits non traitée (en haut) ou dans une plaque micro-puits traitée au PVP (en bas). NL indique les intensités maximales de chaque chromatogramme.....	115
Figure 96 : Chromatogramme en pic de base d'un échantillon de HeLa (1.6ng) dans une plaque micro-puits non traitée (en haut) ou dans une plaque micro-puits traitée au PVP (en bas). NL indique les intensités maximales de chaque chromatogramme.....	116
Figure 97 : Chromatogramme en pic de base d'un échantillon de HeLa (50ng) dans une plaque micro-puits non traitée (en haut) ou dans une plaque micro-puits traitée au PVP (en bas). NL indique les intensités maximales de chaque chromatogramme.....	117
Figure 98 : Photographie du <i>wafer</i> du prototype 2.....	118
Figure 99: Étude de l'effet de la présence d'une surface de PDMS en contact avec une solution de digestat tryptique d'extrait protéique total de cellule HeLa à 10ng/μl. Les analyses ont été réalisées en triplicat pour la solution test (bleu sombre, 10ng_ref_1, 2 et 3) et pour la solution qui est passée dans le dispositif en PDMS (bleu clair, pdms_1h-1, 2 et 3). La moyenne des résultats de chaque série est présentée en hachuré .....	119
Figure 100: Analyse différentielle des identifications de peptides après mise en contact avec une surface PDMS. (À gauche) : Distribution des peptides selon leur indice d'hydrophobicité (échelle Kyte-Doolittle). Plus un peptide est hydrophobe, plus cet indice est grand. En vert, les peptides identifiés à partir de la solution de référence (10ng/μl), en orange, les peptides identifiés après avoir séjourné dans le dispositif microfluidique en PDMS pendant 1h. (à droite) : Étude des peptides identifiés en commun en absence (Reference) et en présence d'une surface PDMS (PDMS).....	120
Figure 101 : Distribution des temps de rétention des peptides identifiés dans la solution de référence (orange) et dans la solution au contact du PDMS (bleu). .....	120
Figure 102: Comparaison des analyses deux à deux. Chaque point correspond à un peptide représenté en échelle d'intensité pour les deux conditions comparées. Pour chaque couple, le coefficient de corrélation de Pearson est indiqué. Le cadre en haut à gauche correspond aux analyses sur PDMS, le cadre en bas à droite correspond aux analyses de l'échantillon contrôle. À droite, l'agrandissement d'une des comparaisons montre la position du nuage de points par rapport à la première bissectrice, cas idéal de reproductibilité et d'une corrélation parfaite.....	121
Figure 103 : Évolution du nombre moyen de peptides détectés après passage dans un dispositif en PDMS selon le traitement de surface. L'échantillon de référence à 10ng/μl (ref) a été injecté dans des dispositifs en PDMS non traités (PDMS) ou traités (PVP, PVA, PD1, PD2, PD3, APTES, bêta-octyl, HA). Les chiffres indiquent le nombre moyen de peptides identifiés et passant les filtres de validation sur l'ensemble des analyses réalisées, indépendamment des aliquots. Ceci explique la plus grande variabilité des résultats ref et PDMS qui sont moyennés sur l'ensemble des campagnes d'analyse pour tous les traitements. .	123
Figure 104 : Effet du traitement de la surface du PDMS à la polydopamine sans traitement oxydant, en présence d'O <sub>2</sub> (PD3) sur la population des peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (PD3) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence. ....	124
Figure 105 : Effet du traitement de la surface du PDMS à la polydopamine avec traitement oxydant au peroxyde et HCl concentré (PD1) sur les peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (PD1) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence. ....	125
Figure 106 : Effet du traitement de la surface du PDMS à la polydopamine avec traitement oxydant au périodate de sodium sur la population des peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (PD2) ou non traité (PDMS). À gauche,	

répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence .....	125
Figure 107 : Effet du traitement de la surface du PDMS au polyacétate de vinyle après activation avec traitement oxydant au peroxyde et HCl concentré (PVA) sur la population des peptides identifiés sans passage (ref) ou avec passage dans le dispositif microfluidique traité (PVA) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence .....	126
Figure 108 : Effet du traitement de la surface du PDMS au polyvinylpyrrolidone PVP après activation au plasma O <sub>2</sub> (PVP) sur la population des peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (PVP) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence .....	126
Figure 109 : Effet du traitement de la surface du PDMS au (3-aminopropyl) triéthoxysilane (APTES) après activation au plasma O <sub>2</sub> sur la population des peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (APTES) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence .....	127
Figure 110 : Évolution du nombre de peptides identifiés pour chaque enregistrement avec ou sans mise en contact avec une surface PDMS- bêta-octylglucoside.....	127
Figure 111: Effet du traitement de la surface du PDMS au bêta-octylglucoside (betaoctyl) après activation H <sub>2</sub> O <sub>2</sub> , HCl concentré sur la population des peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (betaoctyl) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence .....	128
Figure 112 : Effet du traitement de la surface du PDMS à l'acide hyaluronique (HA)après activation H <sub>2</sub> O <sub>2</sub> , HCl concentré sur la population des peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (HA) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence. ....	128
Figure 113 : Évolution de la distribution des peptides selon les paramètres physicochimiques définis par Cruciani <sup>213, 214</sup> . L'axe des x représente des échelles PP1 : polarité, PP2 : Hydrophobicité et taille, PP3 : capacité à former des liaisons-H. ....	130
Figure 114 : Diagramme de dispersion des analyses dans les canaux PDMS traités au PVP. Les intensités obtenues pour chaque peptide sont comparées enregistrement par enregistrement. Le coefficient de corrélation de Pearson est indiqué en haut à gauche de chaque comparaison. ....	131
Figure 115 : Nombre de protéines et peptides identifiés à partir de protéines extraites de levure digérées sur la plaque micro-puits en PDMS. Les barres d'intervalles de confiance à 95% sont annotées des médianes des réplicats. Les * indiquent une p- value <0.05 obtenue par un t test non apparié (avec *<0.05, **<0.005, ***<0.0005). ....	132
Figure 116 : Nombre de protéines et peptides identifiés après digestion de protéines extraites de cellules humaines ou de levure déposées sur une plaque standard 384 puits ou une plaque micro-puits traitées ou non au PVP. Les médianes des réplicats sont indiquées au-dessus des barres d'intervalle de confiance à 95%. Les * indiquent une p-value significative après t-test non apparié (avec *<0.05, **<0.005, ***<0.0005). ....	133
Figure 117 : Nombre de protéines et peptides après analyse de 2, 10 et 40 cellules traitées sur une plaque standard ou une plaque micro-puits en PDMS. Les médianes des réplicats sont indiquées au-dessus de chaque barre d'intervalles de confiance à 95%. Les * indiquent une p-value significative après t-test non apparié (avec *<0.05, **<0.005, ***<0.0005). ....	137

Figure 118 : Nombre de protéines et peptides après analyse de 2, 10 et 40 cellules traitées sur une plaque standard ou une plaque micro-puits en PDMS. Les médianes des réplicats sont indiquées au-dessus de chaque barre d'intervalles de confiance à 95%. Les * indiquent une p-value significative après t-test non apparié (avec *<0.05, **<0.005, ***<0.0005). .....	139
Figure 119 : Nombre de protéines et peptides identifiés pour 2, 10 et 40 cellules analysées après une lyse mécanique par sonication. Le nombre d'identification médian est annoté au-dessus de chaque barre d'intervalles de confiance à 95%. Les * indiquent une p-value significative après un Anova à un facteur (avec *<0.05, **<0.005, ***<0.0005). .....	142
Figure 120 : Nombre de protéines identifiées après dépôt de différentes concentrations d'un digestat de cellules Thp-1 sur une plaque standard 384 puits avec 58µl/puits (PP 58µl) et une plaque micro-puits PP 10µl (PP10µl). Les médianes des réplicats de préparation sont indiquées au-dessus de chaque barre d'intervalle de confiance à 95%. Les * indiquent une p-value significative après t-test (avec *<0.05, **<0.0005, ***< 0.0005). .....	144
Figure 121 : Photographie de l'observation sous microscope optique d'un puits de la plaque micro-puits en PP imprimée en 3D sur lequel on peut observer une surface non lisse avec des cercles concentriques montrant les effets de l'impression 3D donnant la surface rugueuse. ....	144
Figure 123 : photographie à t0 et à t2h de la plaque MALDI et <i>AnchorChip</i> utilisées pour la digestion des échantillons d'extraits protéiques de levure à 37°C et à 20°C avec un système <i>homemade</i> pour limiter l'évaporation consistant en une boîte fermée incluant des bouchons d'eau à 100°C. ....	177
Figure 124 : Nombre de protéines, peptides et PSM identifiés en LC-MS/MS sur un Q Exactive HF à partir de différentes quantités d'un digestat de levure <i>Saccharomyces cerevisiae</i> . La moyenne des 3 réplicats est annotée au-dessus des barres d'écart type standard. ....	177
Figure 125 : spectres théoriques de polymères (en haut) PEG et (en bas) PVP .....	178
Figure 126 : TICs des échantillons, dans l'ordre de haut en bas, un échantillon blanc ayant traversé la puce non traitée, un échantillon blanc ayant traversé la puce traitée au PVP, un échantillon de digestat de HeLa 1.6ng ayant traversé la puce non traitée, un échantillon de digestat de HeLa 1.6ng ayant traversé la puce traitée au PVP, un échantillon de digestat de HeLa 50ng ayant traversé la puce non traitée, un échantillon de digestat de HeLa 50ng ayant traversé la puce traitée au PVP. ....	179
Figure 129 : Dimensions des différents designs de puce d'isolation. ....	180
Figure 130 : Étapes de la fabrication d'une colonne en laboratoire. (a) installation du fritté (b) insertion de la phase dans le capillaire grâce à une pompe de gaz (c) création d'un embout pour la connexion avec la MS (d) sertissage de la colonne pour la connexion avec la LC. ....	181
Figure 131 : schéma de fonctionnement de l'Echo ( <i>Acoustic Liquid Handler</i> ). Le transducteur passe sous la plaque source et émet une énergie acoustique répétée qui produit un ménisque.....	181

## Table des tableaux

Tableau 1 : Quantités en protéines et leur nombre de cellules correspondant.....	30
Tableau 2 : Mesure des canaux du smart Aliquoter photographiés par le microscope confocal. ....	67
Tableau 3: Diamètres des puits centraux et récepteurs en µm.....	69
Tableau 4 : Gradient d'augmentation du débit en fonction du temps.....	87
Tableau 5 : Gradient d'élution de 112 minutes et % d'acétonitrile utilisée pour l'élution. ....	89
Tableau 6 : Gradient d'élution de 30 ou 100 minutes (runs 60min ou 130min respectivement) et % d'acétonitrile utilisée pour l'élution. ....	92



Tableau 7 : Nombre de peptides identifiés en nano LC-MS/MS après dépôt du digestat d'hémoglobine humaine (200ng ou 50ng) sur une plaque MALDI comparé au nombre de peptides théoriques obtenus après digestion par la trypsine avec 2 <i>miscleavages</i> autorisés et compris entre 500 et 4000 m/z. ....	97
Tableau 8 : Dimensions caractéristiques des dispositifs microfluidiques avec surface totale de contact PDMS- Solution et volume d'un canal arrondi à l'inférieur.....	119
Tableau 9 : Liste des traitements de surface PDMS évalués. PVP : polyvinylpyrrolidone, PVA : Poly-acétate de vinyle, PD : Polydopamine, APTES : (3-Aminopropyl)triethoxysilane, Bêta-octyl : bêta-octyl-glucoside, HA : Hyaluronic acid. ....	122
Tableau 10 : Comparaison de la lyse thermique de cellule Jurkat sur plaque PP 120 $\mu$ L commerciale et plaque PDMS 10 $\mu$ L <i>homemade</i> (12 répliqués/condition).....	136
Tableau 11 : Comparaison de la lyse thermique de cellule THP1 sur plaque PP 120 $\mu$ L commerciale et plaque PDMS 10 $\mu$ L <i>homemade</i> (18 répliqués/condition).....	138
Tableau 12 : Comparaison de la lyse thermique de cellule THP1 sur plaque PP 120 $\mu$ L commerciale et plaque PDMS 10 $\mu$ L <i>homemade</i> (18 répliqués/condition). (*) différence significative. ....	140

## Table des équations

Équation 1 : Équation du débit en fonction du diamètre interne de la colonne avec $f$ = débit, $d_c$ = diamètre interne de la colonne, $\mu$ = vitesse linéaire de la phase mobile ( $L$ = longueur de la colonne et $t_m$ = temps d'élution d'un analyte non retenu dans la colonne) et $\epsilon$ = porosité de la colonne. ....	15
Équation 2 : Équation du calcul de la mobilité ionique. ....	24
Équation 3 : L'équation de Michaelis-Menten (ou de Michaelis-Menten-Henri) permet de décrire la cinétique d'une réaction catalysée par une enzyme agissant sur un substrat unique pour donner irréversiblement un produit.....	34
Équation 4 : L'équation de Michaelis-Menten pour une concentration de substrat très faible .....	34

## Liste des abréviations

Dans ce manuscrit, nous avons pris le parti d'employer certaines abréviations anglaises en usage

APTES : (3-Aminopropyl)tri-ethoxy-silane	MRM : multiple reaction monitoring
ASMS : American society of mass spectrometry	MS : mass spectrometry
AUC : area under curve	MS/MS : tandem mass spectrometry
BEH : Bridged ethylene hybrid	NGS : next generation sequencing
BSA : bovine serum albumin	PASEF : parallel acquisition serial fragmentation
CAD : collisionally-activated dissociation	PBS : phosphate buffered saline
CID : collision induced dissociation	PCR : polymerase chain reaction
CCS : collisional cross section	PD : polydopamine
CE : capillary electrophoresis	PEG : polyethylene glycol
DDA : data dependent acquisition	PRM : parallel reaction monitoring
DDM : n-dodécyl- $\beta$ -d-maltoside	PTM : post translational modifications
DI : direct injection	PVA : polyvinyl alcohol
DIA : data independent acquisition	PVP : polyvinylpyrrolidone
DMF : digital microfluidics	Q Ex HF : Q Exactive HF
DTT : dithiothreitol	RF : radiofrequency
EDTA : ethylene diamine tetra acetic acid	RPLC : reversed phase liquid chromatography
ESI : electrospray ionization	RT : retention time
ETD : electron transfer dissociation	scRNA seq : single cell RNA sequencing
FACS : fluorescence activated cell sorting	SDS : sodium dodecyl sulphate
FASP : filter aided sample preparation	SID : surface induced photodissociation
HCD : higher energy collisional dissociation	SIP : surface-initiated polymerization
HEMA : 2-hydroxy-ethyl methacrylate	SWATH : Sequential Acquisition of All Theoretical fragment-ion spectra
HPLC : high performance liquid chromatography	t-SNE : t-distributed stochastic neighbour embedding
IAA : iodoacetamide	TCEP : tri-carboxy-ethyl-phosphine
ICR : ion cyclotron resonance	TIMS : trapped ion mobility spectrometry
IM : ion mobility	TMT : tandem mass tags
IRMPD : infrared multi-photon photodissociation	TOF : time of flight
IT : ion trap	UMI : unique molecular identifiers
LC : liquid chromatography	UV : ultra-violet
LPS : lipo-polysaccharide	XIC : extracted ion chromatogram
MAC : macrophages	
MALDI : matrix assisted laser desorption ionization	

Dans ce manuscrit, d'autres abréviations sont en français :

ADN : acide désoxyribonucléique	d.i. : diamètre interne
ARN : acide ribonucléique	PC : préconcentration
ARNm : ARN messenger	PDMS : polydiméthylsiloxane
ADNc : ADN complémentaire	PP : polypropylène
CAD : collision activée	PS : polystyrène



**THÈSE DE DOCTORAT**  
DE L'UNIVERSITÉ PSL

## I. INTRODUCTION

## Chapitre 1 L'analyse de cellules uniques

<b>Chapitre 1 L'analyse de cellules uniques .....</b>	<b>2</b>
<b>1.A L'hétérogénéité cellulaire .....</b>	<b>3</b>
<b>1.B L'apport des disciplines omiques.....</b>	<b>5</b>
1.B.1. Exemple du séquençage de l'ARN en cellule unique.....	6
1.B.2. Applications biologiques.....	8
1.B.3. Limites .....	8
1.B.4. La protéomique et l'importance des protéines.....	8
<b>1.C Les principales étapes en protéomique <i>bottom-up</i> classique .....</b>	<b>10</b>
1.C.1. La lyse cellulaire.....	11
1.C.2. La préparation des protéines .....	13
1.C.3. La digestion enzymatique.....	14
1.C.4. Enrichissement des protéines .....	14
1.C.5. Les techniques séparatives miniaturisées .....	14
<b>1.D L'analyse des peptides par spectrométrie de masse .....</b>	<b>18</b>
1.D.1. La fragmentation des peptides.....	19
1.D.2. Les modes d'acquisition en spectrométrie de masse.....	21
1.D.3. Les instruments utilisés dans ce projet .....	22
(a) Les Orbitrap, ThermoFisher Scientific .....	22
(b) Le Tims TOF Pro, Bruker.....	24
1.D.4. La quantification des protéines .....	26
1.D.5. Le traitement des données de spectrométrie de masse.....	28

## INTRODUCTION : L'analyse de cellules uniques

Les défis analytiques les plus importants pour la compréhension des systèmes biologiques et qui justifient le développement d'analyses sur cellules individuelles sont liés à l'hétérogénéité cellulaire et la rareté de certains échantillons biologiques.

### 1.A L'hétérogénéité cellulaire

La biologie cellulaire est l'étude structurale et fonctionnelle de la cellule, entité mise en évidence en 1665 par Robert Hooke <sup>1</sup>. L'étude des cellules a donné lieu à des découvertes révolutionnaires pour la compréhension de l'ADN, du cancer, du vieillissement ou de la biologie du développement embryonnaire <sup>2</sup>. La théorie cellulaire, développée en 1839 par Matthias Jakob Schleiden et Theodor Schwann, stipule que la cellule est l'unité fondamentale de structure, de fonction et de reproduction du Vivant <sup>3,4</sup>.

Les systèmes cellulaires tels que les tissus, les cultures cellulaires et plusieurs maladies sont constitués d'une variété de cellules ayant des propriétés moléculaires et fonctionnelles distinctes <sup>5,6</sup>. Cette hétérogénéité cellulaire a été étudiée à différents niveaux moléculaires et fonctionnels et permet à un organe ou à un organisme de se différencier et d'adapter sa réponse aux conditions environnementales <sup>7</sup>. En plus des types cellulaires, il a été également prouvé que des cellules d'un même tissu ou même d'un même type cellulaire peuvent être très hétérogènes et présenter des phénotypes différents. La caractérisation de ces différences à l'échelle de ces sous-populations de cellules est essentielle pour comprendre comment un ensemble de cellules complexe traite les informations et réagit aux perturbations <sup>7</sup>.

En biologie, une distribution de comportements cellulaires peut être estimée à partir des observations d'un petit nombre de cellules sur une longue période, ou d'un grand nombre de cellules à un seul moment <sup>7</sup>. En théorie, ces estimations peuvent donner des informations similaires si les comportements d'une seule cellule moyennés dans le temps et les comportements de la population moyennés dans l'ensemble sont équivalents <sup>8</sup>. En pratique, les analyses de moyennes de populations à un instant donné sont privilégiées et ont permis l'identification des composants et des interactions au sein de réseaux complexes de métabolisme, de signalisation et de transcription. Ces essais se basent sur l'hypothèse que la moyenne de la population cellulaire reflète le mécanisme biologique dominant dans les cellules individuelles, et que cette population est globalement homogène. Cependant, ces informations moyennées à partir de milliers voire de millions de cellules ne définissent pas forcément leurs comportements individuels <sup>7</sup>.

En effet, une même population cellulaire est souvent composée de plusieurs sous-populations de différents phénotypes. Les cellules endothéliales, par exemple, jouent un rôle essentiel dans de nombreux processus physiologiques et présentent une hétérogénéité remarquable dans leur structure et leur fonction, dans les tissus sains ou malades <sup>9</sup>. D'autre part, de petits réservoirs de cellules souches dormantes ont été identifiés au sein de populations de cellules souches hématopoïétiques plus importantes. La réactivation rapide de ces sous-populations en cas de blessure joue un rôle crucial dans le rétablissement de l'homéostasie <sup>10</sup>. Le cancer ainsi que d'autres maladies comme la stéatohépatite du foie ont été établis comme des maladies très hétérogènes <sup>6,11,12</sup>. La population de cellules composant un cancer est hétérogène aux niveaux génétique, épigénétique et phénotypique <sup>13</sup>. D'ailleurs, le microenvironnement tumoral hétérogène influence la réponse thérapeutique (Figure 1) <sup>14</sup>. L'infiltrat immunitaire peut inclure de multiples types de cellules, ces populations cellulaires

## INTRODUCTION : L'analyse de cellules uniques

peuvent avoir des fonctions pro- et antitumorales et peuvent varier dans leur statut d'activation et leur localisation dans la tumeur<sup>14</sup>. Plusieurs cellules cibles de détection comme les cellules des tissus tumoraux, les cellules circulantes dans le sang, les cellules souches cancéreuses et bien d'autres ont été listées par Qian *et al.* et confirmées par Cohen *et al.*, qui révèlent cette hétérogénéité importante à la fois pour le diagnostic et le traitement du cancer<sup>15,16</sup>. En effet, pour le diagnostic, la compréhension de l'hétérogénéité unicellulaire est essentielle pour garantir une utilisation précise des biomarqueurs<sup>15</sup>. Par exemple, les fibroblastes associés au cancer peuvent avoir une plasticité importante et diverger en ce qui concerne leur statut d'activation, leur localisation dans le tissu, leur réponse au stress et leur origine<sup>14</sup>. Ces niveaux d'hétérogénéité, en plus de la variabilité inter-patient, rendent difficile un traitement spécifique et ciblé<sup>17</sup>. La compréhension de l'hétérogénéité unicellulaire peut aider à choisir les médicaments appropriés et à améliorer l'efficacité du traitement<sup>15</sup>. Même au sein de populations clonales, dans des conditions contrôlées, les réponses individuelles des cellules cancéreuses aux médicaments peuvent varier considérablement et des états de résistance aux traitements peuvent apparaître<sup>6,16,18</sup>.

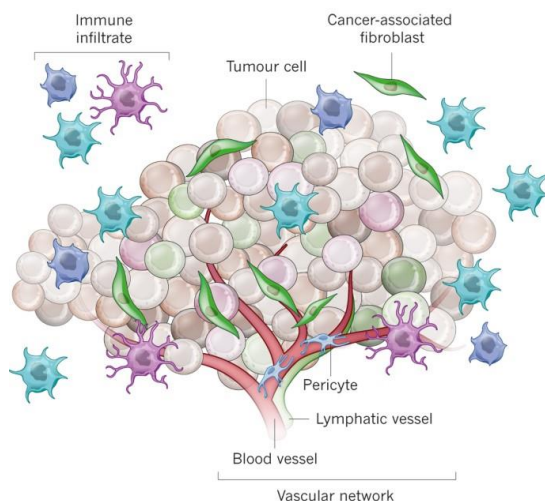


Figure 1 : Schéma d'une tumeur et de son environnement tumoral composé de la matrice extracellulaire, des cellules endothéliales, stromales et immunitaires vasculaires<sup>14</sup>.

Il faut cependant distinguer l'hétérogénéité qui provient du bruit de fond de détection, qui pourrait n'avoir aucune fonction importante, de celle qui a une fonction biologique<sup>19,20,21</sup>. Des analyses à haute résolution aux niveaux cellulaire, transcriptomique, protéomique ou métabolomique, se basant sur des techniques d'analyses innovantes en cellule unique, mettent en évidence des hétérogénéités à l'origine de réponses biologiques allant jusqu'au destin cellulaire<sup>22,23</sup>.

## 1.B L'apport des disciplines omiques

Les analyses en cellule unique se sont beaucoup développées durant les dix dernières années et ont permis d'explorer cette hétérogénéité cellulaire. Les analyses de génomes, transcriptomes ou protéomes de millions de cellules s'appellent analyses en « *bulk* » contrairement aux analyses en cellules uniques. Ces dernières offrent un vaste potentiel pour faire progresser notre compréhension de la biologie et des maladies <sup>24</sup>. Les méthodes manuelles initiales permettaient l'analyse de quelques cellules.

Par ailleurs, la microfluidique est un domaine de recherche qui s'intéresse à la miniaturisation de systèmes fluidiques (système de production de gouttelettes par exemple) qui seront utilisés dans des applications biologiques, physiques ou mécaniques. En biologie, le développement des DMF pour « *Digital MicroFluidics* » qui met en jeu la microfluidique de gouttes est utilisé pour le séquençage de l'ADN et de l'ARN, leur quantification, l'étude structurale des protéines, le screening de médicaments, et les analyses en cellules uniques <sup>25</sup>. La microfluidique en gouttelettes se base sur la non-miscibilité de l'huile et des solvants aqueux comme le PBS dans lesquels baignent les cellules. Les cellules sont amenées à débit constant à un carrefour où l'eau croise l'huile et cette dernière forme une coupure entre les gouttelettes d'eau. Les cellules sont en faible concentration de départ. Ce phénomène permet de créer des gouttelettes d'eau dans l'huile dans lesquelles les cellules sont isolées à une cellule par goutte (Figure 2) <sup>26</sup>.

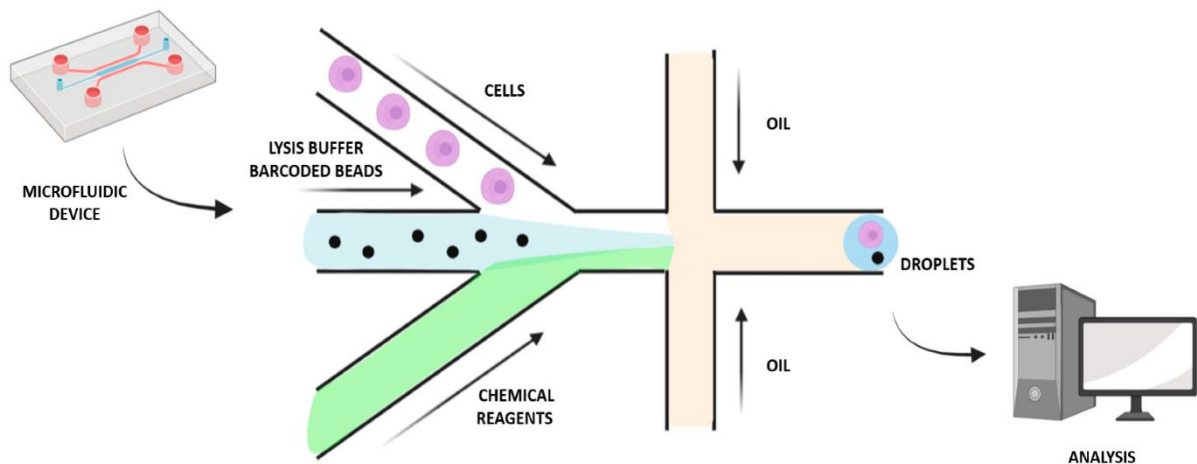


Figure 2 : Schéma du principe de l'isolement des cellules par un système de création de gouttelettes par DMF ou « *Digital MicroFluidics* » <sup>27</sup>

Le développement de ces dispositifs microfluidiques intégrés et l'introduction de robots de manipulation de liquides dans le processus ont permis de monter le nombre de cellules à plusieurs milliers, puis de l'augmenter encore avec les technologies des nano-gouttelettes et des pico-puits. L'introduction d'un code-barres in situ a permis de garantir une bonne traçabilité pour des centaines de milliers de cellules et les derniers développements technologiques permettent d'intégrer la localisation spatiale dans les coupes de tissus des informations transcriptomiques <sup>24</sup>.

La nouvelle génération de séquençage de l'ADN et de l'ARN se base donc sur les avancées de la microfluidique en gouttelettes et permet aujourd'hui d'établir le profil génétique de milliers de cellules en une seule expérience. Il existe plusieurs méthodes, basées sur des puces microfluidiques commerciales ou sur des plaques multi-puits qui dépendent de la question



biologique de départ et de l'échelle de l'expérience et diffèrent en sensibilité et en coût. Elles permettent de détecter des types cellulaires rares ou de déconvoluer des populations très hétérogènes comme des tissus entiers <sup>28</sup>.

### 1.B.1. Exemple du séquençage de l'ARN en cellule unique

Il existe plusieurs méthodologies pour le séquençage de l'ARN en cellule unique mais la plupart comportent les différentes étapes suivantes (Figure 3) <sup>29</sup>. Brièvement, la première étape est constituée par l'isolement de cellules viables et uniques depuis un tissu ou une culture cellulaire d'intérêt. Cela peut se faire de plusieurs manières allant de la microfluidique en gouttelettes à la cytométrie en flux ou FACS (pour « *fluorescence activated cell sorting* ») ou par microdissection par laser. Il faut ensuite lyser chaque cellule d'une façon qui préserve l'ARN messager (ARNm). Des sondes poly T sont utilisées pour capturer tous les brins d'ARNm polyadénylés par leur queue poly A. Cette capture évite l'analyse des ARNs ribosomiaux qui requièrent un protocole plus spécifique pour leur analyse. En utilisant une reverse transcriptase, l'ARNm est converti en ADN complémentaire (ADNc) qui sera ensuite amplifié par plusieurs cycles de PCR (pour « *polymerase chain reaction* »). À l'étape de transcription inverse, d'autres séquences nucléotidiques peuvent être ajoutées aux amorces de la transcriptase comme des identificateurs moléculaires uniques (UMIs pour « *unique molecular identifiers* ») afin de marquer chaque ARN messager et préserver les informations sur leur origine cellulaire <sup>30</sup>. Par la suite, les ADNc amplifiés et marqués de chaque cellule sont regroupés et séquencés par NGS pour « *next generation sequencing* » grâce à la préparation de bibliothèques, de plateformes de séquençage et d'outils d'alignement génomique semblables à ceux utilisés pour les échantillons en analyse en masse. L'analyse bio-informatique des résultats obtenus par séquençage de l'ARN en cellule unique s'est rapidement développée, pour vérifier le contrôle qualité des données et leur variabilité intra et inter échantillons. Des méthodes de regroupement hiérarchique ou non sont utilisées pour interpréter ces données <sup>29</sup>.

Les UMIs cités plus haut, agissent comme des codes-barres et permettent de multiplexer les analyses et d'analyser plusieurs échantillons à la fois. Ainsi, des puces commerciales basées sur la microfluidique de goutte et intégrant ces codes-barres ont été développées pour le séquençage à haut débit. Aujourd'hui, il existe plusieurs plateformes pour le séquençage d'ARN en cellule unique à haut débit basé sur la microfluidique en gouttelettes. Zhang *et al.* ont comparé trois de ces plateformes, à savoir « 10X-Genomics », « Drop-Seq » et « in Drop », en termes d'abondance de cellules, de flexibilité, de sensibilité, de précision et de coût et ont déterminé que le choix dépendait de ces différents paramètres et de la question biologique des chercheurs (Figure 4Figure 3) <sup>31,32</sup>.

## INTRODUCTION : L'analyse de cellules uniques

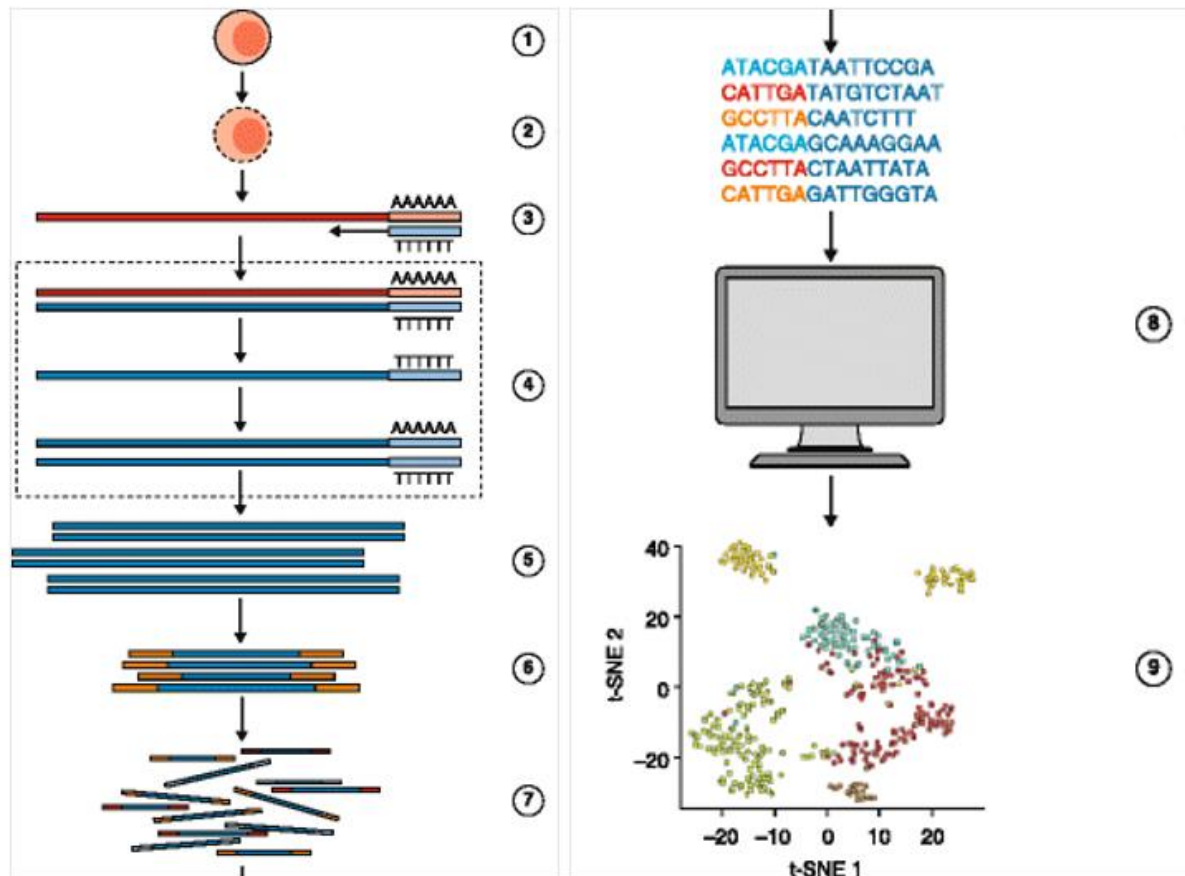


Figure 3 : Protocole de séquençage des ARNs en cellule unique (sc RNA-Seq). Les différentes étapes sont : 1) isolement de cellules uniques, 2) lyse cellulaire préservant l'ARNm, 3) capture de l'ARNm, 4) transcription inverse de l'ARNm en ADN complémentaire (ADNc), 5) amplification de l'ADNc, 6) préparation des bibliothèques d'ADNc, 7) regroupement des bibliothèques, 8) utilisation d'outils bio-informatiques pour vérifier la qualité et la variabilité des données, 9) utilisation d'outils pour analyser et visualiser les datas. *t-SNE t-distributed stochastic neighbour embedding.*<sup>29</sup>

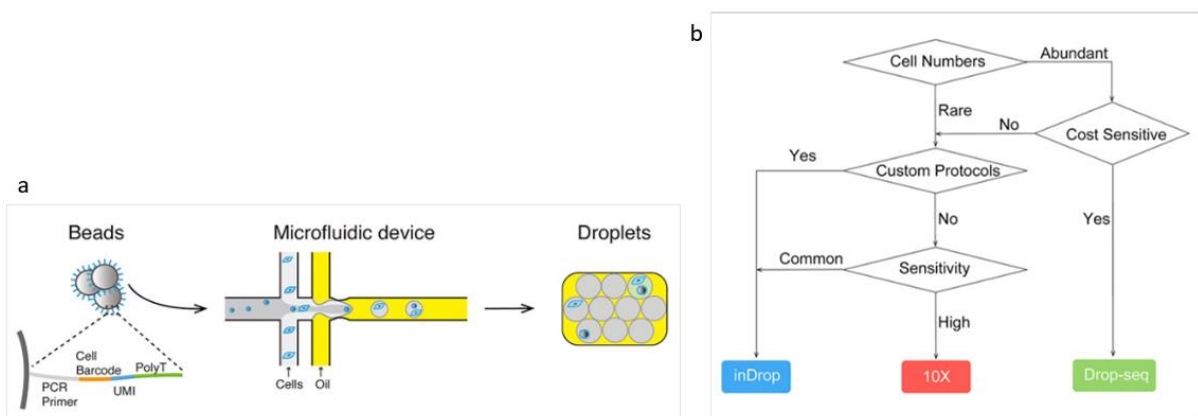


Figure 4 : (a) Protocole d'isolation cellulaire par microfluidique utilisé en séquençage de l'ARN en cellules uniques avec un encodage des ARNm de différentes cellules. (b) Stratégie de comparaison de 3 méthodes de séquençage de l'ARN en fonction des questions biologiques posées (inDrop, 10X ou Drop-seq)<sup>31</sup>.

## INTRODUCTION : L'analyse de cellules uniques

### 1.B.2. Applications biologiques

Afin de mieux comprendre l'hétérogénéité cellulaire de la moëlle osseuse, en 2014, Jaitin *et al.* ont développé une méthode de clustering hiérarchique afin de classer des cellules spléniques, analysées en séquençage d'ARN en cellule unique, en groupes transcriptionnellement distincts. En mimant une infection bactérienne par lipopolysaccharide (LPS), ils ont également identifié des gènes de réponse spécifiques à chaque type de cellule ainsi que des réponses générales à toutes les sous-populations <sup>33</sup>.

De la même façon, en utilisant le séquençage de l'ARN en cellule unique, MacParland *et al.* ont prouvé en 2018 l'existence de 20 populations distinctes dont 5 différents clusters d'hépatocytes au sein des cellules hépatiques humaines (Figure 5) <sup>34</sup>.

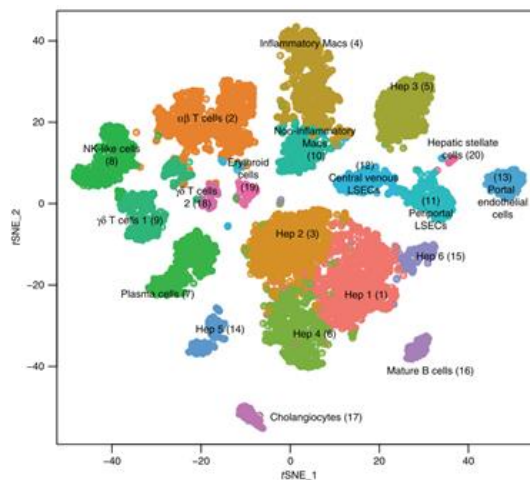


Figure 5 : 20 populations distinctes observées par regroupement des t-SNE obtenus par le séquençage de l'ARN des cellules uniques de foies humains sains. (MACs: macrophages, t-SNE: t-distributed stochastic neighbour embedding) <sup>34</sup>.

### 1.B.3. Limites

Cependant des limitations techniques font que les ARN, notamment non codants, bactériens ou en très faible quantité, ne sont pas tous séquencés par ces méthodes d'analyse. De plus, ces mesures transcriptomiques n'expliquent pas l'entièreté de l'expression génique. Les résultats de séquençage de l'ARN en cellule unique ne décrivent pas directement le phénotype moléculaire d'une cellule, c'est-à-dire la nature et concentration des protéines présentes qui définissent ses caractéristiques biochimiques. Ils ne peuvent capturer l'hétérogénéité protéomique au niveau post-traductionnel qui peut déterminer le comportement cellulaire <sup>28</sup>.

### 1.B.4. La protéomique et l'importance des protéines

Les protéines sont des acteurs fonctionnels et structuraux clés de la cellule qui traduisent la fonction potentielle des gènes via des catalyses enzymatiques, des signalisations moléculaires ou encore des interactions physiques <sup>35</sup>. Plus de 400 modifications post traductionnelles de ces protéines, modifient leur fonction, régulent divers processus cellulaires et sont des biomarqueurs très importants dans le cadre de plusieurs maladies <sup>36</sup>. Les protéines sont des macromolécules <sup>37</sup> dont la structure primaire est un enchaînement de plusieurs chaînes polypeptidiques constituées de l'enchaînement de résidus d'acides aminés liés entre eux par des liaisons amides (groupe -CONH-) (Figure 6). La chaîne polypeptidique

## INTRODUCTION : L'analyse de cellules uniques

comporte une extrémité N-terminale (groupe -NH<sub>2</sub>) et une extrémité C-terminale (groupe -COOH). Un acide aminé est constitué d'une fonction amine, d'un groupe carboxyle et d'un groupe caractéristique variable R formant ainsi un répertoire de 20 acides aminés.

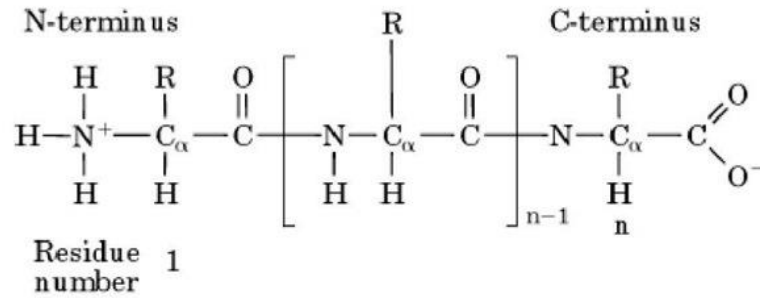


Figure 6 : Structure primaire d'un résidu protéique : chaîne polypeptidique.

Parmi les diverses approches "omiques" en sciences, la protéomique étudie l'ensemble des protéines (le protéome) en visant à déterminer leurs identités, leurs modifications post traductionnelles, leurs quantifications et leurs localisations<sup>38</sup>. Le protéome est dynamique et son étude peut révéler des processus moléculaires non identifiables à l'échelle du génome. Les connaissances du génome seules ne sont pas suffisantes pour une caractérisation moléculaire de la cellule : il faut donc étudier le génome, le transcriptome et le protéome en parallèle<sup>38</sup>. Si le génotype d'une cellule donnée délivre des informations sur un type de cellule potentiel, son phénotype reflète sa fonction selon son environnement et peut varier d'une cellule à l'autre pour un génotype donné. L'étude du phénotype est liée à l'étude de son contenu protéique. Elle pourrait aider au suivi de l'activation cellulaire, de la différenciation ou de la classification des cancers, par exemple.

L'analyse globale de l'expression des protéines a connu une évolution fulgurante grâce au développement des technologies pour la protéomique au cours des 20 dernières années<sup>35</sup>. Les analyses protéomiques requièrent des stratégies analytiques très sensibles et spécifiques pour permettre une identification discriminante et sans ambiguïté d'espèces très peu abondantes<sup>39</sup>. La sensibilité analytique représente la plus petite quantité de substance dans un échantillon qui peut être mesurée avec précision par un test et la spécificité analytique fait référence à la capacité d'un test à mesurer une substance, plutôt que d'autres, dans un échantillon<sup>39</sup>. En outre, l'identification et la quantification des protéomes cellulaires nécessitent de travailler avec une gamme dynamique élevée s'étendant sur 10 ordres de grandeur, allant des 22 protéines les plus abondantes dans la cellule à plusieurs milliers d'autres protéines à des concentrations beaucoup plus faibles. Plus exactement, dans les tissus ou les lignées cellulaires de mammifères, cet ordre de grandeur varie entre 6 et 9 ordres de grandeur. Pour les fluides comme le plasma ou le sang, cet ordre de grandeur peut aller jusqu'à 12 avec l'albumine représentant plus de 50% des protéines présentes dans ces échantillons. Or, les protéines peu abondantes peuvent être les plus intéressantes en termes fonctionnels pour le diagnostic précoce des maladies<sup>40</sup>.

Les approches unicellulaires tentent également de surmonter la disponibilité limitée des échantillons, en particulier les biopsies de tumeurs, où il est nécessaire de travailler avec une petite quantité de matériel. Il est déjà possible d'éviter ces méthodes invasives et de rechercher les cellules tumorales circulant dans le sang. Or, les fluides biologiques comme le

## INTRODUCTION : L'analyse de cellules uniques

plasma contiennent la plupart des protéines humaines solubles ainsi que des protéines dérivées de virus, de bactéries et de champignons à des concentrations très variables <sup>41</sup>.

Une expérience réussie en protéomique nécessite l'intégration d'une bonne préparation des échantillons, d'une bonne instrumentation et de bons logiciels (Figure 7) <sup>42</sup>. Pour atteindre une détection optimale, il est donc crucial d'optimiser et de maîtriser la préparation des échantillons en amont de l'analyse qui influence le temps, la robustesse et le coût de l'analyse aussi bien que les résultats. La clé est une optimisation des 3 domaines en parallèle afin d'assurer une amélioration de la détection des peptides et des protéines. La mise en œuvre de ces stratégies diffère selon la complexité de l'échantillon et les objectifs de l'analyse.

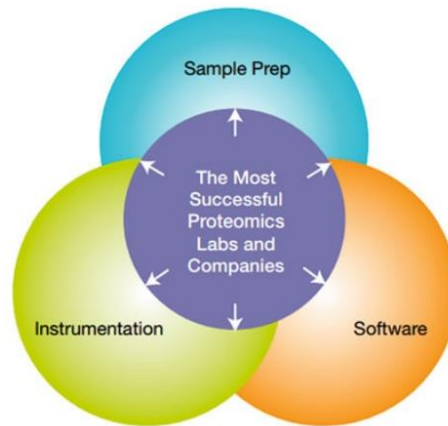


Figure 7 : Les bases du succès en protéomique sont un équilibre entre l'amélioration de l'instrumentation, de la préparation d'échantillons et des logiciels de traitement de données <sup>42</sup>.

### 1.C Les principales étapes en protéomique *bottom-up* classique

Afin d'étudier des échantillons complexes comme le protéome de millions de cellules, le protocole de protéomique classique comporte plusieurs étapes qui sont : une préparation de l'échantillon (lyse, dénaturation, réduction/alkylation et digestion des protéines), une analyse en couplage chromatographie-spectrométrie de masse (séparation, ionisation, analyse) et une acquisition et un traitement des données (Figure 8).

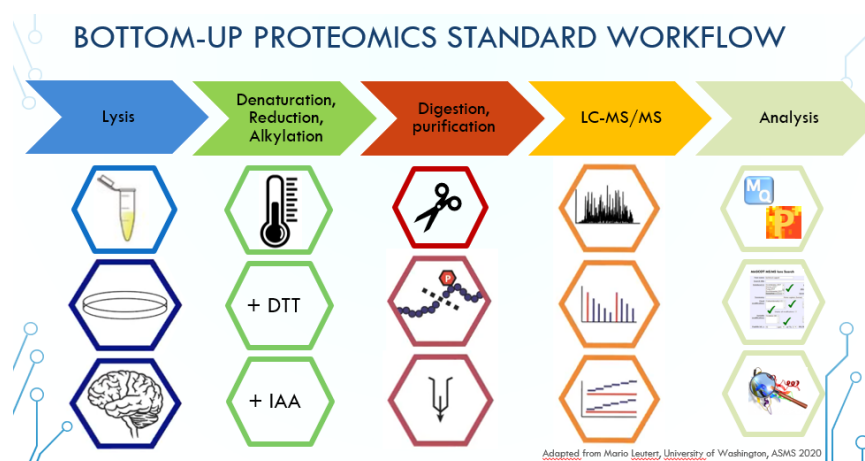


Figure 8 : Exemple des étapes détaillées d'un protocole en protéomique classique

En effet, si quelques groupes utilisent avec succès la spectrométrie de masse couplée à la chromatographie liquide (LC-MS, pour *Liquid Chromatography-Mass Spectrometry*) pour

## INTRODUCTION : L'analyse de cellules uniques

analyser des protéines intactes via la protéomique « *top-down* », la plupart des travaux actuels portent sur la protéomique « *shotgun* » ou « *bottom-up* »<sup>43</sup>. La protéomique *bottom-up* est une approche dans laquelle les protéines sont identifiées non pas par analyse directe, mais par identification des peptides protéolytiques issus de ces protéines<sup>35</sup>. La préparation des échantillons pour les stratégies en *bottom-up* est décrite dans les paragraphes suivants.

### 1.C.1. La lyse cellulaire

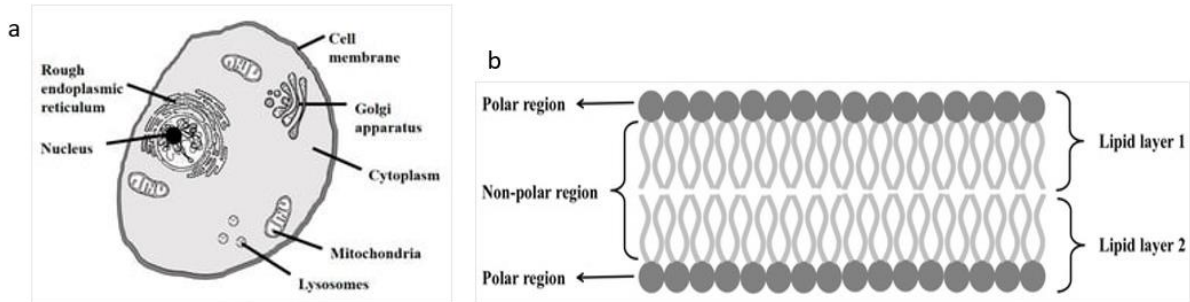


Figure 9 : (a) Cellule animale et ses organites (b) membrane plasmique formée d'une bicouche lipidique.

Le pré-traitement de l'échantillon, c'est-à-dire la lyse cellulaire pour extraire les protéines cellulaires est une étape très importante pour leur analyse. Les cellules de mammifères étudiées dans ce projet sont des cellules eucaryotes animales entourées d'une membrane plasmique composée d'une bicouche phospholipidique (Figure 9). Leur lyse cellulaire peut se faire de façon mécanique ou non mécanique (Figure 10). Les techniques non mécaniques peuvent être physiques, chimiques ou biologiques<sup>44</sup>. Nous aborderons les techniques les plus souvent utilisées en protéomique que sont les techniques thermiques ou chimiques.

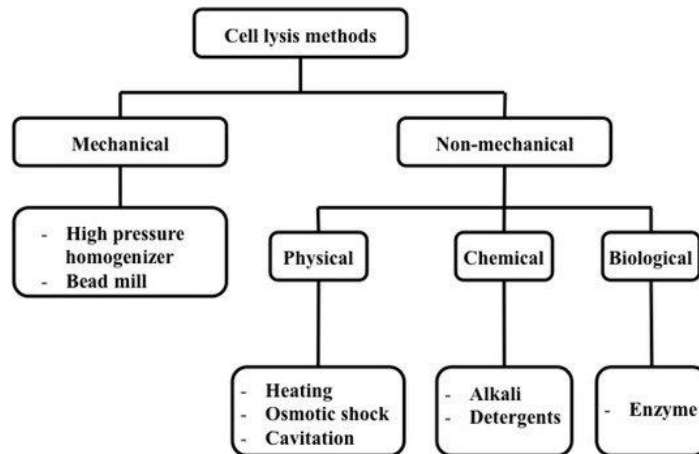


Figure 10 : Classification des différents types de lyses cellulaires.

La lyse physique est une méthode utilisant des sources d'énergie externes afin de rompre la membrane plasmique. Les différentes sources comprennent la chaleur, la pression et l'énergie sonore. Les méthodes associées sont respectivement la lyse thermique, le choc osmotique et la cavitation par ultrasons<sup>44</sup>.

La lyse par choc thermique est effectuée par des cycles de congélation/décongélation. Cela entraîne la formation de glace sur la membrane cellulaire, contribuant à sa destruction. Selon la revue de Shehadul, cette méthode est lente et incompatible avec certains composés

## INTRODUCTION : L'analyse de cellules uniques

cellulaires sensibles à la température <sup>44</sup>. Cependant, Johnson *et al.* ont montré qu'en utilisant les cycles de congélation/décongélation, il est possible de séparer des protéines recombinantes exprimées chez *E. coli* et ont constaté que la méthode de congélation/décongélation était plus efficace que d'autres méthodes comme la sonication ou la lyse enzymatique <sup>45</sup>. La lyse thermique peut aussi être effectuée avec une augmentation de la température seule. Cela déstructure la bicouche lipidique membranaire et dénature les protéines membranaires et entraîne la libération des protéines intracellulaires. Un chauffage à 90°C permet de libérer une grande quantité de protéines d'*E. coli* <sup>46</sup>. D'après Shehadul, le chauffage prolongé peut endommager les protéines <sup>44</sup>.

La sonication par ultrasons génère un phénomène de cavitation qui induit la formation puis la rupture de bulles créées en réduisant la pression locale en augmentant la vitesse et la vibration ultrasonique. La réduction de la pression provoque l'effondrement de cette bulle qui libère une grande énergie mécanique sous la forme d'une onde de choc qui se propage dans le milieu <sup>47</sup>. Cette onde de choc est utilisée pour désintégrer la membrane plasmique des cellules. Il a été démontré qu'une vibration ultrasonique de 15 à 20 kHz pouvait être utilisée pour générer une onde de pression sonore et que cette onde de perturbation était proportionnelle à la puissance absorbée et indépendante de la concentration de la biomasse. Cependant, comme cette technique produit beaucoup de chaleur, elle souffre des mêmes limitations que la lyse thermique et les enzymes extraites par cavitation ultrasonique ont été signalées comme dégradées <sup>48</sup>.

Les méthodes de lyse chimique utilisent des tampons de lyse afin de déstabiliser la membrane cellulaire et de libérer le contenu des cellules de mammifères (Figure 11) <sup>44</sup>.

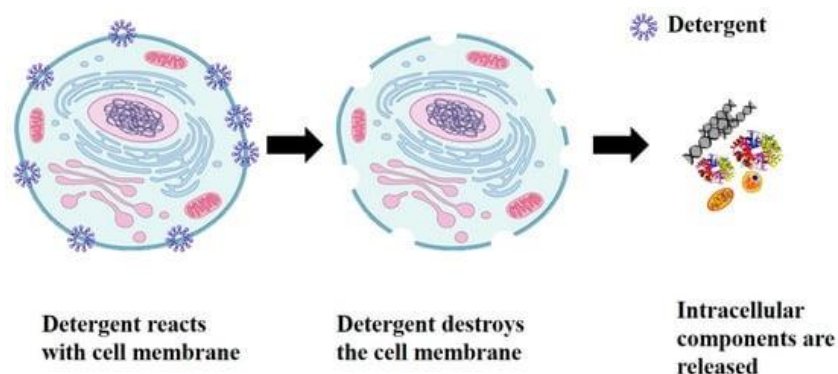


Figure 11 : Lyse cellulaire avec utilisation de détergent pour détruire la membrane plasmique et libérer les composants intracellulaires.

Les détergents, également appelés tensio-actifs, ont la capacité de perturber les interactions hydrophobes-hydrophiles de la membrane cellulaire. Ils sont capables de perturber les interactions lipide-lipide, lipide-protéine et protéine-protéine. Ils peuvent être divisés en détergents cationiques, anioniques et non ioniques. Les détergents non ioniques (CHAPS, le Triton-X et Tween) sont préférés car ils sont les moins délétères pour les protéines. Le sodium dodécylsulfate (SDS) est un détergent ionique largement utilisé pour lyser les cellules en raison de sa grande affinité avec les protéines en les dénaturant rapidement. Il est utilisé dans les techniques d'électrophorèse sur gel et de *western blotting*. Outre les détergents ioniques et non ioniques, des agents chaotropiques peuvent également être utilisés pour la lyse cellulaire, en modifiant la structure tridimensionnelle des macro-biomolécules. Il s'agit notamment de l'urée, de la guanidine et de l'acide éthylène diamine

tétra acétique (EDTA), qui peuvent détruire le réseau de liaisons hydrogène entre les molécules d'eau et rendre le solvant moins hydrophile, ce qui affaiblit les interactions hydrophobes. Ils évitent également l'agrégation des protéines et permettent une meilleure solubilisation <sup>49</sup>. Cependant, ces détergents sont aussi des contaminants délétères pour les analyses protéomiques car ils perturbent l'analyse en spectrométrie de masse. Une étape supplémentaire de purification doit donc être intégrée au protocole de lyse cellulaire lors de leur utilisation <sup>50</sup>. La purification de ces petites molécules peut se faire par plusieurs méthodes comme la dialyse, la chromatographie, l'électrophorèse sur gel, la précipitation de protéines, le FASP (pour « *filter assisted sample preparation* ») ou encore la méthode SP3. Cette étape supplémentaire va induire un coût en rendement global.

Il existe aussi des solutions commerciales de détergents MS-compatibles comme l'Invitrosol (Thermo Fisher Scientific) qui contient plusieurs détergents mono-dispersés, élués en chromatographie à polarité de phase inversée (RPLC) dans des régions du gradient de phase mobile qui n'interfèrent pas avec les peptides. D'autres exemples de détergents MS-compatibles sont les détergents clivables comme le Rapigest (Waters), le Protease Max (Promega), le PPS Silent surfactant (Expedeon) ou le Progenta (Protea) qui, après dégradation à haute température ou à faible pH, n'interfèrent plus avec la LC-MS.

Lors de l'extraction des protéines est ajouté aussi un cocktail d'inhibiteurs de protéases et de phosphatases afin d'éviter les protéolyses non spécifiques ou la perte de modifications post traductionnelles sur les protéines comme la phosphorylation.

### 1.C.2. La préparation des protéines

Les protéines une fois extraites doivent être dénaturées en rompant leur configuration 3D (par exemple par la chaleur) sauf en cas de lyse thermique. Préalablement à la digestion enzymatique et afin de faciliter la dénaturation des protéines, les ponts disulfures qui se créent entre les résidus des cystéines (acide aminé constituant des protéines) sont réduits par ajout d'agents réducteurs comme le dithiothréitol (DTT) ou le tri-carboxy-éthyl-phosphine TCEP (Figure 12). Afin d'empêcher la reformation de ces ponts, les fonctions thiol des cystéines désormais libres sont alkylées par l'ajout d'iodoacétamide (IAA), par exemple (Figure 12).

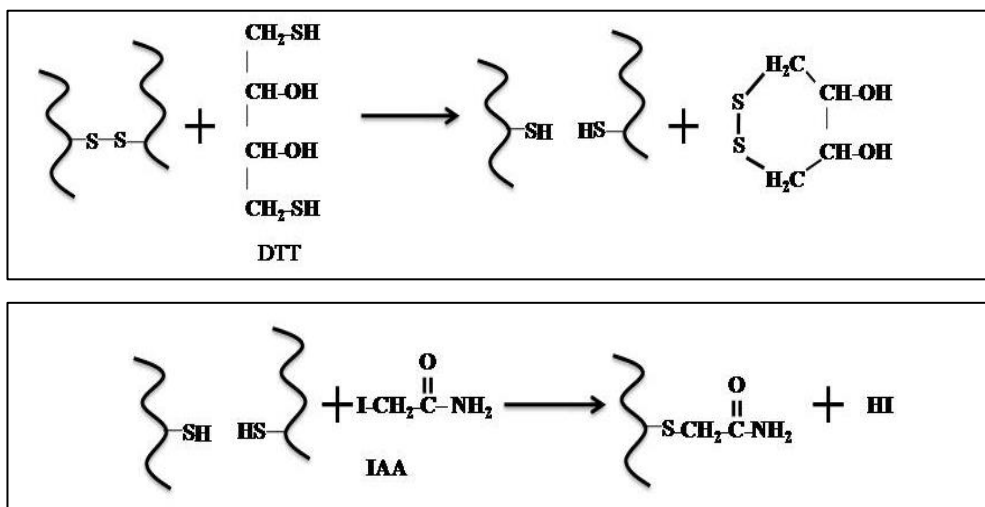


Figure 12 : Schéma des réactions de réduction (en haut) et d'alkylation (en bas) des protéines.



### 1.C.3. La digestion enzymatique

Enfin, les protéines sont digérées par des endoprotéases qui coupent la chaîne peptidique générant des fragments de longueur variable selon les sites de coupure préférentiels (Figure 13). Cette étape de digestion peut être faite de plusieurs façons, sur gel, sur colonne ou en solution <sup>49</sup>. En protéomique, l'utilisation de la trypsine est privilégiée car elle a la spécificité de couper en C-terminal des résidus basiques, la lysine (K) et l'arginine (R). Ces acides aminés basiques chargés favorisent une fragmentation protéolytique car ils sont bien répartis dans les protéines donnant des peptides de 7 à 15 résidus d'acides aminés <sup>51</sup>. De temps à autre, certains sites potentiels restent intacts (clivages manqués ou *miscleavages*) notamment si les résidus cibles sont liés à une Proline (P) par exemple. Afin d'augmenter le taux de couverture protéique des analyses, d'autres endoprotéases comme la chymotrypsine ou l'endoprotéase Asp-N ou Glu-C peuvent être utilisées.

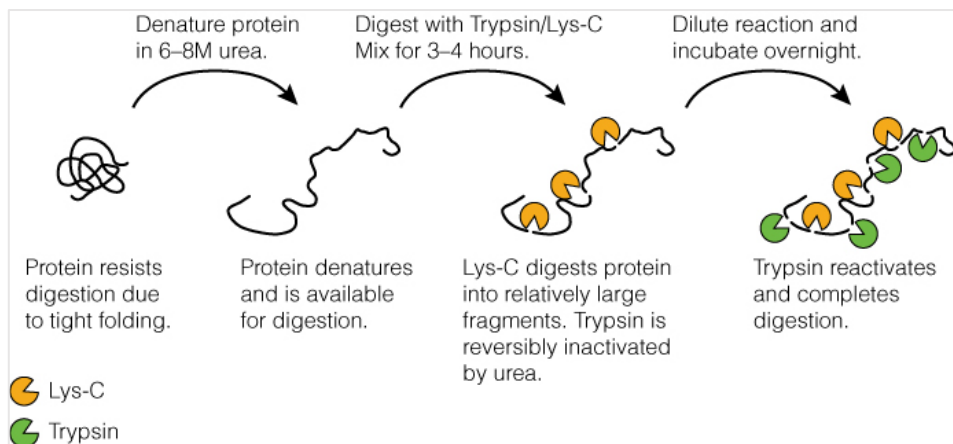


Figure 13 : Combinaison des enzymes trypsine et Lys-C pour la digestion enzymatique d'une protéine.

### 1.C.4. Enrichissement des protéines

Comme indiqué plus haut dans l'introduction, les abondances des protéines varient avec un ordre de grandeur allant de 6 à 12 dans un tissu avec certaines protéines majoritaires. Afin d'enrichir les protéines de faible abondance, il est possible de faire une déplétion des protéines majoritaires. Une chromatographie d'affinité peut être utilisée afin de retirer l'albumine ou d'autres protéines abondantes comme les immunoglobulines ou les lipoprotéines par exemple. Sinon, un processus de préfractionnement de l'échantillon en chromatographie peut permettre d'ajouter une dimension de séparation et de répartir les protéines sur plusieurs fractions ce qui permet une meilleure détection des protéines minoritaires.

### 1.C.5. Les techniques séparatives miniaturisées

Afin d'assurer une bonne couverture de l'échantillon lors de l'acquisition par spectrométrie de masse, il faut une bonne séparation des peptides en amont car les mélanges complexes de peptides souffrent de plusieurs effets limitants pour leur détection : effet de suppression du signal des espèces entre elles, rapidité de séquençage limitée, dilution et chute du rapport signal/bruit <sup>52</sup>. Plusieurs types de techniques chromatographiques comme l'échange d'ions, la phase inverse, l'interaction hydrophobe, l'affinité et l'exclusion de taille, ont été utilisés pour le fractionnement des protéines <sup>49</sup>. Lorsqu'une protéine particulière doit être obtenue avec un haut degré de pureté, une combinaison de plusieurs étapes

## INTRODUCTION : L'analyse de cellules uniques

chromatographiques suivie d'un SDS-PAGE est fréquemment réalisée. L'HPLC (chromatographie liquide à haute performance) est souvent privilégiée pour l'analyse de mélanges peptidiques en séparations mono ou multidimensionnelles car elle peut être directement couplée à un analyseur UV ou MS.

La chromatographie liquide est une technique de séparation des analytes entre une phase liquide qui transporte les analytes à travers la colonne et une phase stationnaire contenue dans la colonne ou sur les billes de résine dans la colonne. La phase stationnaire interagit de manière non covalente avec chaque analyte en fonction de sa structure chimique, ce qui entraîne la séparation des analytes en fonction du temps passé dans la colonne de séparation. Moins les analytes interagissent avec la phase stationnaire, plus ils sont transportés rapidement dans le système. Le choix de la phase stationnaire est important car une phase ayant des forces de liaisons intermoléculaires similaires à celles qui caractérisent l'analyte permettra une plus grande interaction entre la colonne et les analytes et augmentera leur temps de rétention dans la colonne. Il en résultera une meilleure séparation des analytes structurellement similaires. Ainsi, les nombreux analytes d'un échantillon sont identifiés par le temps de rétention dans le système pour un ensemble donné de conditions. Le temps mort pendant lequel une espèce chimique non retenue restera dans le système est  $t_m$  (c'est la ligne de solvant). Tous les composés seront élués en ce temps minimum. Cela correspond au volume total de l'injecteur et de la colonne. Toute affinité pour la phase stationnaire entraîne la rétention du composé dans la colonne et son élution de la colonne à un temps supérieur à  $t_m$ . La résolution chromatographique est une mesure de la séparation de deux pics de temps de rétention  $t$  différents dans un chromatogramme. La forme du pic chromatographique a également une importance. Les pics sont comparés en fonction de leur largeur à mi-hauteur. Un élargissement des pics est souvent dû à la diffusion de l'analyte dans la colonne et peut entraîner une mauvaise séparation ou cacher un analyte qui aurait un temps de rétention proche.

Plusieurs paramètres entrent en jeu pour la résolution chromatographique. En HPLC, la pression de la phase mobile, le débit, la vitesse linéaire et la polarité de la phase mobile affectent le temps de rétention d'un composé.

En séparation par chromatographie liquide, un débit plus petit a été lié directement à une meilleure séparation des peptides jusqu'à un certain débit. En effet, la réduction du débit de la séparation à un faible débit de quelques nanolitres par minute peut augmenter considérablement l'efficacité de l'ionisation et réduire davantage les problèmes de suppression de signal pendant l'ionisation<sup>53</sup>. La formule du débit est décrite ci-dessous (Équation 1).

$$F = \frac{(\pi * dc^2)}{4} * \mu * \varepsilon \text{ Avec } \mu = L / t_m$$

Équation 1 : Équation du débit en fonction du diamètre interne de la colonne avec  $f$  = débit,  $dc$  = diamètre interne de la colonne,  $\mu$  = vitesse linéaire de la phase mobile ( $L$  = longueur de la colonne et  $t_m$  = temps d'élution d'un analyte non retenu dans la colonne) et  $\varepsilon$  = porosité de la colonne.

Le premier paramètre est le débit ( $F$ ). Selon cette équation, il dépend directement du diamètre interne de la colonne ( $dc$ ), de la vitesse de la phase mobile dans la colonne ( $\mu$ ) et de la porosité de la colonne ( $E$ ). La vitesse de la phase mobile  $\mu$  est décrite comme étant égale à la longueur de la colonne ( $L$ ) divisée par le temps d'élution d'un analyte non retenu ( $t_m$ ). La porosité des colonnes capillaires comme celles utilisées en nano LC est considérée égale à 1.

## INTRODUCTION : L'analyse de cellules uniques

Le diamètre interne (id.) de la colonne analytique est un facteur crucial dans l'analyse des échantillons surtout des échantillons de trace. En réduisant le diamètre interne de la colonne, la concentration des composés élués est plus élevée et la sensibilité de l'analyse est accrue<sup>54</sup>. Les colonnes à id. réduits comprennent la variante à micro-échelle (environ 1 mm d.i.), la variante LC capillaire (environ 0.2-0.5 mm d.i.), et la variante nano LC, qui sont en général des colonnes capillaires avec un diamètre interne inférieur à 100  $\mu\text{m}$ <sup>55</sup>. La longueur de la colonne joue également un rôle dans la séparation. Plus la colonne est longue, meilleure sera la séparation.

La taille des particules qui tapissent la colonne est également un facteur important. Jusqu'au début des années 2000, les particules entre 3 et 5  $\mu\text{m}$  et une porosité de 300 Å étaient les plus utilisées. Plus récemment, des particules de l'ordre du micron sont utilisées avec une porosité de 100 Å. La réduction de la taille des particules augmente la surface de contact entre la phase stationnaire et les analytes et cela permet d'avoir des colonnes avec une meilleure résolution<sup>56</sup>. Dans le commerce, des colonnes capillaires avec un diamètre interne de 50 ou 75  $\mu\text{m}$  sont trouvées auprès de plusieurs fournisseurs (Agilent, Thermo Fisher, Waters, Sigma Aldrich, Dr. Maisch). Ces colonnes sont remplies de particules de 1.7 à 3.5  $\mu\text{m}$  et sont disponibles en différentes longueurs. Les colonnes à diamètre interne très étroit (10-30  $\mu\text{m}$ ), destinées aux analyses ultra-sensibles, sont disponibles dans le commerce auprès d'au moins un fournisseur (CoAnn Technologies) et sont tapissées de particules de 1.7  $\mu\text{m}$ . Les colonnes de 75  $\mu\text{m}$  d.i. sont devenues un standard dans les analyses protéomiques<sup>55</sup>.

Comme le débit dépend des caractéristiques de chaque colonne analytique : le diamètre interne de la colonne et sa longueur, il doit être adapté à chaque colonne selon les pertes de charges tolérables pour le système, en fonction des dimensions de la colonne et de la granulométrie de la phase stationnaire, pour éviter les surpressions et la destruction de celle-ci. De plus, en LC, la colonne doit être maintenue à une température stable (en général 50°C) afin d'éviter que la reproductibilité ne soit affectée par les températures changeantes des laboratoires et éviter les problèmes de surpression. En effet, dans les colonnes capillaires, la pression peut monter jusqu'à 600-800 bar sur une configuration standard avec un débit standard. L'augmentation de la température permet une réduction de la pression dans la colonne à 300-400 bar.

En protéomique, pour une colonne standard de 75  $\mu\text{m}$  de diamètre interne et 25cm de longueur, le débit suggéré est de 300nl/min. Ce débit a été adapté sur les colonnes Aurora (Ion Opticks) de 75  $\mu\text{m}$  de diamètre interne et 25cm de longueur à 220nl/min. Ces colonnes sont utilisées en connexion avec la source nano ESI du Tims TOF Pro (Bruker). Avec des colonnes commerciales de 50  $\mu\text{m}$  de diamètre interne et 50cm de longueur, le débit utilisé pour la séparation est de 115 nl/min. En protéomique de la cellule unique, le débit a pu être divisé à 20nl/min par l'équipe de Ryan Kelly travaillant sur des colonnes à 20  $\mu\text{m}$  de diamètre interne avec un diviseur de débit<sup>53</sup>.

La chromatographie à polarité de phases inversée (RPLC pour « *Reversed Phase Liquid Chromatography* ») est la méthode la plus répandue pour la séparation de peptides. Les molécules sont séparées selon un gradient de phase mobile en fonction de leur hydrophobicité et leur taille, par interaction avec la phase stationnaire apolaire hydrophobe<sup>57</sup>. La phase stationnaire apolaire est souvent constituée de particules de silice apolaire greffée de deux à dix-huit carbones. En protéomique, en général, on utilise une phase en C18.

## INTRODUCTION : L'analyse de cellules uniques

La séparation des peptides est généralement effectuée grâce à un gradient d'élution constitué d'un mélange eau/acétonitrile/acide.

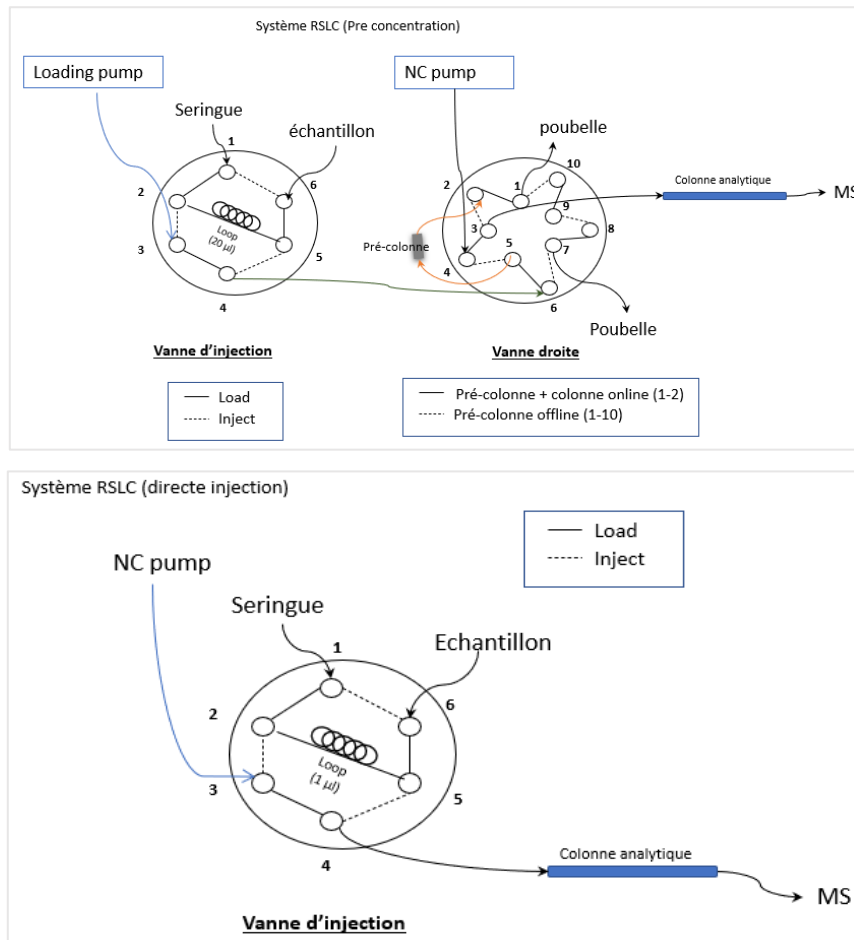


Figure 14 : Schéma d'un exemple de montage en HPLC U3000 (Thermo Fisher) (en haut) avec une préconcentration, (en bas) en injection directe.

Les principaux composants d'un système HPLC sont : la pompe, l'injecteur, la colonne, le détecteur et le système de données. Les réservoirs d'éluant (phase mobile) sont gérés par la pompe afin de distribuer le gradient de solvant d'élution. La connectique et les matériaux des tubes sont très importants et influent sur les performances et la durée de vie du système<sup>58</sup>. Du fait de la faible abondance des échantillons en protéomique, l'ensemble du système est miniaturisé. On parle de nano LC.

La colonne analytique peut être précédée d'une colonne de piégeage qui permet d'injecter un volume d'échantillon équivalent à plusieurs volumes de colonne, de préconcentrer et dessaler les échantillons avant leur arrivée sur la colonne analytique (Figure 14). C'est une configuration en préconcentration (PC). Il est possible d'éviter cette étape de dessalage et de connecter la colonne analytique directement à l'injecteur dans un montage appelé « en injection directe » (DI) (Figure 14). Il est alors essentiel que l'ensemble des éléments respectent cette échelle pour minimiser les volumes morts qui feraient immédiatement chuter la capacité de pic de la séparation. Les capillaires de connexion ont donc des diamètres internes égaux (avant colonne) ou inférieurs (après colonne) au diamètre de la colonne utilisée. L'injecteur est composé d'une boucle de 1, 5, 10 ou 25 µl, ce qui provoque une injection non ponctuelle qui est corrigée par l'utilisation d'un gradient de la phase mobile.

## INTRODUCTION : L'analyse de cellules uniques

Ce système LC est ensuite couplé à un spectromètre de masse pour l'analyse des peptides d'intérêt.

### 1. DL'analyse des peptides par spectrométrie de masse

Le spectromètre de masse est composé de 3 principaux éléments : la source ionique et les optiques, l'analyseur de masse et le détecteur (Figure 15) <sup>35</sup>.

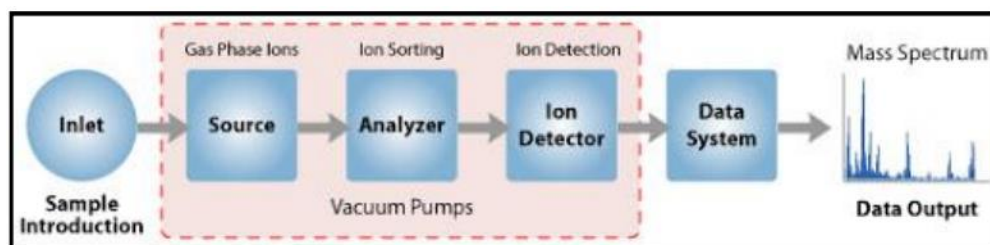


Figure 15 : Schéma des composants d'un spectromètre de masse.

La source d'ion sert à convertir les molécules d'intérêt en solution sortant de la LC en ions en phase gazeuse avant leur arrivée devant le détecteur. Les ions sont d'abord volatilisés puis ionisés. Les techniques d'ionisation ont beaucoup évolué et presque toutes les macromolécules non volatiles et thermiquement labiles peuvent être ionisées en phase gazeuse. Deux de ces techniques se sont imposées en protéomique, il s'agit de la désorption laser assistée par matrice (MALDI pour « *matrix-assisted laser desorption ionization* ») et l'ionisation par électro-nébulisation (ESI pour « *electrospray ionization* »). Les développements de ces techniques ont valu le prix Nobel de chimie à John B. Fenn et Koichi Tanaka en 2002 <sup>59</sup>. Le développement de la micro ESI et de la nano ESI ont permis de réduire le débit à quelques nanolitres par minute et d'améliorer la sensibilité de la méthode <sup>32</sup>. Différentes configurations de sources nano ESI ont été utilisées pendant ce projet. Les analytes sont présentés à l'extrémité d'un capillaire de quelques  $\mu\text{m}$  de diamètre interne et connecté électriquement pour qu'une différence de potentiel d'environ 2 kV soit appliquée entre cet émetteur et l'entrée du spectromètre de masse <sup>35</sup>. La tension nécessaire à l'établissement de l'électro-spray dépend du diamètre interne de l'émetteur du capillaire ainsi que des solvants utilisés. Le principe de l'ESI ne sera pas détaillé ici car il n'a pas été l'objet de l'étude et est le sujet de nombreux travaux depuis sa première utilisation par Yamashita, Fenn et indépendamment Gall en 1984 <sup>60,61,62,63,64</sup>. La caractéristique principale de l'ESI par rapport aux autres sources d'ionisation est la création d'espèces ioniques multichargées, la sensibilité à la concentration des analytes et leur débit d'arrivée. Une des limitations de l'ESI à ne pas négliger est sa susceptibilité aux effets de suppression des ions. Si la solution contient une concentration élevée de sels ( $> \sim 1\text{mM}$ ), la formation des ions est entravée. De plus, lors de l'analyse d'un mélange complexe, la présence d'analytes à forte concentration peut supprimer la formation d'ions à plus faible concentration <sup>51</sup>.

Par la suite, les ions formés sont transférés dans l'analyseur. Les analyseurs de masse vont conférer leur spécificité à chaque instrument. Les analyseurs de masse à piège Orbitrap et à résonance cyclotron (ICR pour *ion cyclotron resonance*) séparent les ions en fonction de leur fréquence de résonance dans un champ électrique ou magnétique, les pièges à ions (IT pour *ion trap*), les quadripôles (Q) utilisent leur stabilité dans un champ électrique et les analyseurs à temps de vol (TOF) utilisent le temps de vol des ions dans une zone libre de champ <sup>35</sup>.

## INTRODUCTION : L'analyse de cellules uniques

Chaque analyseur de masse possède des propriétés uniques, telles que la gamme de masse, la vitesse d'analyse, la résolution, la sensibilité, la gamme dynamique. Les nouveaux spectromètres de masse sont hybrides et combinent plusieurs analyseurs pour répondre à des besoins plus spécifiques. La présentation et le détail des analyseurs utilisés durant ce projet se fera dans la partie « instruments utilisés ».

La spectrométrie de masse (MS) utilise la donnée du rapport masse/charge ( $m/z$ ) des ions pour la caractérisation des analytes. C'est l'outil le plus complet et le plus polyvalent de la protéomique à grande échelle puisqu'elle allie sensibilité et précision afin d'identifier et quantifier le maximum de protéines<sup>35</sup>. C'est aussi la méthode qui permet de quantifier le plus de protéines par cellule avec une spécificité de détection très élevée (Figure 16). Ici, trois paramètres sont importants : la résolution, la justesse et la précision. Ces deux derniers paramètres sont indépendants (une mesure exacte ne nécessite pas d'être une mesure précise et vice versa). En MS, la justesse est liée à la résolution de l'appareil, c'est-à-dire sa capacité à résoudre deux pics adjacents. La résolution peut être calculée grâce à la largeur à mi-hauteur des pics (FMWH pour « *full width at half maximum* ») et est définie comme  $m/\Delta m$ , où  $m$  est la masse entière des pics à résoudre et  $\Delta m$  est la différence de masse entre les deux pics<sup>59</sup>. La précision est liée à la robustesse de l'analyse.

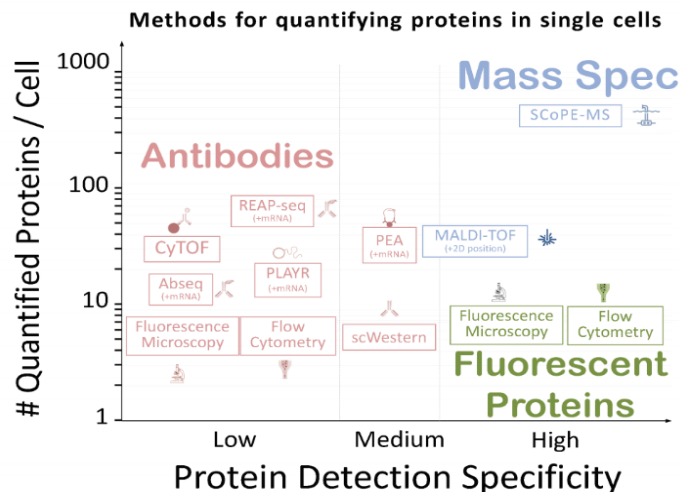


Figure 16 : Comparaison des méthodes de quantification des protéines dans les cellules uniques<sup>65</sup>.

### 1.D.1. La fragmentation des peptides

Comme indiqué précédemment, la protéomique fait appel à un couplage de type nano LC-MS pour limiter les effets de suppression. Ainsi, à chaque temps  $t$  donné, une fraction du mélange de peptides est éluée de la chromatographie liquide dans des conditions compatibles pour être ionisés et analysés en MS (Figure 17). Leur  $m/z$  est déterminé par un premier analyseur de masse en mode MS1 (ou MS). Les peptides d'intérêt sont isolés des autres ions et fragmentés dans une trappe à ions ou une cellule de collision. Les ions fragmentés résultants sont analysés dans un second analyseur en mode MS en tandem (MS/MS). Les spectres correspondant à ces ions fragmentés permettent d'obtenir l'information de la masse des peptides analysés<sup>35</sup>.

Pour obtenir une fragmentation, l'énergie interne des ions est augmentée, ce qui entraîne leur dissociation. Les différents modes de fragmentation utilisés seront présentés dans la partie « instruments utilisés ». La fragmentation peut se faire selon différents modes comme

## INTRODUCTION : L'analyse de cellules uniques

le CID pour « *collision induced dissociation* », IRMPD pour « *infrared multi-photon photodissociation* », SID pour « *surface induced photodissociation* » ou HCD pour « *higher energy collisional dissociation* »<sup>59</sup>. En dissociation induite par collision (CID), également appelée dissociation activée par collision (CAD), les ions sont accélérés dans l'analyseur par l'application d'un potentiel électrique puis ils entrent en collision avec des molécules neutres (souvent de l'hélium, de l'azote ou de l'argon)<sup>66</sup>. Lors de la collision, une partie de l'énergie cinétique est convertie en énergie interne, ce qui entraîne la fragmentation de l'ion moléculaire en fragments plus petits. Ces ions fragmentés peuvent ensuite être analysés par spectrométrie de masse en tandem. La dissociation par collision à haute énergie (HCD) se fait dans une cellule de collision. La fragmentation HCD utilise une plus haute énergie que la fragmentation CID. De plus, elle permet de détecter des fragments de plus faibles rapports masses/charges par rapport à la fragmentation CID en trappe à ions<sup>67</sup>. Enfin, dans les appareils à mobilité ionique, une autre technique d'analyse appelée PASEF® pour « *Parallel Accumulation–Serial Fragmentation* » utilise la séparation additionnelle des ions dans le TIMS (décrite plus tard dans la partie « instruments utilisés) pour fragmenter ces ions en série.

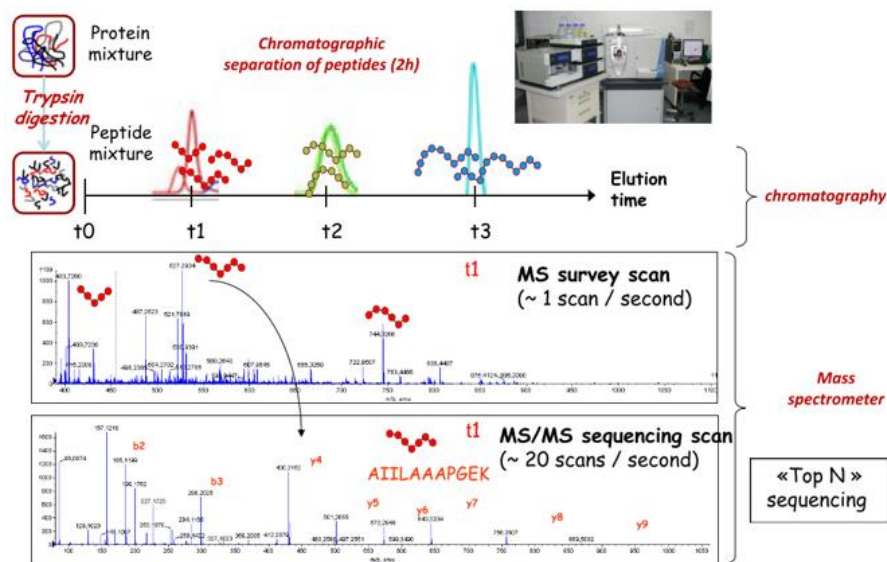


Figure 17 : Schéma explicatif de la séparation des peptides par chromatographie liquide à leur analyse en spectrométrie de masse en tandem (d'après le cours du Pr Marie-Pierre Bousquet).

Les ions fragments caractéristiques sont nommés grâce à la nomenclature de Roepstorff et Biemann<sup>68,69</sup>. Les différents fragments résultant de ruptures le long de la chaîne peptidique sont présentés dans la Figure 18. Les fragmentations classiques (CID, HCD) sur les peptides tryptiques donnent naissance à des ions fragments majoritairement de type y et b. Il s'agit de fragments générés après rupture des liaisons amides du peptide et qui contiennent respectivement l'extrémité C- ou N-terminale du précurseur. S'il y a rupture avant la liaison amide, des fragments a et x seront générés. S'il y a rupture après la liaison amide, des fragments c et z sont générés. Ces fragments surviennent lors de combinaisons d'acides aminés particuliers ou à l'utilisation de modes de fragmentation particuliers comme l'ETD (pour « *electron transfer dissociation* »).

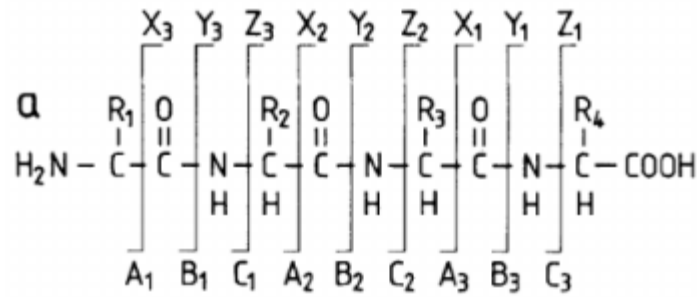


Figure 18 : Schéma de fragmentation original de Roepstorff et Fohlmann de 1984 <sup>68</sup>.

### 1.D.2. Les modes d'acquisition en spectrométrie de masse

Dans le cas de l'analyse de mélanges complexes, il est indispensable d'automatiser ces acquisitions. En effet un protéome complet comptant quelques milliers de protéines, le digestat contient plusieurs centaines de milliers de peptides protéolytiques. À chaque instant des dizaines de peptides sont donc élués de la colonne analytique. L'obtention des spectres de masse MS/MS permettant le séquençage des peptides peut être faite de manière ciblée en PRM (pour « *parallel reaction monitoring* ») ou MRM (pour « *multiple reaction monitoring* ») ou bien de manière non ciblée. Les acquisitions non ciblées peuvent être dépendantes (DDA pour « *Data Dependent Acquisition* ») ou indépendantes (DIA pour « *Data Independent Acquisition* ») des données, signifiant que les informations provenant d'un scan MS en cours déterminent ou non les paramètres des scans MS/MS suivants. La différence entre les acquisitions en DDA et en DIA survient à l'étape de sélection des ions précurseurs avant leur fragmentation dans la cellule CID ou HCD (Figure 19).

En DDA, après la première acquisition MS, une fenêtre autour de chaque masse d'intérêt est sélectionnée (modes Top N ou Top Speed). Ainsi les 20 précurseurs les plus intenses en mode Top 20 par exemple sont sélectionnés pour être fragmentés en MS/MS. Ces masses sont ensuite exclues de la sélection pendant un certain temps. Cependant, en raison de son mode stochastique de sélection des peptides, la DDA manque de reproductibilité et ne parvient pas à identifier toutes les protéines peu abondantes <sup>70</sup>. Sur un appareil Q Exactive par exemple, tous les peptides (environ 30%) ne sont pas identifiables dépendant de la gamme dynamique de l'échantillon, de la quantité de départ et les peptides identifiés peuvent varier entre les répétitions <sup>71</sup>. Par conséquent, de meilleures reproductibilité et sensibilité sont nécessaires pour identifier un nombre plus important de peptides.

Les progrès techniques de la MS avec une sensibilité, une résolution et une précision de masse accrues ont rendu possible le développement de méthodes d'acquisition indépendantes des données (DIA) comme le SWATH (*Sequential Acquisition of All Theoretical fragment-ion spectra*) <sup>72</sup>.

Contrairement aux mesures DDA, l'acquisition en DIA se caractérise par une sélection sans a priori de fenêtres successives plus larges d'ions précurseurs pour fragmentation en MS/MS (Figure 19) <sup>73</sup>. De ce fait, tous les peptides générés lors du premier cycle MS seront fragmentés et analysés lors du deuxième cycle MS, ce qui conduit théoriquement à l'identification d'un plus grand nombre de peptides qu'en DDA et à des niveaux de multiplexage plus élevés qu'en protéomique ciblée <sup>73</sup>. L'acquisition DIA permet une identification et une quantification précises, ainsi que des mesures reproductibles pour un grand nombre d'échantillons <sup>74</sup>. Cependant, les spectres de fragmentation obtenus en DIA sont multiplexés et très difficiles à



## INTRODUCTION : L'analyse de cellules uniques

déconvoluer car ils résultent de la fragmentation simultanée de nombreuses espèces ioniques. Le développement de logiciels basés sur des bibliothèques spectrales pour l'extraction des données ou sur des algorithmes de réseaux neuronaux comme Spectronaut (Biognosys) ou DIA-NN récemment développés pour les analyses DIA est très important pour faciliter la compréhension de ces données <sup>75,76</sup>.

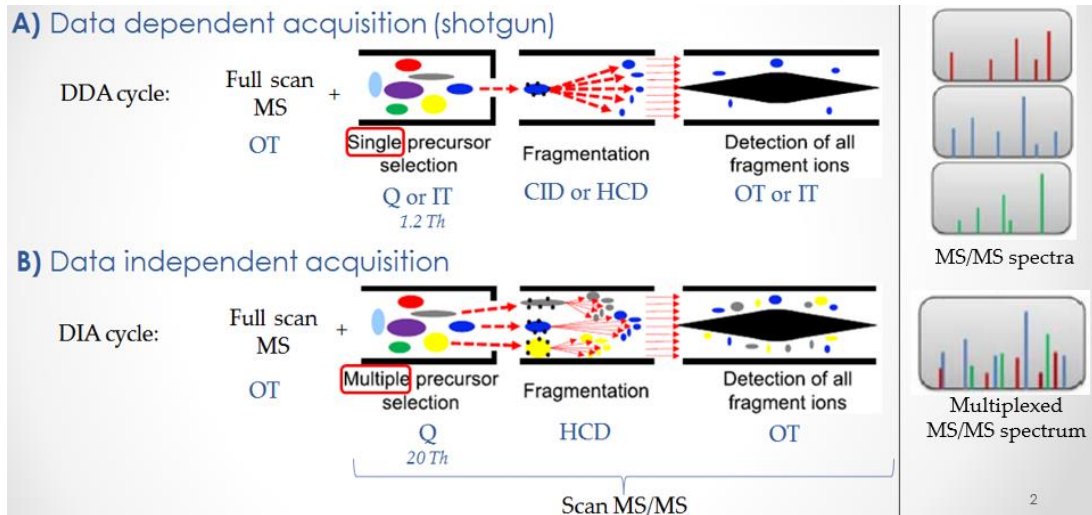
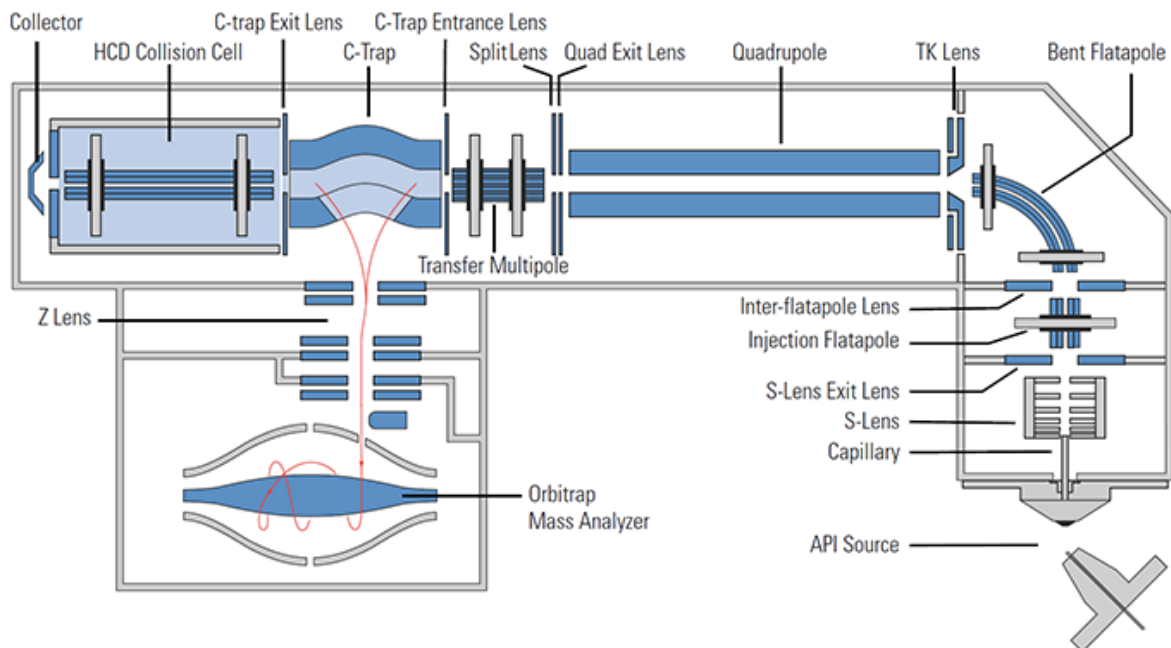


Figure 19 : Exemple de DDA versus DIA sur un instrument Orbitrap.

### 1.D.3. Les instruments utilisés dans ce projet

Il existe de nombreux appareils composés de différentes combinaisons de sources, analyseurs et détecteurs, mais seuls seront décrits les spectromètres de masse utilisés durant cette étude : les Orbitrap Q Exactive HF et Tribrid Eclipse (ThermoFisher Scientific), et le temps de vol TIMS TOF Pro (Bruker).

#### (a) Les Orbitrap, ThermoFisher Scientific



## INTRODUCTION : L'analyse de cellules uniques

Figure 20 : Schéma d'un Orbitrap HF (ThermoFisher Scientific).

Le Q Exactive HF<sup>TM</sup> du constructeur ThermoFisher Scientific est un spectromètre de masse hybride alliant quadripôle et Orbitrap (Figure 20). À l'entrée du spectromètre de masse après la source nano ESI, les ions traversent une série de lentilles de transfert *S-lens* et un flatapôle qui est un octopôle courbé conçu pour accélérer les ions et éliminer un maximum d'espèces neutres qui auraient pu entrer dans le spectromètre de masse avant l'entrée dans le quadripôle. Le quadripôle est un analyseur à faisceau ionique qui différencie les ions en fonction de leur masse/charge selon un potentiel de radio fréquences (RF). Il peut également accélérer les ions, ou les filtrer selon leur  $m/z$ . En effet, la résolution du quadripôle étant plus basse que l'Orbitrap, mais sa rapidité de balayage étant très élevée, il reste un analyseur de choix pour sélectionner les ions. En mode MS/MS, les ions précurseurs sont sélectionnés dans le quadripôle puis envoyés dans la cellule HCD pour être fragmentés avant d'être envoyés dans la C-trap (trappe à ions) afin de les accumuler avant leur analyse dans l'Orbitrap pour l'analyse. L'Orbitrap est un analyseur par transformée de Fourier, fonctionnant avec un champ électrostatique, présenté pour la première fois à la conférence annuelle de l'ASMS (*American Society for Mass Spectrometry*) en 1999. Il se base sur un piégeage orbital, le piège de Kingdon sous vide poussé ( $10^{-10}$  torr)<sup>77</sup>. Il est composé de deux électrodes, une électrode centrale en forme de fuseau et une électrode axiale qui possède une forme de tonneau coupé en deux parties égales (Figure 21). Les ions entrent dans l'Orbitrap par l'interstice de séparation entre les deux moitiés de l'électrode axiale. Les électrodes sont connectées de façon indépendante, ce qui permet de faire varier leur potentiel électrique pour créer un champ électrique. Ce champ électrique est maintenu stable car nécessaire pour mettre en orbite les ions entrants. Les ions oscillent en orbite autour de l'électrode centrale dans les directions axiales et radiales sous l'action concomitante du champ électrique. L'oscillation axiale est une harmonique indépendante de l'instrument et des conditions initiale et dont la fréquence est inversement proportionnelle à la racine carrée de la valeur  $m/z$  des ions. Le courant induit par le passage des ions selon leur oscillation axiale est mesuré et une transformée de Fourier est appliquée afin d'obtenir les fréquences d'oscillation et par suite le  $m/z$  des ions présents<sup>78</sup>.

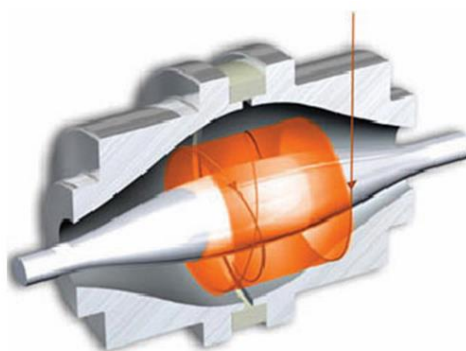


Figure 21 : L'analyseur Orbitrap composé d'une électrode centrale et d'une électrode axiale.

L'Eclipse Tribrid<sup>TM</sup> ThermoFisher Scientific fait partie des dernières générations d'instruments hybrides (contenant 3 types d'analyseur) présenté en 2019 (Figure 22). L'optique ionique a été améliorée afin d'augmenter la sensibilité. En plus d'un quadripôle QR5, qui est segmenté par rapport aux anciens quadripôles, augmentant sa sensibilité, cet appareil est composé d'un Orbitrap et d'un piège à ion (IT pour « *Ion Trap* »). Il possède

## INTRODUCTION : L'analyse de cellules uniques

également plusieurs modes de fragmentation dont l'ETD. Le piège à ions composé de deux cellules, une première cellule sous haute pression et une seconde à basse pression. Elles permettent une fragmentation CID puis une stabilisation pour analyse directe des ions fragments à très haute fréquence, simultanément à l'accumulation du signal en MS à très haute résolution dans l'orbitrap. En protéomique de la cellule unique, l'Eclipse a été comparé à l'ancienne génération des Tribrid, le Fusion Lumos, et 20% de groupes de protéines ont été identifiés en plus avec l'Eclipse.

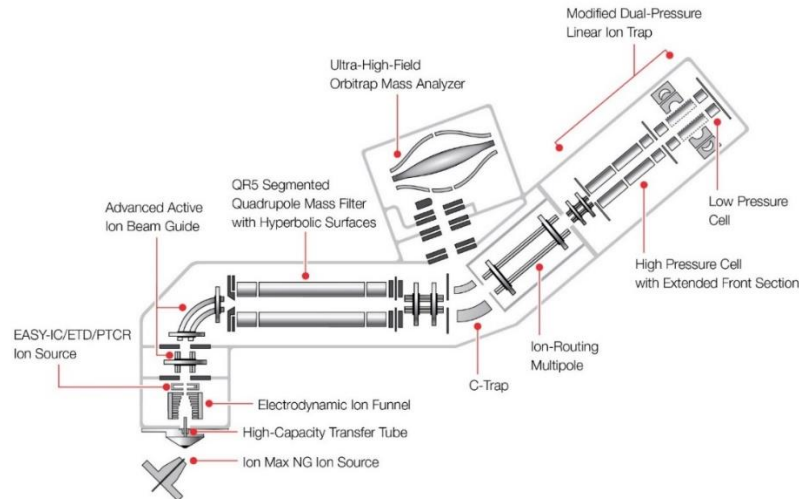


Figure 22 : Schéma d'un Orbitrap Tribrid Eclipse (ThermoFisher Scientific).

### (b) Le Tims TOF Pro, Bruker

La spécificité du Tims TOF Pro<sup>TM</sup> du constructeur Bruker est l'addition d'une dimension de séparation à l'intérieur du spectromètre de masse grâce à la mobilité ionique par TIMS (pour « *Trapped Ion Mobility Spectrometry* ») et une analyse par temps de vol dans le TOF (pour « *time of flight* »). La mobilité ionique (IM pour « *ion mobility* ») est une propriété intrinsèque d'un ion/une molécule qui est utilisée en combinaison avec d'autres paramètres comme le rapport masse/charge ou le temps de rétention. Elle est mesurée expérimentalement en mesurant la dérive d'un ion entraîné par un flux de gaz neutre à travers un champ électrique à contre-courant<sup>79</sup>. Cette vitesse d'analyte  $v_d$  en régime permanent permet de définir la mobilité ionique  $K$  (Équation 2) qui est donc une mesure de la friction liée à un observable : le temps  $t_d$  que les ions mettent à parcourir la longueur  $l$  de la cellule de mobilité. Dans cette formule,  $E$  est le champ électrique<sup>79</sup>.

$$K = \frac{v_d}{E} = \frac{l}{t_d E}$$

Équation 2 : Équation du calcul de la mobilité ionique.

Ceci permet l'ajout d'une dimension de séparation et d'un analyseur en plus. Le TIMS est composé de 3 régions principales : l'entonnoir d'entrée, la section de l'analyseur de mobilité et l'entonnoir de sortie (Figure 23)<sup>80</sup>. En premier lieu, les ions sont injectés et focalisés vers la section de séparation de la mobilité par un entonnoir à ions. Ensuite, un champ électrique croissant passe dans la section axiale tandis qu'une RF est appliquée aux électrodes ce qui

## INTRODUCTION : L'analyse de cellules uniques

confine les ions radialement. Les ions sont séparés selon le rapport taille/charge et sont piégés dans les régions de la section où leur force de dérive est compensée par la force du champ électrique. Puis, en phase d'élué, les ions sont élués des rapports taille/charge plus élevés aux plus faibles<sup>80</sup>. Après une séparation selon la section efficace de collision (ccs *collisional cross section*) des ions le TIMS peut être facilement couplé à un analyseur MS. Dans le Tims TOF Pro, cet analyseur est un temps de vol ou TOF.

Cette séparation en TIMS peut être utilisée afin de fragmenter les ions en série en mode MS/MS également appelé PASEF (pour « *Parallel Accumulation – Serial Fragmentation*»). La séparation par TIMS permet aux précurseurs de même section efficace de collision (ccs) de s'immobiliser lorsque les forces exercées par le flux gazeux et le champ électrostatique s'équilibrent. Ils sont ensuite élués séquentiellement en abaissant la force du champ électrique vers le spectromètre de masse pour leur analyse en MS puis MS/MS. Les spectres MS des précurseurs et MS/MS des fragments sont alignés selon les ccs.

Le TOF ou *time of flight*, sépare les ions en fonction de leur vitesse ou temps de vol qui est caractéristique de leur  $m/z$ <sup>59</sup>. En effet, les ions qui arrivent à un temps  $t$  donné dans l'analyseur, sont tous initialement accélérés par un potentiel fixe. Comme tous les ions de même charge  $z$  obtiennent la même énergie cinétique, la vitesse des ions est inversement proportionnelle à la racine carrée de  $m/z$ . Après avoir été accélérés, les ions parcourent une distance fixe, généralement comprise entre 0,5 et 2 mètres, avant d'atteindre le détecteur. Ainsi, en mesurant le temps qu'il faut pour atteindre le détecteur après la formation de l'ion, on peut déterminer le  $m/z$  de l'ion<sup>59</sup>. Une autre caractéristique de l'analyseur dans cet appareil est qu'il contient un réflecteur. À mi-chemin dans l'analyseur, les ions entrent dans un miroir électrostatique qui les renvoie sur une seconde course vers le détecteur. Ce réflecteur permet d'augmenter la résolution de l'analyseur en compensant les différentes de vitesse initiale entre les ions de même  $m/z$ .

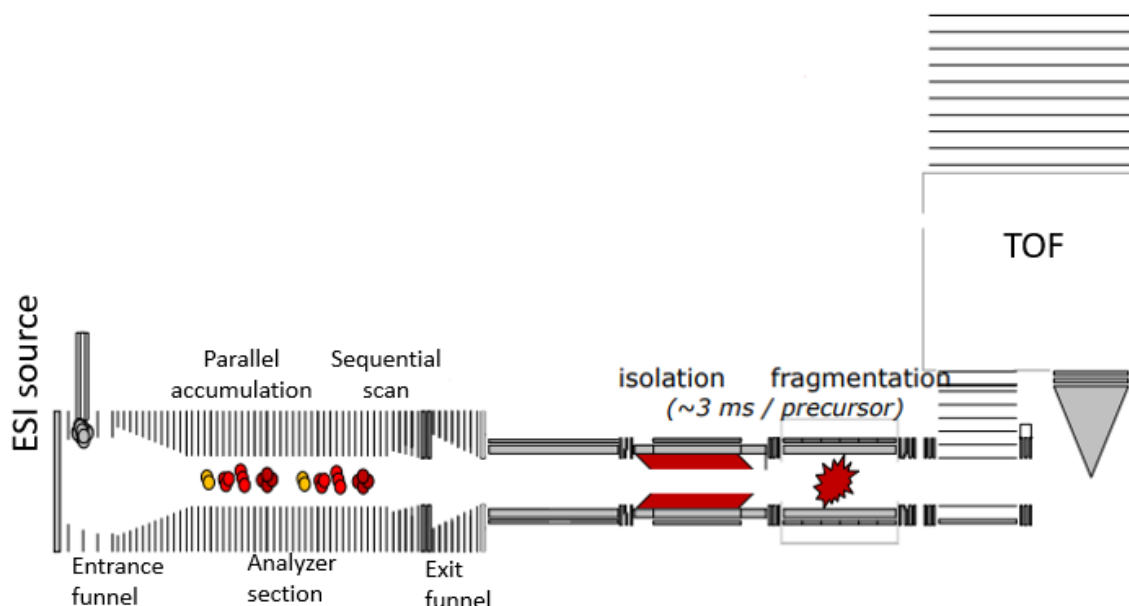


Figure 23 : Schéma d'un TIMS TOF Pro (Bruker).

#### 1.D.4. La quantification des protéines

Aujourd'hui, l'identification d'un nombre maximal de protéines dans un échantillon donné n'est plus le seul objectif des études protéomiques. Dans le domaine de la protéomique clinique notamment, l'étude des profils d'expression protéique variables peut fournir des informations utiles sur les biomarqueurs et les cibles médicamenteuses liés à une maladie particulière.

La protéomique sans marquage ou « *label free* » est devenue une méthode de choix pour les études quantitatives de protéomique clinique à haut débit. Les approches sans marquage de l'échantillon peuvent être divisées en deux stratégies de quantification présentées rapidement (Figure 24). En quantification basée sur le comptage des spectres, l'abondance des peptides et des protéines peut être estimée sur la base du nombre de d'occurrences de spectres MS/MS acquis pour un même peptide. Selon l'hypothèse établissant que le nombre de spectres MS/MS par peptide augmente avec son intensité, une estimation de la quantité relative entre différents échantillons de la protéine associée est possible <sup>81</sup>. Cependant, comme cette quantification repose sur un simple comptage des spectres acquis plutôt que sur la mesure de données physiques, la méthode du comptage spectral est controversée <sup>82</sup>. En quantification basée sur l'intensité ionique, les variations de l'abondance des peptides sont déterminées en mesurant et en comparant les intensités des pics mono-isotopiques dans un chromatogramme d'ions extraits (XIC pour « *extracted ion chromatogram* ») et leur aire sous la courbe (AUC pour « *area under curve* ») peut être calculée. Les aires des pics chromatographiques présentent théoriquement une corrélation linéaire avec l'abondance des protéines, ce qui rend leur mesure possible pour les études quantitatives <sup>83,84</sup>. Les limitations aux approches de quantification relative par mesure de l'ion parent sont de plusieurs ordres. Tout d'abord il se peut que dans une des conditions, une espèce passe sous le seuil de détection ce qui proscrit toute quantification relative. Lors d'analyse sans marquage (dites « *label free* ») les échantillons sont analysés séparément. Il est possible que dans un échantillon d'autres espèces co-éluées viennent perturber le signal du peptide d'intérêt et suppriment ou à l'inverse se superposent à l'espèce étudiée. Enfin les variations entre enregistrements peuvent aussi biaiser les résultats.

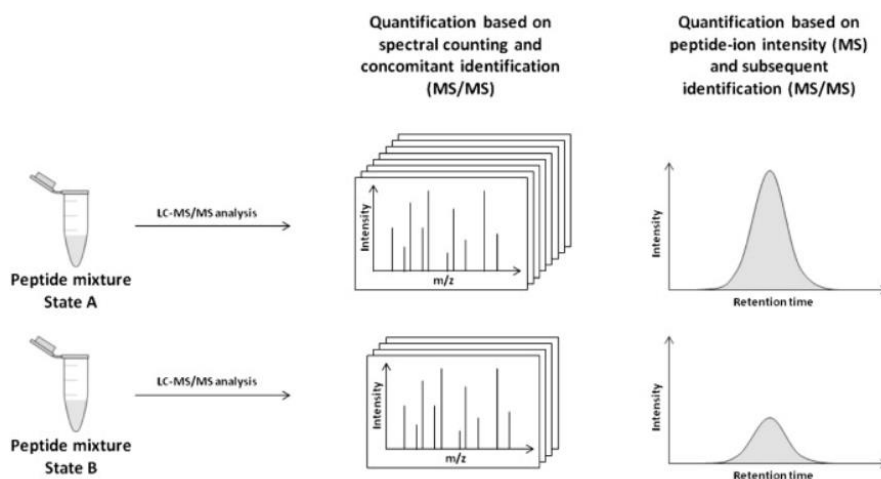


Figure 24 : Représentation de deux stratégies de quantification et leur application dans deux conditions A et B. Les conditions A et B sont comparées et les deux résultats de quantification indiquent une plus grande abondance des peptides dans la condition A <sup>85</sup>.

## INTRODUCTION : L'analyse de cellules uniques

Afin de multiplexer l'analyse des échantillons en spectrométrie de masse, notamment pour s'affranchir de précédentes limitations, il est possible de marquer les protéines ou les peptides à différents niveaux de la préparation d'échantillons. Il existe plusieurs stratégies de marquage choisies selon la question biologique posée (Figure 25)<sup>86</sup>. Dans la première stratégie, les intensités des ions précurseurs en MS sont utilisées pour quantifier les analytes en rapportant l'intensité du signal de l'analyte à celle d'une molécule de référence qui est un peptide standard marqué avec des isotopes stables lourds. Il est aussi possible de comparer les intensités relatives des fragments au niveau des spectres MS/MS comme avec les marqueurs iTRAQ pour « *isobaric tag for relative and absolute quantification* » ou TMT pour « *tandem mass tags* » ou marqueur de masse en tandem. Ce sont des marqueurs isobares, ayant la même masse globale, composés d'un groupe réactif qui s'attache aux peptides, un groupe balance et qui produisent des ions rapporteurs de masse différente après leur fragmentation en MS2 (Figure 26). Le rapport relatif de l'intensité de ces ions rapporteurs représente l'abondance relative de la molécule marquée<sup>87</sup>. Ces marquages chimiques permettent de rassembler jusqu'à 16 échantillons différents dans la même analyse. Cependant, le coût de l'expérience n'est pas négligeable<sup>88</sup>.

De plus la superposition peut provoquer la co-sélection de plusieurs espèces (portant toutes le même type de marquage), ce qui va induire un effet de lissage atténuant les variations individuelles par moyennage. En effet tous les peptides produisent le même jeu d'ions rapporteurs qui vont s'additionner dans les spectres MS/MS. Pour augmenter la spécificité de la quantification, une solution proposée est de réaliser une seconde fragmentation des ions en mode MS3 de fragments spécifiques du peptide d'intérêt, mais cela a un coût en temps d'acquisition et en sensibilité<sup>89</sup>.

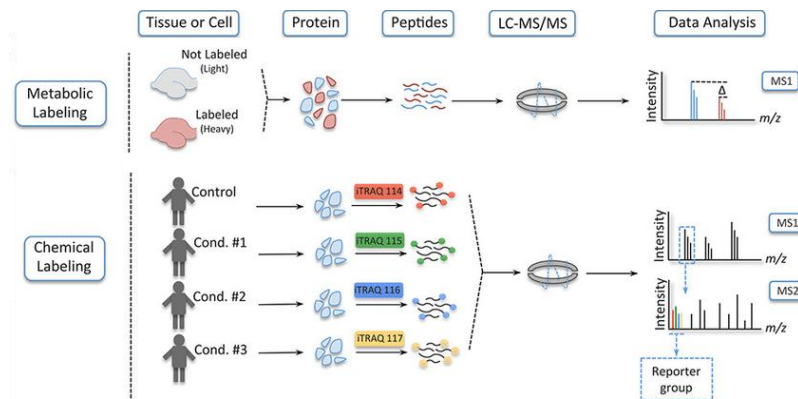


Figure 25 : Stratégie pour les analyses quantitatives par l'utilisation d'isotopes lourds (en haut) et de marqueurs de masse (en bas). Dans la figure, un exemple de 4 canaux est montré avec des isotopes aux m/z égales à 114, 115, 116 et 117.

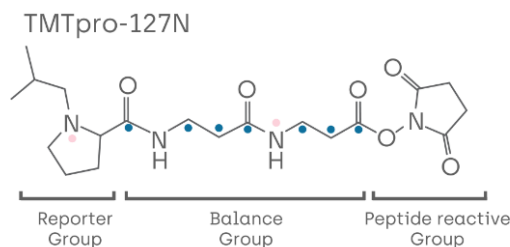


Figure 26 : Structure des marqueurs TMT pour « *Tandem mass tags* ».

### 1.D.5. Le traitement des données de spectrométrie de masse

L'analyse des données de protéomique classique se fait à l'aide de différents logiciels utilisés pour un processus d'identification et de quantification des spectres obtenus par spectrométrie de masse (Figure 27).

L'information contenue dans les spectres MS et MS/MS peut permettre de remonter à l'information protéique. Brièvement, la digestion et la fragmentation *in silico* de toutes les protéines de la banque de données utilisée va permettre de générer des profils MS/MS théoriques qui seront comparés aux données expérimentales grâce à des moteurs de recherche et selon les conditions de l'expérience. Dans cette étude, les moteurs de recherche utilisés sont principalement Mascot et MaxQuant<sup>90,91</sup>. Ces logiciels fonctionnent à partir de paramètres définis par l'utilisateur : l'erreur de masse tolérée sur l'ion parent et sur les ions fils, l'enzyme de digestion utilisée (coupures spécifiques ou non spécifiques), les modifications des acides aminés, le type de série de fragmentation recherchée et le format de la liste de masses. Une banque de séquences virtuelles appelée « *decoy* » générée aléatoirement à partir de séquences inversées est également utilisée et permet de vérifier le taux d'erreurs et de faux positifs de l'analyse. Elle est construite afin de ne contenir aucune séquence connue dans le modèle considéré.

De plus, les données LC-MS brutes générées dans les expériences doivent être retraitées (détection des caractéristiques, alignement des temps de rétention, normalisation des intensités MS, sélection des pics, réduction du bruit). Des logiciels comme MaxQuant permettent d'effectuer l'ensemble de ces opérations. Nous n'avons pas cherché dans le cadre de ce projet à évaluer d'autres solutions. Par la suite, ces données peuvent être retraitées pour être normalisées, et comparées. Des logiciels comme Excel, Perseus ou des packages en R peuvent être utilisés.

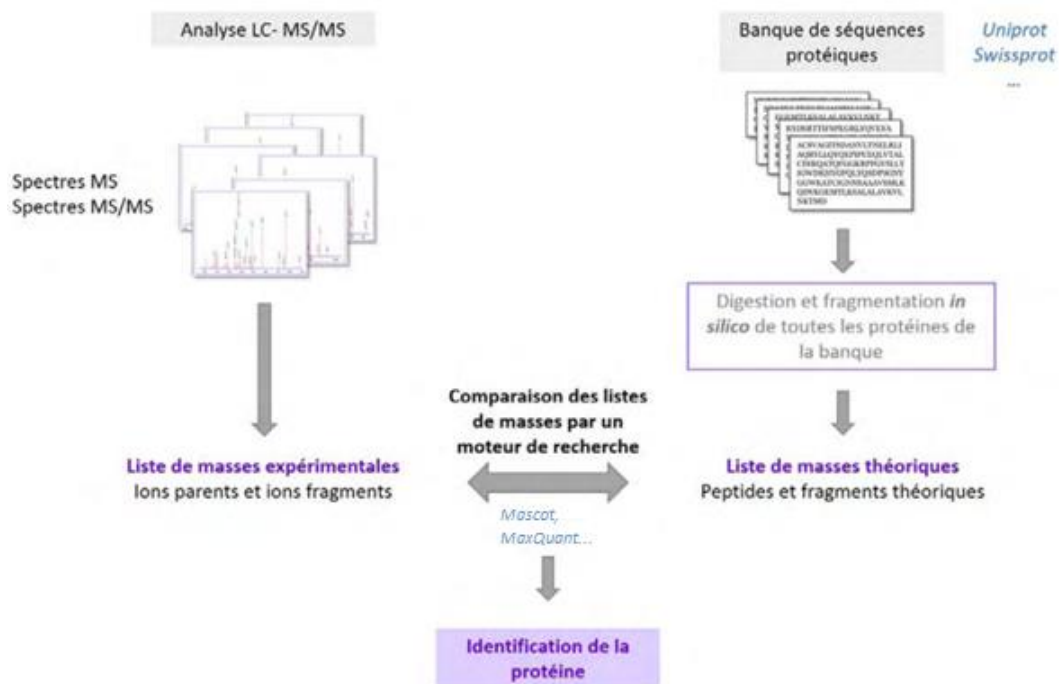


Figure 27 : Analyse des données expérimentales des données obtenues par LC-MS/MS.

## **Chapitre 2 Défis et problèmes de la protéomique en cellule unique**

<b>Chapitre 2 Défis et problèmes de la protéomique en cellule unique.....</b>	<b>29</b>
<b>2.A Un problème de taille... ..</b>	<b>30</b>
<b>2.B Les défis de la protéomique classique .....</b>	<b>31</b>
<b>2.C La différence avec l'analyse des ARNs .....</b>	<b>32</b>
<b>2.D Les propriétés des analytes.....</b>	<b>33</b>



## INTRODUCTION : Défis et problèmes de la protéomique en cellule unique

Les défis en protéomique de la cellule unique sont nombreux et divers.

### 2.A Un problème de taille...

Une cellule HeLa a un volume moyen de  $3000 \mu\text{m}^3$ , un diamètre de  $15 \mu\text{m}$  et contient  $10^9$  protéines soit environ  $100 \text{ pg}$  de protéines<sup>92,93</sup>. Les lymphocytes et monocytes utilisés dans cette étude font entre  $7$  à  $10 \mu\text{m}$  de diamètre et  $16$  à  $23 \mu\text{m}$  de diamètre respectivement en fonction de leur état de croissance<sup>94</sup>. Il est possible d'estimer que les lymphocytes étudiés contiennent environ  $50 \text{ pg}$  de protéines et les monocytes étudiés environ  $100 \text{ pg}$  de protéines<sup>95</sup>. Le tableau suivant indique les équivalences entre les cellules étudiées et la quantité de protéines (Tableau 1).

<i>Quantité de protéines</i>	<i>Lymphocytes : équivalent en nombre de cellules</i>	<i>Monocytes : équivalent en nombre de cellules</i>
<i>200 ng</i>	4000	2000
<i>10 ng</i>	200	100
<i>2 ng</i>	40	20
<i>1 ng</i>	20	10
<i>500 pg</i>	10	5
<i>100 pg</i>	2	1
<i>50 pg</i>	1	

Tableau 1 : Quantités en protéines et leur nombre de cellules correspondant.

Dans les analyses de spectrométrie de masse classiques, l'échantillon de référence est un digestat de protéines de cellules HeLa de  $200 \text{ ng}$  ce qui correspond à environ  $2000$  à  $4000$  cellules. Avec ce digest de référence, une analyse LC-MS/MS identifie en moyenne  $3000$  protéines avec notre setup expérimental. Notre mise au point a été effectuée sur des injections réalisées à partir d'échantillons standard contenant entre  $20$  et  $400 \text{ ng}$  de protéines. Bien sûr, la sensibilité des instruments de mesure et l'amélioration de la préparation et séparation de l'échantillon doivent permettre de descendre la limite de détection en spectrométrie de masse à quelques attomoles<sup>96</sup>.

Pour s'approcher des conditions analytiques nécessaires pour les analyses de cellules uniques, il faut commencer par diminuer la quantité d'échantillons de départ et donc travailler avec un ordre de grandeur d'échantillon  $1000$  fois plus petit que d'habitude. L'analyse robuste de ces échantillons de très faible concentration constitue l'un des principaux défis à relever pour le développement actuel de nouveaux processus protéomiques basés sur la spectrométrie de masse, de techniques de préparation et de séparation des échantillons.

## 2.B Les défis de la protéomique classique

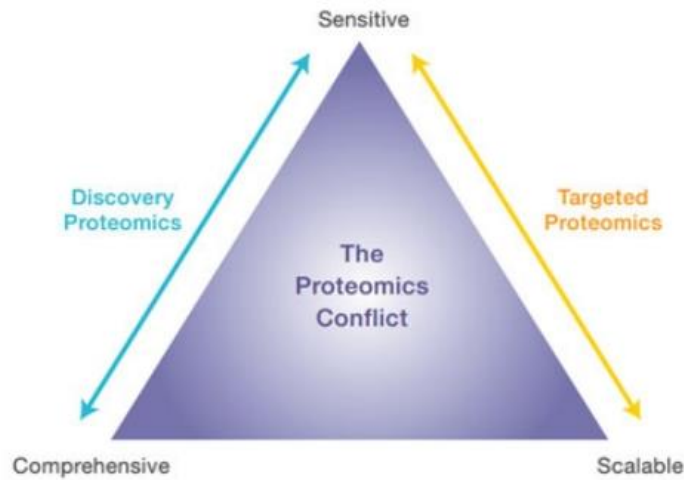


Figure 28 : Conflits en protéomique pour l'obtention d'une analyse à la fois sensible, à haut débit et exhaustive.

Les défis de la protéomique en cellule unique débutent par les défis de la protéomique classique. En effet, selon le schéma de Rogers et Bomgardner, il est impossible d'optimiser la sensibilité, le débit et l'exhaustivité de l'analyse protéomique simultanément (Figure 28). Il faut donc prioriser.

Par exemple, les stratégies de protéomique ciblée optimisent la sensibilité et l'évolutivité en limitant le nombre de caractéristiques surveillées. Les stratégies de découverte en protéomique tentent l'optimisation d'analyses sensibles et exhaustives avec peu d'échantillons. Rogers et Bomgardner notent qu'une analyse complète avec un débit raisonnable est rendue possible par le multiplexage des échantillons avec des réactifs de marquage de masse <sup>42</sup>.

Comme discuté dans la partie (« importance des protéines »), les abondances protéiques dans les cellules de mammifères varient avec un ordre de grandeur allant jusqu'à 10. De plus, le protéome de ces cellules est composé de 22 protéines très abondantes et de milliers de protéines à des concentrations beaucoup plus faibles. Il est donc nécessaire de noter que l'identification et la quantification de l'ensemble du protéome grâce aux techniques de spectrométrie de masse et de séparation des peptides est encore un défi difficile à atteindre à ce jour (Figure 29). Les protéines cellulaires couvrent une large gamme d'expression et les technologies actuelles n'échantillonnent qu'une fraction de toutes les protéines présentes dans un échantillon. En effet, la couverture pour les organismes supérieurs dépasse rarement 10 %. Pour la quantification des protéines, ces chiffres sont nettement inférieurs, car la qualité des données requise pour la quantification dépasse de loin celle de l'identification des protéines <sup>82</sup>.

## INTRODUCTION : Défis et problèmes de la protéomique en cellule unique

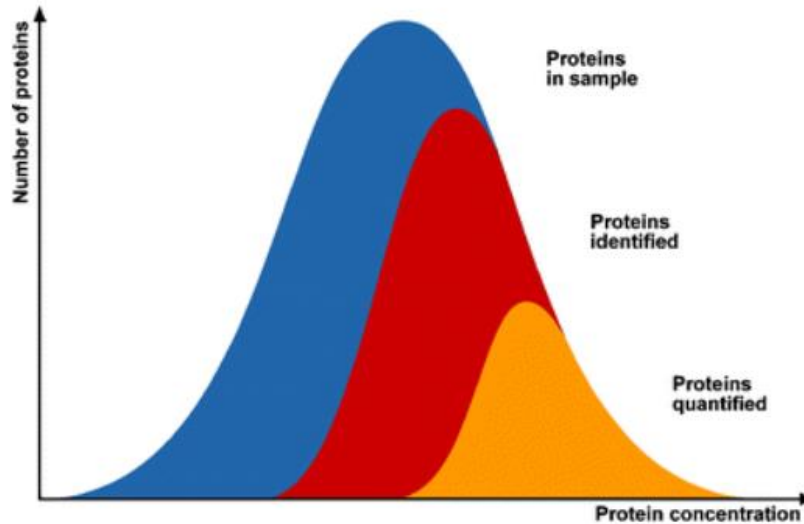


Figure 29 : Schéma représentant la fraction des protéines quantifiées ou identifiées sur un protéome total par spectrométrie de masse en protéomique.

### 2.C La différence avec l'analyse des ARNs

Contrairement aux transcriptomes ou aux génomes, les protéomes sont plus diversifiés et plus dynamiques. À partir d'un gène, de nombreuses protéoformes dérivent en raison des modifications post-traductionnelles (Figure 30). Les protéoformes sont un ensemble de toutes les formes moléculaires d'une protéine produite à partir d'un gène. La quantification complète des protéoformes dans des milliers de cellules uniques peut permettre de modéliser les réseaux de signalisation avec plus de précision <sup>65</sup>. Par conséquent, des milliers, voire des millions d'espèces protéiques différentes sont trouvées dans des échantillons de traces <sup>97</sup>. En quantifiant des milliers de protéoformes dans des milliers de cellules uniques, cela permettrait de contourner les problèmes d'échantillonnages en réduisant la quantité d'échantillon nécessaire pour chaque analyse.

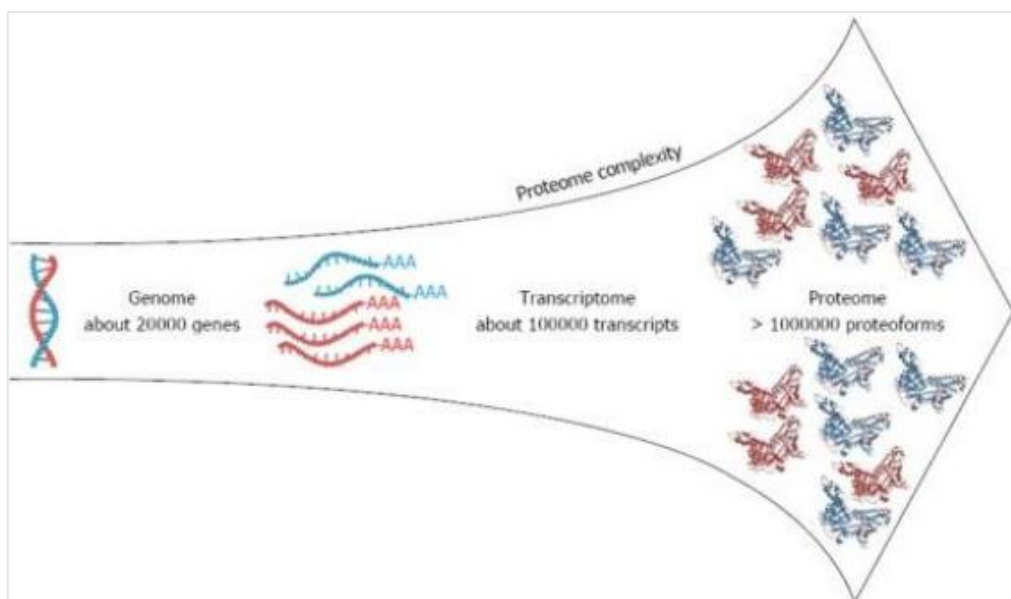


Figure 30 : Représentation de la complexité du protéome comparé au génome ou au transcriptome <sup>98</sup>.

## INTRODUCTION : Défis et problèmes de la protéomique en cellule unique

Cependant, en protéomique, nous n'avons pas la possibilité d'extraire les protéines et de les multiplier pour les rendre détectables comme il est possible de le faire avec les ARN messagers en transcriptomique. Il est tout de même important de noter que le nombre de copies de protéines est beaucoup plus important que celui des ARN dans une cellule. Il y a 1000 à 10.000 copies d'une protéine par cellule pour 20 ARN messagers seulement.

### 2.D Les propriétés des analytes

La préparation d'échantillons est une des étapes les plus importantes en protéomique surtout en cellule unique. Elle présente beaucoup de défis dus notamment aux propriétés intrinsèques des analytes.

Le premier est la contamination. Les interférences ou contaminants non protéiques en spectrométrie de masse proviennent principalement des plastiques utilisés dans les nano-connecteurs, les capillaires, c'est l'exemple du polyéthylène glycol (PEG). D'autres proviennent des surfactants utilisés lors de la préparation des échantillons comme le PEG ou le polypropylène glycol (PPG). D'autres sont trouvés dans l'environnement ou les produits de soins pour la peau ou encore les polymères utilisés en microfluidique comme le polydiméthylsiloxane (PDMS). Ces contaminants sont particulièrement perturbants car ils entraînent une suppression des ions et interfèrent souvent avec les ions cibles d'intérêt. La chimie analytique moderne utilise des produits chimiques et des réactifs ultrapurs ainsi que des récipients ultra-propres pour la manipulation des échantillons afin de minimiser toute interférence potentielle et indésirable <sup>99</sup>.

De plus, lors de l'étude de cellules humaines, des contaminants protéiques présents dans l'environnement comme les kératines sur la peau ou dans l'air par exemple peuvent devenir un véritable problème s'ils saturent les analyses. Sur de grandes quantités d'échantillons, cette contamination n'est pas visible mais elle commence à devenir gênante lorsque de très petites quantités d'échantillons sont mises en jeu. En effet, la miniaturisation des procédures de transfert d'échantillons et des outils de manipulation augmente de façon exponentielle le rapport entre la surface et le volume de l'échantillon et, par conséquent, toute interférence résultant de surfaces contaminées se multipliera également <sup>100</sup>. La solution idéale serait une automatisation complète de la préparation d'échantillons afin d'éviter tout contact avec les échantillons.

Le second défi est la miniaturisation de cette préparation d'échantillons. Elle est essentiellement impactée par deux paramètres : la diminution des concentrations et la perte des échantillons pendant leur préparation.

En effet, l'impact des concentrations d'échantillons et de réactifs sur l'efficacité des réactions chimiques et enzymatiques lors de la préparation des échantillons est non négligeable. Si une cellule unique est traitée dans le volume conventionnel du traitement d'un million de cellules (environ 100µl), le contenu protéique de cette cellule serait dilué par un facteur de  $10^8$ . Ryan Kelly compare cette dilution avec la dilution d'un raisin dans une très large piscine. Cette dilution devient problématique quant à l'ajout des réactifs pour les réactions de réduction, alkylation et digestion pendant la préparation des échantillons <sup>53</sup>.

## INTRODUCTION : Défis et problèmes de la protéomique en cellule unique

Prenons l'exemple de la digestion tryptique. Cette digestion suit la cinétique de Michaelis-Menten qui décrit une réaction catalysée par une enzyme agissant sur un substrat unique pour donner un produit (Équation 3). Selon cette équation, dans des conditions de faible concentration de substrat, la vitesse de digestion devient directement proportionnelle à la concentration de l'enzyme (trypsine) et du substrat (échantillon de protéines) (Équation 4)<sup>101</sup>.

$$v = k_{\text{cat}} [\text{E}]_0 \frac{[\text{S}]}{K_M + [\text{S}]}$$

Équation 3 : L'équation de Michaelis-Menten (ou de Michaelis-Menten-Henri) permet de décrire la cinétique d'une réaction catalysée par une enzyme agissant sur un substrat unique pour donner irréversiblement un produit.

$$[\text{S}] \ll K_M, \quad v = k_{\text{cat}} [\text{E}]_0 \frac{[\text{S}]}{K_M}$$

Équation 4 : L'équation de Michaelis-Menten pour une concentration de substrat très faible

Ainsi, les conditions de digestion qui fonctionnent bien pour les échantillons en « *bulk* » dans des volumes de traitement standard de 100µl peuvent devenir inefficaces lors du traitement d'une cellule unique dans le même volume en raison de la concentration fortement réduite en protéines. En effet, la dilution mentionnée précédemment a pour conséquence une très faible concentration en protéines sur laquelle la trypsine aura du mal à agir. L'augmentation considérable de la concentration de trypsine peut compenser partiellement les effets de la réduction des concentrations de substrat, mais il y a une limite au-delà de laquelle l'autolyse et l'activité chymo-tryptique résultant de l'excès de trypsine interfèrent avec l'analyse<sup>53,102</sup>.

Le plus grand défi avec la miniaturisation reste la perte d'échantillons. En effet, plus la quantité de matière analysée par LC-MS est faible, plus les identifications protéiques se font rares. Cela est dû à une adsorption des protéines et des peptides sur les surfaces. Les expériences de protéomique sur des échantillons de traces sont entravées par des pertes indésirables dues à l'adsorption des protéines, principalement en raison de l'immobilisation des analytes sur les surfaces de contact (plastique, verre) (Figure 31)<sup>97</sup>. Cette adsorption des protéines et des peptides limite le rendement de récupération et est délétère pour identifier le maximum de peptides.

Les pertes dues à l'adsorption nuisent davantage à la sensibilité et à la précision des études unicellulaires que des études en « *bulk* ». En effet, ces pertes peuvent être négligeables lors des études de grandes quantités d'échantillons mais pas sur des quantités d'échantillon réduites. De plus, les transferts multiples des échantillons sur plusieurs surfaces multiplient ces pertes.

Pour les formes protéiques de faible abondance, l'adsorption non spécifique de ces protéines et de leurs peptides associés sur les surfaces de contact est un problème majeur et agit comme un seuil de sensibilité<sup>53</sup>.

## INTRODUCTION : Défis et problèmes de la protéomique en cellule unique

D'ailleurs, l'étude des surfaces du support en contact avec les échantillons est souvent négligée et pourtant très importante. Sun et Kumar expliquent que l'adsorption spontanée en surface lors d'interactions moléculaires est fréquente car elle est énergétiquement favorisée et diminue l'énergie de surface <sup>103</sup>. Ce processus dépend des propriétés physicochimiques des protéines, des solvants et des surfaces <sup>97</sup>. Ainsi, un objectif important de la science des matériaux est de produire des matériaux répulsifs aux peptides et aux protéines, qui seront utilisés dans des supports spécifiques conçus pour minimiser les surfaces de contact <sup>104</sup>.

De ce fait, la réduction du volume de réaction offre un double avantage d'augmentation des concentrations d'échantillons et de réduction de l'exposition aux surfaces et donc l'adsorption des protéines sur ces surfaces, ce qui devrait considérablement améliorer la protéomique en cellule unique <sup>53</sup>.

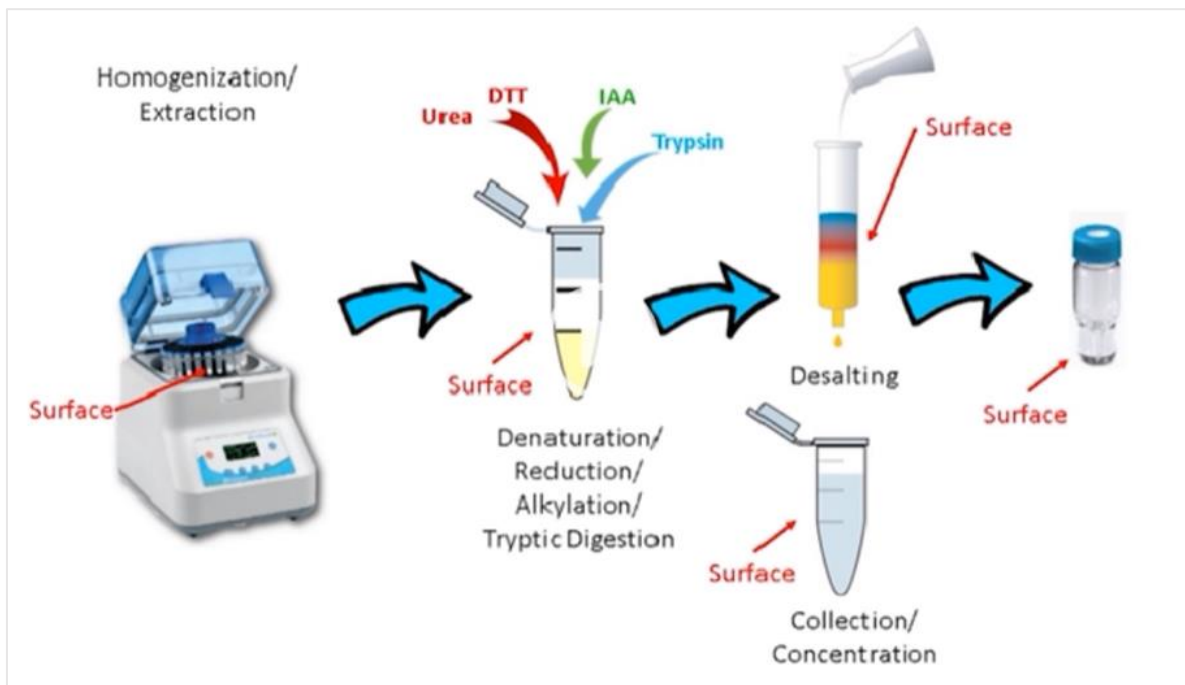


Figure 31 : Exemples de surfaces sur lesquelles les protéines et peptides peuvent s'adsorber pendant la préparation des échantillons en protéomique.

## Chapitre 3 Les avancées en protéomique de la cellule unique

<b>Chapitre 3 Les avancées en protéomique de la cellule unique.....</b>	<b>36</b>
<b>3.A Les solutions proposées.....</b>	<b>37</b>
3.A.1. La microfluidique et les protéines .....	37
3.A.2. La microfluidique dans ce projet .....	40
(a) La micro-fabrication.....	40
(b) La nature des matériaux utilisés.....	41
(c) Les techniques de micro-fabrication utilisées dans ce projet .....	42
3.A.3. Traitements de surface existants .....	44
(a) L'activation de surface.....	44
(b) La physisorption.....	45
(c) Les brosses de polymères .....	46
<b>3.B Les débuts de la protéomique en cellule unique .....</b>	<b>47</b>
3.B.1. Approches basées sur les anticorps .....	48
(a) CyTOF.....	48
(b) Western Blot en cellule unique .....	49
(c) siMOA .....	50
(d) CITE-Seq.....	50
(e) Abseq .....	50
3.B.2. Approches basées sur la spectrométrie de masse .....	51
(a) iPAD-1 .....	51
(b) « Oil Air Droplets » (OAD).....	52
(c) SCOPE-MS .....	52
(d) NanoPOTs .....	54
(e) ProteoCHIP.....	56
3.B.3. Bilan des stratégies actuelles pour l'analyse protéomique de très petits nombres de cellules	57

## INTRODUCTION : Les avancées en protéomique de la cellule unique

Comme discuté précédemment, pour répondre à de nombreux défis médicaux et biologiques, l'analyse du protéome sur cellule unique ou proche de la cellule unique est devenue une approche incontournable de la protéomique car elle fournit des informations sur le changement dynamique ou l'hétérogénéité cellulaire<sup>105</sup>. En théorie, elle doit permettre de quantifier des milliers de protéines et leurs modifications post-traductionnelles déconvoluées au niveau de la cellule unique. C'est comme si on étudiait la composition d'un smoothie, le goût du smoothie donne une idée sur la composition générale mais on perd une information importante sur la composition exacte notamment les ingrédients les plus rares (Figure 32).



Figure 32 : Le smoothie représente un protéome total dont l'analyse donne une idée des ingrédients mais sans la profondeur d'analyse, on ne connaîtrait jamais la quantité ou la représentation de chaque fruit du smoothie (métaphore utilisée par Ryan Kelly en conférence).

Les systèmes de détection capables d'identifier et d'analyser les protéines de manière efficace et rapide peuvent considérablement améliorer les capacités de diagnostic moléculaire<sup>106</sup>.

Cependant, plusieurs défis technologiques sont apparus lors du développement de ces méthodologies en cellule unique. Ils portent tous sur la petitesse des échantillons, le risque de pertes non spécifiques et les limites de détection et de quantification nécessaires. Les solutions proposées seront discutées dans ce chapitre.

### 3.A Les solutions proposées

#### 3.A.1. La microfluidique et les protéines

Pour relever les défis posés par le niveau de sensibilité requis pour l'utilisation de la MS et des techniques séparatives en amont, la miniaturisation est incontournable. Certains verrous peuvent être levés grâce à la microfluidique. Des systèmes microfluidiques ont été développés en miniaturisant des systèmes existants dans des domaines comme la biologie, la recherche contre les maladies infectieuses et même la sécurité alimentaire<sup>107</sup>. Les dispositifs miniaturisés les plus complexes sont des « laboratoires sur puce » (*Lab-on-a-Chip*) et intègrent plusieurs opérations dans une plateforme microfluidique unique. Ils peuvent être adaptés à des besoins analytiques spécifiques<sup>108</sup>.

La miniaturisation permet d'utiliser des volumes d'échantillons plus petits, de réduire les distances de diffusion tout en limitant l'adsorption sur les surfaces de dimensions réduites, de réaliser des analyses à haut débit, d'automatiser les systèmes et de les paralléliser. La limitation de la manipulation manuelle réduit le risque de contaminations, qui a un impact



## INTRODUCTION : Les avancées en protéomique de la cellule unique

majeur sur la fiabilité des analyses protéomiques. Enfin, la capacité des dispositifs microfluidiques à manipuler des cellules individuelles a ouvert les possibilités d'analyse des protéines au niveau de la cellule unique <sup>108</sup>.

Différentes méthodes de micro-fabrication sont utilisées pour la création de plateformes microfluidiques bio-analytiques, qui seront discutées dans le paragraphe suivant. Selon les besoins, divers matériaux ont été utilisés comme substrats, bien que les matériaux à base de silice ou de polymères soient généralement préférés <sup>108</sup>. La méthode de prototypage la plus courante a recours à l'utilisation de polydiméthylsiloxane (PDMS), car il est polyvalent et peu coûteux. En général, le design du dispositif est gravé à partir d'un substrat de silicium par dépôt d'une résine photorésistante, puis le PDMS est durci et collé sur une lame de verre. Une complexité supplémentaire peut être ajoutée au dispositif microfluidique en intégrant des valves et des chambres pour éliminer les problèmes d'ajustement, de fuite et de blocage <sup>107</sup>. Ceci sera approfondi dans les paragraphes suivants.

Dans la littérature, Chao Tzu-Chiao et Hansmeier Nicole ont établi une liste des différents dispositifs microfluidiques utilisés pour la protéomique à haut débit <sup>108</sup>. Bien qu'il n'existe pas encore un dispositif qui rassemble toutes les étapes de la culture cellulaire à l'analyse LC-MS, il existe plusieurs dispositifs qui améliorent le rendement de chaque étape séparément. Pour la culture cellulaire, des exemples de chambres de culture microfluidiques sont présentés par Groisman *et al.*, qui permettent la culture en parallèle de plusieurs cellules confinées dans des chambres <sup>109</sup>. Une alternative à ce système avec 100 micro-chambres de culture a vu le jour commercialisée par l'entreprise TebuBio®, sous la forme du Smart Aliquoter dont l'utilisation sera discutée plus tard (voir « isolement des cellules »). Ensuite, des exemples de puces microfluidiques ont été développés pour les différentes lyses existantes (voir « introduction : lyse cellulaire »). La lyse chimique peut être réalisée en microfluidique en ajoutant une ligne d'entrée de tampon de lyse au dispositif et en permettant un mélange par diffusion de la solution cellulaire avec les réactifs de lyse dans les canaux ou dans des chambres de réaction comme dans l'exemple d'Irimia *et al.* ou celui de Jen *et al.* qui ont développé un système de lyse en parallèle pour plusieurs cellules HeLa (Figure 33) <sup>110,111</sup>. Il existe également des dispositifs de lyse en continu dans des canaux <sup>112</sup>. La lyse par sonication a également été adaptée en microfluidique <sup>113,114</sup>. Cependant, dans ces différents exemples, la cavitation par ultrasons est provoquée en utilisant des électrodes ou des pulsations laser souvent coûteuses et difficiles à mettre en place.

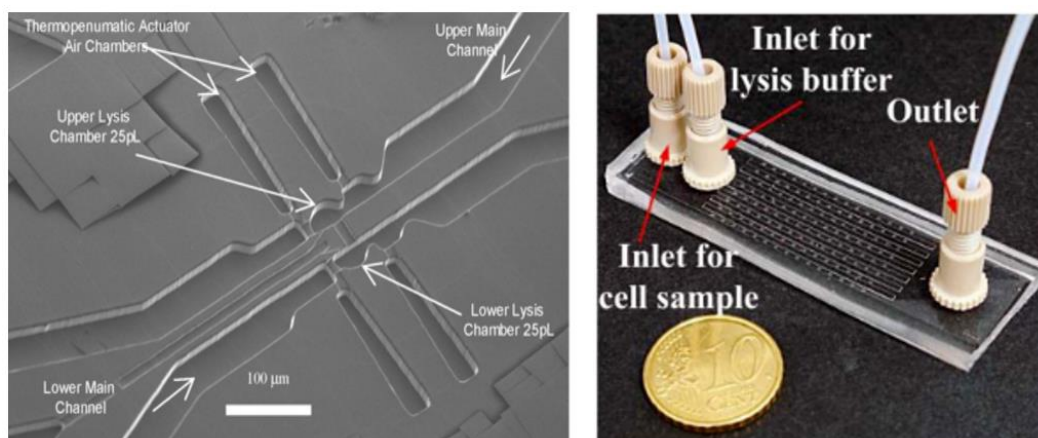


Figure 33 : Mise en place de l'étape de lyse dans des dispositifs microfluidiques montrée dans deux exemples (à gauche) Irimia *et al.* (à droite) Jen *et al.* <sup>110,111</sup>.

## INTRODUCTION : Les avancées en protéomique de la cellule unique

À la suite de l'étape de lyse, a lieu l'étape de digestion enzymatique dont les temps de réaction durent plusieurs heures et empêchent un haut débit d'analyse. Une possibilité pour réduire ce temps de digestion est de confiner les réactions dans des volumes plus petits pour augmenter l'interaction protéase-protéine<sup>108</sup>. En microfluidique, la digestion sur puce peut être classée en 2 approches principales, une basée sur l'utilisation d'enzyme immobilisée et l'autre avec une enzyme non immobilisée. Les dispositifs basés sur l'utilisation d'enzymes immobilisées permettent une digestion des plus rapides (10 secondes à 15 minutes)<sup>115</sup> par exemple être immobilisées sur des billes aimantées ou sur les parois des dispositifs microfluidiques en PDMS<sup>116,117</sup>. Un avantage des réacteurs à base de billes est la possibilité de changer les billes une fois que l'activité enzymatique se réduit. Cependant, l'intégration des billes dans les dispositifs peut être complexe et des rétentions non spécifiques et des effets de *carry over* ont pu être observés<sup>118</sup> *et al.*, grâce à un système d'hydrogels sur lesquels les protéases sont immobilisées<sup>119</sup>. D'autre part, les systèmes DMF (*digital microfluidics*) se prêtent très bien à la digestion enzymatique basée sur les enzymes non immobilisées car chaque gouttelette sert de chambre de réaction individuelle. Ainsi, le traitement parallèle des échantillons et le contrôle individuel du temps de digestion sont possibles. La fusion de gouttelettes de protéase avec l'échantillon donne des temps de digestion efficaces entre 2 et 10 minutes<sup>120</sup>.

Enfin, comme les débits de la nano ESI sont faibles, ils sont compatibles avec les dispositifs microfluidiques. Les premières interfaces microfluidiques-ESI étaient des puces en verre avec une ouverture (appelée « *channel opening* » sur la Figure 34) servant de sortie pour l'électro-nébulisation<sup>121,122</sup> (Figure 34). Certains systèmes, comme celui présenté par Ji *et al.*, analysent les peptides à la sortie de la LC et ont même été couplés à une source nano ESI faisant de ces dispositifs une interface de préparation des protéines couplée à la MS<sup>118</sup>.

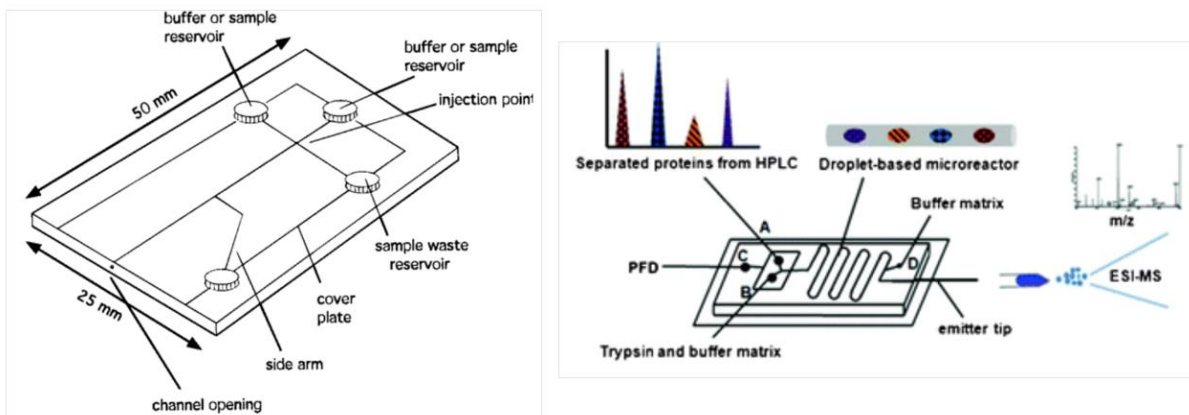


Figure 34 : Puces microfluidiques servant d'interface avec (à gauche) l'ESI par Ramsey *et al.* (à droite) avec la HPLC par Ji *et al.*

Toutes ces avancées ont permis l'amélioration et la miniaturisation des différentes étapes de préparation des échantillons en protéomique. Cependant, ces systèmes microfluidiques ne permettaient pas la prise en charge de protéomes complexes provenant de cellules différentes traitées en parallèle et de faire l'interface avec la MS sans trop de complexité dans leur mise en œuvre jusqu'à très récemment. En 2022, la publication d'iProChip a montré la possibilité d'intégrer le protocole dans une puce microfluidique très complexe couplée à une analyse des peptides en DIA et d'identifier en moyenne 1500 protéines pour 20 cellules de mammifères analysées en protéomique de la cellule unique (Figure 35)<sup>123,124</sup>.

## INTRODUCTION : Les avancées en protéomique de la cellule unique

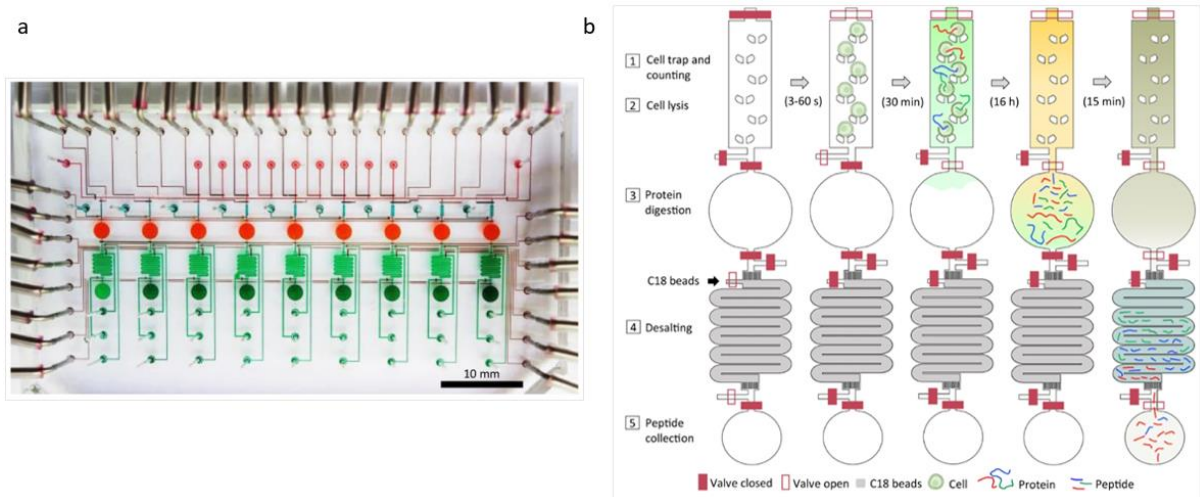


Figure 35 : iProChip permettant un protocole complet de protéomique dans une puce microfluidique (à gauche) photographie (à droite) schéma des étapes du protocole.

### 3.A.2. La microfluidique dans ce projet

#### (a) La micro-fabrication

La micro-fabrication est un ensemble de procédés permettant de fabriquer des objets dont les dimensions sont comprises entre le micromètre et le millimètre utilisés en microfluidique. Elle est basée sur des procédés comme la photolithographie, la gravure, l'usinage laser, le fraisage ainsi que la lithographie douce pour le moulage de polymères qui peuvent être utilisés seuls ou en combinaison. Ces techniques étant très sensibles à l'humidité et à la température, il est nécessaire de contrôler ces paramètres dans l'environnement de travail. De plus, la micro-fabrication doit être réalisée dans un environnement propre avec un minimum de particules au mètre cube.

Or, comme discuté dans la partie « propriétés des analytes », la contamination est le premier défi rencontré en protéomique des protéines mais également en micro-fabrication. En effet, en plus des contaminations de protéines et de polymères, la contamination par l'air et par l'humain est un paramètre non négligeable. En plus de l'opérateur qui est une source non négligeable de particules (protéines, poussières...), un échantillon typique d'air intérieur peut contenir entre 500 000 et 1 million de particules (poussières, insectes, aérosols) par mètre cube <sup>125</sup>. Ces particules ont tendance à s'adsorber sur les surfaces et à induire des défauts de fabrication.

Tout cela nécessite un environnement spécifique retrouvé dans des salles spécifiques appelées « salles propres » dans lesquelles la température est maintenue à 20°C, l'humidité à 25% et l'air est filtré en permanence pour maintenir le nombre de particules à faible niveau. En fonction du nombre de particules dans l'air, les salles propres sont classées en différents niveaux adaptés pour certaines techniques de micro-fabrication plutôt que d'autres. Par exemple, la salle de lithographie est spécifiquement utilisée pour la fabrication des *wafers*. La salle grise est utilisée pour le coulage de PDMS et le découpage des objets micro-fabriqués.

En microfluidique, les objets obtenus par micro-fabrication permettent la manipulation de très faibles quantités de fluides <sup>126</sup>. Ils peuvent inclure des structures mobiles (valves) ou des structures statiques (canaux microfluidiques, chambres de réaction) <sup>127</sup>.

(b) La nature des matériaux utilisés

Selon les conditions expérimentales, divers matériaux peuvent être utilisés comme substrats pour construire ces supports, bien que les matériaux à base de silice ou de polymères sont généralement préférés<sup>108</sup>. Le matériau le plus courant dans les plastiques de laboratoire est le polystyrène (PS). Cependant, le PS n'est pas utilisé en protéomique car incompatible avec les solvants, peu stable et induisant des pertes d'échantillons considérables. Le verre est également utilisé pour plusieurs matériels comme les lames de microscopie ou les flacons d'injection en LC-MS car il présente des propriétés similaires à celles de la silice fondue, couramment utilisée pour l'électrophorèse capillaire ou la chromatographie. Par rapport aux matériaux à base de silice, les substrats polymères sont plus polyvalents, car ils peuvent être facilement et à moindre coût être façonnés dans les structures requises. En protéomique, les tubes à faible adsorption ou les plaques multi-puits préférés sont fabriqués en polypropylène (PP). Le PP, dont la structure est  $(-\text{CH}_2-\text{CH}(\text{CH}_3)-)_n$ ,  $n$  étant le nombre de répétition du monomère, est un polymère thermoplastique semi-cristallin fabriqué à partir de la polymérisation de monomères de propylène  $[(\text{CH}_2=\text{CH}-\text{CH}_3)]$  en présence de catalyseurs suivant la catalyse de Ziegler-Natta (Figure 36). Le PP présente de nombreux avantages, il a un faible coût, est non toxique, très résistant, chimiquement inerte et recyclable. Il a été prouvé que le polypropylène était égal ou supérieur au polystyrène pour minimiser les pertes de récupération de peptides testés de différentes propriétés physico-chimiques<sup>128</sup>. Cependant, il reste hydrophobe et nécessite des traitements de surface pour le rendre « *low binding* » ou adapté pour limiter l'adsorption des peptides.

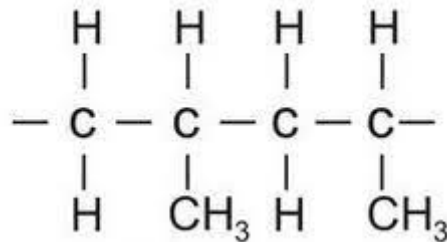


Figure 36 : Structure du polypropylène.

L'utilisation de dispositifs jetables est importante afin d'éliminer le risque de contamination croisée entre les échantillons. Cela signifie que les contenants doivent être simples et rapides à fabriquer avec un coût le plus faible possible. De ce point de vue, l'étude de dispositifs par prototypage en PDMS est plus intéressante<sup>108</sup>. Le polydiméthylsiloxane (PDMS) de structure  $\text{CH}_3[\text{Si}(\text{CH}_3)_2\text{O}]_n-\text{Si}(\text{CH}_3)_3$ ,  $n$  étant le nombre de répétition du monomère, est un élastomère, et reste le matériau le plus répandu et le plus polyvalent utilisé dans la construction de dispositifs microfluidiques, en particulier pour le prototypage rapide (Figure 37). L'une des raisons de la popularité du PDMS est sa méthode de fabrication particulièrement simple : le précurseur et le catalyseur du PDMS peuvent être coulés contre un moule sous forme liquide, suivi d'une étape de réticulation permettant d'aboutir à un matériau élastomérique. De nombreuses sociétés commercialisent du PDMS à deux composants, qui se compose d'un composant de base et d'un agent de durcissement. Le Sylgard 184 de Dow Corning est largement connu et utilisé pour la fabrication de microsystèmes<sup>129</sup>.

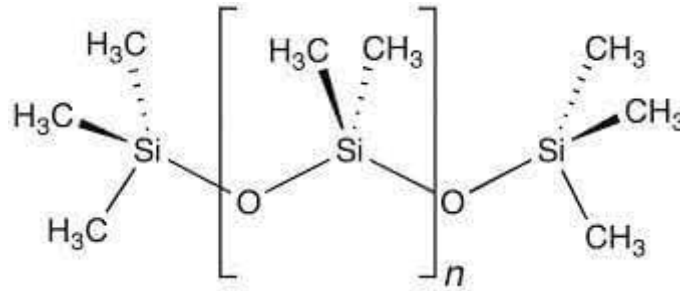


Figure 37 : Structure du PDMS, n représente un monomère.

Les avantages indéniables de l'utilisation du PDMS sont sa polyvalence, son faible coût, sa réticulation à basse température (70°C), son excellente fidélité de réplication et sa résolution nanométrique. Il est également transparent, biologiquement et chimiquement inerte, non toxique, biocompatible et a une bonne stabilité thermique <sup>130</sup>. De plus, les matrices utilisées comme supports pour créer les motifs avec le PDMS peuvent varier grâce à sa compatibilité. Grâce à son élasticité, la réplique en PDMS peut être retirée des moules sans endommager le matériau de base. Le PDMS se scelle également de manière réversible à lui-même ou de manière irréversible à d'autres matériaux, ce qui le rend adapté à la construction de puces hybrides hétéro-matérielles <sup>129</sup>.

Néanmoins, la surface du PDMS nécessite une fonctionnalisation pour une variété d'applications avant son utilisation. En effet, la surface du PDMS est naturellement hydrophobe ce qui favorise l'adsorption protéique <sup>131</sup>. Le PDMS a été décrit comme induisant des pertes d'adsorption plus élevées pour des molécules spécifiques telles que les protéines <sup>132</sup>. L'interaction est encore plus forte avec les analytes polaires en raison de la liaison hydrogène entre les parties « siloxane » du PDMS et les parties hydroxyle et carboxyle de l'analyte ou par des interactions polaires-polaires <sup>133</sup>. Cette interaction contenant-contenu doit être évitée pour minimiser la perte rédhibitoire de l'analyse des échantillons à l'état de traces. Beaucoup de techniques et traitements sont utilisés pour modifier la surface du PDMS, afin d'améliorer son hydrophobie qui limite son applicabilité pour les dispositifs en microfluidique <sup>129</sup>.

### (c) Les techniques de micro-fabrication utilisées dans ce projet

Les techniques de micro-fabrication choisies dépendent de la nature du matériau utilisé, du coût et de la précision nécessaire pour la fabrication des motifs souhaités. Il existe des méthodes dures comme la gravure, l'usinage ou la lithographie utilisant des matériaux comme le verre ou le silicium et des méthodes douces comme les méthodes de moulage utilisant des polymères ou des plastiques. Le principe est de produire un moule unique sur lequel il sera possible de faire de la réplication d'objets en masse. Dans ce projet, la fabrication du moule unique a été faite soit par photolithographie soit par micro-usinage. Celle-ci commence par le dessin à l'aide de logiciels de design des motifs qui seront utilisés en photolithographie ou du modèle qui sera utilisé pour creuser les moules en micro-usinage.

La photolithographie est une technique très précise (résolution de l'ordre du micromètre). Les étapes sont décrites dans le schéma suivant (Figure 38). Un film photorésistant (résine négative) est laminé uniformément sur un support lisse appelé *wafer* (qui est une plaque de silicium) puis exposé à la lumière à travers un masque optique. Ce masque est un support transparent sur lequel les motifs sont opaques et recouverts de chrome et sert à protéger une

## INTRODUCTION : Les avancées en protéomique de la cellule unique

partie du film des rayonnements incidents. Le film est exposé à des photons puis la polymérisation de la résine photosensible est activée ou inhibée (résine négative ou positive) selon les motifs du masque. Les parties non polymérisées sont ensuite supprimées du *wafer* dans une solution développante.

L'autre technique de micro-fabrication est le micro-usinage. Il s'agit d'une technique moins précise (100  $\mu\text{m}$ ) que la photolithographie. Dans ce projet, la micro-usineuse utilise un foret de micro-perçage dont la vitesse de rotation est très importante (supérieure à 24000 tours/minute) afin de creuser un bloc de laiton et le transformer en moule incluant le modèle préalablement dessiné. L'avantage réside dans la nature du moule qui sera très résistant aux chocs, comparé au *wafer* de photolithographie.

Les moules fabriqués par ces techniques représentent le reflet négatif de l'objet à fabriquer et pourront être réutilisés pour la suite du processus. Il existe plusieurs techniques de réplication. Dans ce projet, elle se fait par moulage, c'est-à-dire le dépôt d'un polymère thermo-réticulant comme le PDMS sur le moule (Figure 38). La fabrication d'un moule par photolithographie suivie de la réplication par moulage est désignée par le terme photolithographie douce <sup>134</sup>. Après réticulation par chauffage, la structure élastomérique constituée par le polymère va reproduire le négatif du moule et être séparé du moule pour avoir l'objet en PDMS.

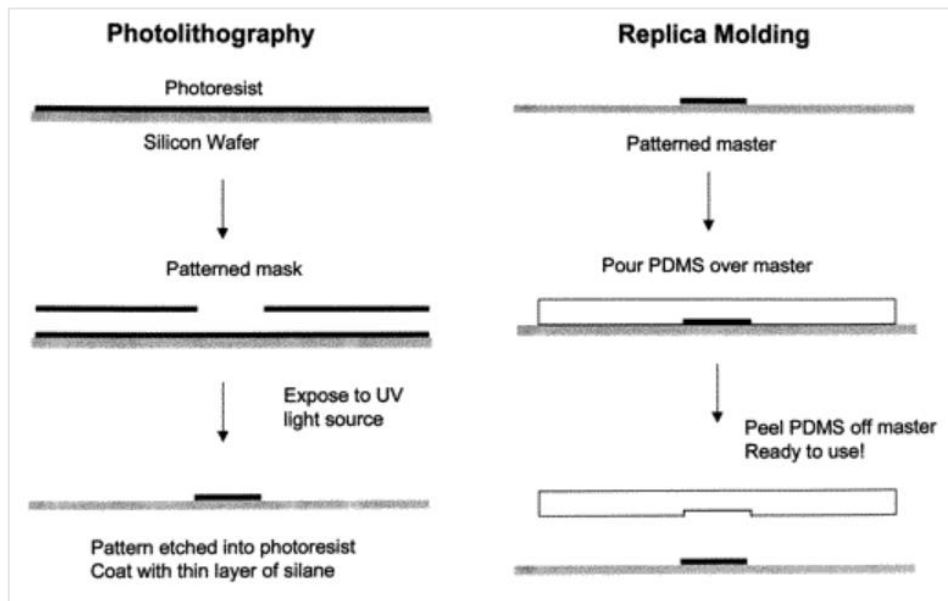


Figure 38 : Schéma de la fabrication d'un dispositif en PDMS en 2 étapes par photolithographie (à gauche) et moulage par réplication (à droite) <sup>122</sup>.

### 3.A.3. Traitements de surface existants

Pour les expériences sur des cellules uniques et d'autres échantillons à l'état de traces, les lames de verre ou les plaques de 384 puits sont le format préféré<sup>135,136</sup>. Mais ces plaques souffrent toujours de l'adsorption des peptides. Ainsi, un traitement de surface adapté est essentiel quel que soit le support utilisé, pour améliorer les propriétés de mouillage du PDMS et pour diminuer le phénomène d'adsorption non spécifique des protéines à la fois sur le PDMS et le PP.

Des traitements de surface dits « *low retention* » ou « *low binding* », utilisés en protéomique, visent à rendre la surface de contact avec les échantillons plus hydrophile. Sur les surfaces hydrophiles, la protéine doit traverser la couche d'eau d'hydratation de la surface pour interagir avec la surface sous-jacente. Ainsi, l'affinité d'interaction des protéines pour une surface hydrophile est plus faible que pour une surface hydrophobe, ce qui entraîne une moindre réorganisation de la conformation des protéines et donc une moindre adsorption<sup>104</sup>.

De nombreuses approches ont été élaborées pour rendre les surfaces plus hydrophiles<sup>137,131</sup>. Ces approches de modification des surfaces peuvent être catégorisées en deux types : des approches physiques comme la physisorption de polymères ou des approches chimiques comme l'activation de surface ou une combinaison des deux avec les brosses de polymères<sup>138</sup>. L'objectif de ces approches est de créer une surface de polymères ou de protéines tampons entre la surface du support et les échantillons d'intérêt.

Un exemple de polymère utilisé pour ces approches est le polyéthylène glycol (PEG), un matériau connu pour empêcher l'adsorption non spécifique des protéines ainsi que pour sa biocompatibilité et sa faible toxicité<sup>138,139,140</sup>. En outre, de nombreux polymères hydrophiles synthétiques ou naturels sont utilisés pour le revêtement statique et dynamique, tels que le polyacrylamide, le polyalcool vinylique (PVA), l'hydroxy-éthyl cellulose (HEC), le poly N-hydroxy-éthyl acrylamide (PHEA), l'hydroxy-propyl méthyl cellulose (HPMC), le poly(2-hydroxyéthyl méthacrylate) (pHEMA), le polyvinylpyrrolidone (PVP), le polyacide acrylique (PAA), le dextrane, l'acide hyaluronique et le poly(2-méthacryloyloxyéthyl-phosphoryl choline) (PMPC) ont été utilisés pour des modifications physiques ou covalentes de surfaces<sup>138,141,142</sup>.

#### (a) L'activation de surface

L'activation de surface par plasma à oxygène à radiofréquence (RF) est une méthode simple et très répandue, mais dont les effets sont temporaires sur certaines surfaces (environ 30 minutes sur le PDMS). Le plasma à oxygène a été largement utilisé pour modifier les surfaces (verre, polydiméthylsiloxane, polystyrène) depuis plusieurs décennies<sup>143,144,129</sup>. Une exposition variable en puissance et en temps a été étudiée plusieurs fois dans la littérature. La récupération hydrophobe des surfaces modifiées a été suivie en fonction du temps<sup>129</sup>.

Une analyse des surfaces pendant l'activation a montré que le plasma d'oxygène entraîne la formation d'une couche superficielle oxydée lisse (<10 nm) d'une épaisseur de 130-160 nm. La couche oxydée contient un mélange du polymère d'origine et de silicium lié à quelques atomes d'oxygène<sup>145</sup>. Ces études indiquent que le traitement par plasma à oxygène se propage profondément sous la surface du polymère (environ plusieurs centaines de nanomètres) et provoque des modifications chimiques dans la région proche de la surface (dans cette étude celle du PDMS) : des groupes fonctionnels polaires contenant des groupes

hydroxyles (groupes Si-OH) sont formés à sa surface, ce qui conduit à une surface hydrophile (Figure 39)<sup>129</sup>.

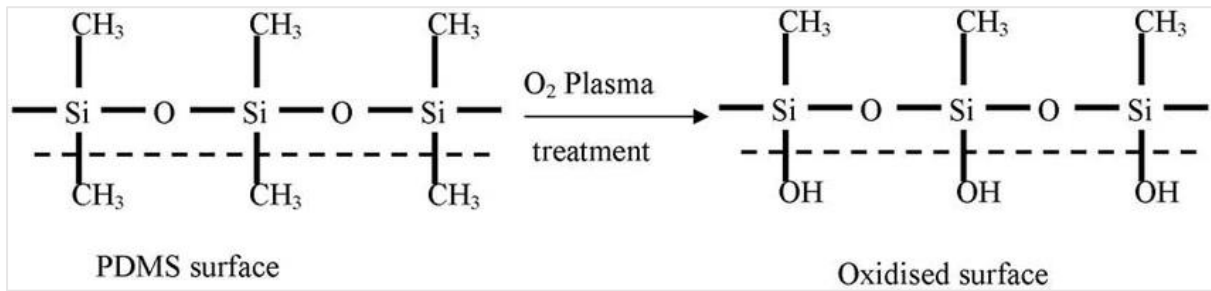


Figure 39 : Traitement au plasma à oxygène de la surface du PDMS qui devient oxydée.

Tan *et al.* rapportent un protocole simple et facile sur une surface de PDMS combinant un traitement prolongé au plasma d'oxygène et un stockage approprié pour produire des dispositifs PDMS hydrophiles. Les résultats montrent qu'avec une puissance de plasma de 70 W, un traitement prolongé de plus de 5 minutes permet à la surface du PDMS de rester hydrophile pendant plus de 6 heures<sup>146</sup>.

Au contraire, une autre étude montre que les caractéristiques de surface changent progressivement au cours du vieillissement et que la surface retrouve son hydrophobie après quelques heures<sup>129</sup>. La similitude de la cinétique de récupération de l'hydrophobie des échantillons vieillis à l'air sec, à l'argon sec et sous vide a montré que la récupération de l'hydrophobie n'est pas due à une contamination par adsorption à partir de l'atmosphère mais à la migration d'espèces de PDMS de faible masse molaire n'ayant pas réagi vers la surface<sup>147</sup>. Une analyse en microscopie électronique à balayage a également montré la présence de fissures en surface dans les spécimens fortement oxydés<sup>145</sup>.

Par conséquent, l'hydrophilie apportée par l'activation des surfaces par plasma à oxygène reste seulement temporaire<sup>131</sup>.

### (b) La physisorption

Les traitements de physisorption consistent en l'adsorption de molécules par interactions faibles hydrophobe ou électrostatique. Ils sont fréquemment utilisés pour les applications protéomiques. Le revêtement par physisorption est produit en laissant des monomères ou polymères (HEMA, PEG, PVP) pendant quelques minutes ou plusieurs heures au contact du support à traiter (Figure 40)<sup>131</sup>. Ces supports sont ensuite rincés abondamment avec de l'eau pour éliminer les résidus chimiques non attachés<sup>131</sup>.

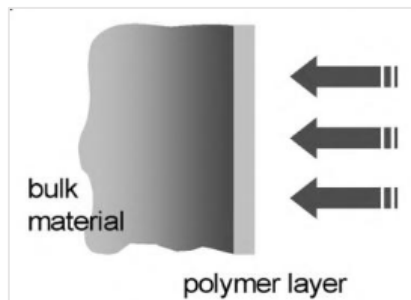


Figure 40 : Schéma du revêtement d'une surface par une fine couche de polymères pour mieux contrôler les interactions de la surface avec son environnement<sup>148</sup>.



## INTRODUCTION : Les avancées en protéomique de la cellule unique

Les surfaces de contacts sont souvent traitées par saturation avec des protéines de lait ou de l'albumine de sérum bovin (BSA) pour les tests ELISA ou pour les Western blots par exemple. En effet un ELISA est une technique immunochimique principalement utilisée pour détecter la présence d'une protéine à l'aide d'anticorps spécifiques. La première étape consiste à fixer l'antigène d'intérêt ou l'anticorps de capture à la surface de la plaque de test. Afin d'éviter les phénomènes d'adsorption, la BSA est utilisée comme agent bloquant <sup>149,150</sup>. De même saturer les membranes des *Western Blots* en protéines connues (caséine) permet de limiter l'adsorption des protéines d'intérêt. Dans un autre exemple, Bratcher et Gaggari ont montré que le traitement des tubes avec la BSA diminuait de manière significative la quantité de protéine d'intérêt collée au plastique (polypropylène), fournissant ainsi une méthode peu coûteuse et efficace pour prévenir l'adsorption <sup>151</sup>.

D'autres méthodes suggèrent un traitement avec des oligopeptides d'alanines et de lysines pour minimiser l'adsorption d'un antigène étudié dans des matrices sur les matériaux à base de polydiméthylsiloxane (PDMS), et indiquent que ce traitement donne de meilleurs résultats que le traitement BSA <sup>130</sup>. Dans une étude sur les infections, pour éviter la croissance bactérienne sur les cathéters en PDMS, Mohan *et al.* ont décrit une approche en conditions aqueuses en fixant de l'acide hyaluronique carboxylé et du fucose sulfaté sur des implants en PDMS. Ils ont montré que cette approche offrait une stabilité à long terme, une absence d'adsorption des protéines et une résistance accrue à la croissance des agents pathogènes <sup>152</sup>.

Cependant, la saturation du contenant par des produits liés par des liaisons non covalentes n'est pas une option pour la protéomique car les molécules libérées de façon régulière satureraient le signal en spectrométrie de masse et masqueraient le signal des molécules d'intérêt moins abondantes. En effet, la physisorption repose sur des interactions faibles entre la surface formée et la couche superficielle du matériau. Cela entraîne des instabilités thermiques dans le matériau et la densité de greffage est donc moins bien contrôlée <sup>131,153</sup>.

### (c) Les brosses de polymères

Les brosses de polymères sont une stratégie alternative pour fabriquer des surfaces hydrophiles plus robustes à long terme et forment une combinaison des méthodes citées précédemment. Le greffage covalent de chaînes de polymères à la surface peut surmonter les inconvénients des deux précédentes étapes <sup>153</sup>. Il peut être réalisé soit par une méthode de « *grafting to* », où une chaîne polymère fonctionnalisée peut réagir avec un groupe réactif sur la surface polymère non modifiée, soit par une méthode de « *grafting from* », où les monomères fonctionnalisés polymérisent à partir d'un initiateur présenté à la surface du polymère non modifié (Figure 41) <sup>154,155</sup>.

La méthodologie « *grafting to* » ou « greffage sur » est la méthode la plus simple pour former des brosses de polymère. Des polymères fonctionnalisés sont attachés à des groupes réactifs sur la surface d'un matériau par liaison chimique ou par physisorption. Une première étape consiste à générer les groupes réactifs ou ancres en utilisant par exemple l'activation de la surface par plasma à oxygène. Ensuite, la surface du matériau est traitée avec des polymères (HEMA, PEG, PVP) pour obtenir une brosse de polymères en surface <sup>131</sup>. Bien que la brosse présente une stabilité thermique et solvolytique, elle présente une faible densité de greffage et une faible épaisseur de film à la surface. Cette observation est due à l'incapacité

des grandes chaînes de polymères à diffuser vers les groupes réactifs à la surface à cause de l'encombrement stérique des chaînes greffées <sup>153</sup>.

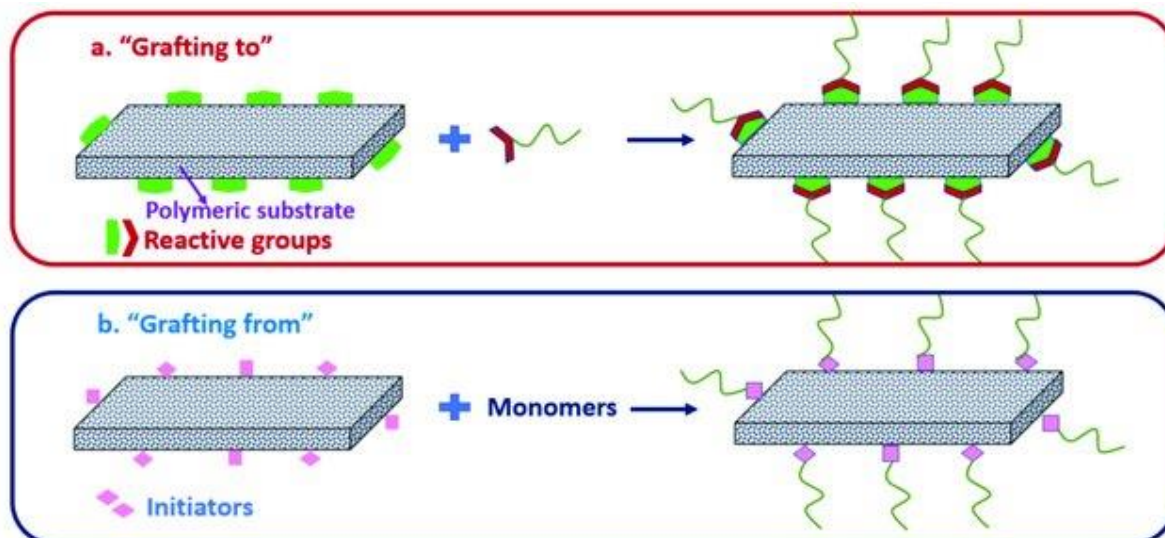


Figure 41 : Stratégies de création des brosses de polymères (a) méthodologie « *grafting to* » (b) méthodologie « *grafting from* » <sup>156</sup>.

La méthodologie par greffage ou « *grafting from* » est une polymérisation initiée à la surface (SIP pour « *surface-initiated polymerization* ») <sup>131</sup>. C'est la méthodologie la plus robuste pour former des brosses de polymère. La formation de ces brosses par SIP commence par l'ancrage d'une espèce initiatrice sur la surface du matériau à traiter. Les monomères ajoutés aux chaînes de polymères en développement ne subiront pas autant l'effet d'encombrement stérique que des polymères. Cela facilite la production d'un revêtement plus épais et plus dense <sup>131</sup>. La longueur de la brosse étant directement proportionnelle au poids moléculaire du polymère, le contrôle de la polymérisation permet d'obtenir une épaisseur de brosse uniforme et ajustable <sup>153</sup>. Dans cette méthode, la surface doit être modifiée pour générer les groupements chimiques initiateurs adaptés à la synthèse de la brosse polymère. En outre, selon la méthode de polymérisation (activation au plasma à oxygène, etc.), l'initiateur peut être de différents types <sup>131</sup>. En variant le substrat (or, silice, nanoparticules, etc.), la technique de dépôt de l'initiateur et la voie de synthèse du polymère diffèrent et offrent des possibilités pratiquement illimitées pour la formation de brosses de polymères <sup>153</sup>.

Un exemple de l'utilisation de cette technique est présenté dans les travaux de Carneiro *et al.* qui ont combiné une oxydation de la surface par activation au plasma à oxygène avec une liaison chimique de l'alcool polyvinylique (PVA pour « *polyvinyl alcohol* ») afin d'obtenir une surface de PDMS hydrophile et stable à long terme <sup>157</sup>.

### 3.B Les débuts de la protéomique en cellule unique

La préparation des échantillons pour les analyses en protéomique de la cellule unique est au cœur des publications depuis une dizaine d'années. Ces approches sont toutes confrontées à plusieurs défis tels que la miniaturisation, la sensibilité et le débit d'analyse.

Une cellule unique étudiée ne doit pas être traitée dans un même volume de 50 ou 100  $\mu$ l utilisés pour analyser un million de cellules (voir partie « un problème de taille »). Cela signifie

## INTRODUCTION : Les avancées en protéomique de la cellule unique

que le protocole de traitement de l'échantillon en protéomique classique n'est absolument pas adapté. Il est donc essentiel de l'adapter en le miniaturisant.

Contrairement à la génomique ou à la transcriptomique, il n'existe pas encore de méthode universelle "standard" de préparation des échantillons pour la protéomique en cellule unique<sup>42</sup>. Deux grands types d'approches ont déjà été développés pour l'analyse protéomique sur cellule unique.

La première concerne les analyses ciblées et est basée sur l'utilisation d'anticorps à l'instar des puces à protéines qui découlent de la transcriptomique. Dans ce cas sont utilisés des marqueurs spécifiques reposant sur des interactions en immuno-affinité, et des sondes de détection comme la fluorescence. Les méthodes utilisant des fluorophores ou des anticorps se limitent à un petit nombre de protéines cibles et souffrent de limitations dues à la disponibilité d'anticorps spécifiques ou de faible précision quant à la quantification des protéines<sup>158</sup>. L'intérêt de ces méthodes réside dans leur capacité à faire du multiplexage.

La seconde basée sur la spectrométrie de masse cherche à réaliser un phénotype moléculaire des cellules sans *a priori*. Il est intéressant de mieux caractériser l'hétérogénéité cellulaire au sein d'une même lignée et cela nécessite des analyses structurales plus détaillées. La protéomique *bottom-up* est une stratégie de choix. Mais elle se heurte à de nombreuses limitations, dues d'une part à la très faible quantité de matériel disponible pour l'analyse, les faibles rendements induits par le traitement et la manipulation des échantillons avec une absence d'amplification du signal, un débit d'analyse insuffisant, la sensibilité de détection limitée de la MS et de la MS/MS, et la complexité intrinsèque du mélange.

### 3.B.1. Approches basées sur les anticorps

Les premières se basent sur l'utilisation d'anticorps ou de protéines fluorescentes comme le CyTOF, le *Western Blot* en cellule unique, le siMoA (pour « *single molecule array technology* ») ou encore le CITE-Seq, l'Abseq (combinant des anticorps marqués et du séquençage d'ARN en cellule unique)<sup>159</sup>. Ces méthodes ont mesuré jusqu'à quelques dizaines de protéines par cellule. Elles se limitent à la détection et au dosage de protéines connues.

#### (a) CyTOF

La cytométrie de masse, ou CyTOF (cytométrie par spectrométrie de masse à temps de vol), est devenue un outil puissant dans le domaine de la protéomique unicellulaire. Inspiré de l'ICP-MS, le CyTOF a été mis au point dans le laboratoire de Scott Tanner à l'université de Toronto en 2008<sup>160</sup>. La cytométrie de masse combine les principes de la spectrométrie de masse et de la cytométrie de flux. Le CyTOF peut évaluer simultanément le contenu de l'ADN et des protéines et a la capacité de mesurer 40 à 100 paramètres dans chaque cellule<sup>161</sup>. Dans cette méthode, les protéines cellulaires sont marquées par des anticorps conjugués avec des isotopes métalliques lourds non présents naturellement dans les échantillons biologiques étudiés. La suspension de cellules marquées est ensuite nébulisée en aérosol de gouttelettes contenant des cellules isolées qui passent dans un plasma et sont ionisées pour analyse par spectrométrie de masse à temps de vol (TOF pour « *time of flight* ») (Figure 42)<sup>162,163</sup>. Chaque isotope métallique unique associé à un anticorps marqué peut être identifié. La quantité d'anticorps liée à chaque cellule est détectée dans le spectromètre de masse et correspond aux niveaux d'expression de l'antigène cible<sup>163,161</sup>.

## INTRODUCTION : Les avancées en protéomique de la cellule unique

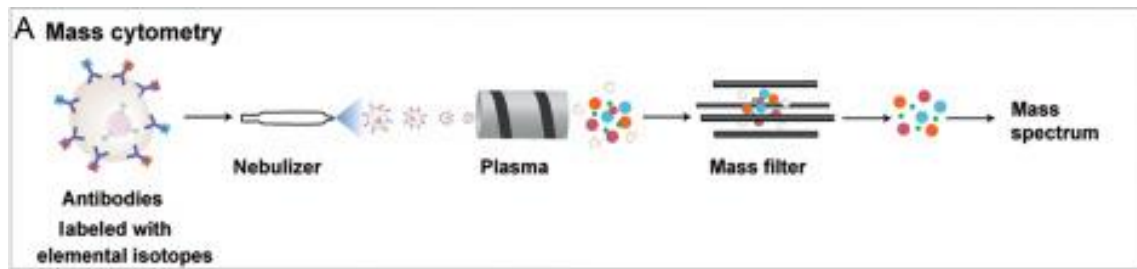


Figure 42 : Principe de la cytométrie de masse. Les cellules sont marquées par des anticorps conjugués à des isotopes métalliques de masses différentes. Les cellules sont ensuite séparées en gouttelettes monocellulaires et nébulisées dans un plasma puis analysées par spectrométrie de masse <sup>164</sup>.

Une équipe comprenant des chercheurs de l'Université d'Ottawa et de l'Université d'Oxford a utilisé la cytométrie de masse unicellulaire (CyTOF) pour capturer les variations relatives au statut cellulaire au cours de l'hématopoïèse, et a suivi l'évolution de l'expression des facteurs de transcription d'une cellule dans une lignée à différents temps. Ils ont mesuré 27 protéines simultanément dans des cellules individuelles. Une équipe de l'université de Zurich, en collaboration avec des collègues d'autres institutions, a réalisé un autre projet CyTOF unicellulaire pour établir le profil des cellules tumorales et immunitaires de 144 échantillons de tumeurs mammaires humaines <sup>165</sup>.

### (b) Western Blot en cellule unique

Pour mesurer la variation des protéines d'une cellule à l'autre, une approche permettant de réaliser 1000 *Western Blots* monocellulaires simultanés (sc Westerns) a été développée. Brièvement, une lame microscopique supportant un gel de polyacrylamide photo-actif permet la réalisation de Western blots. Les étapes dans le gel sont l'isolement de cellules individuelles dans des micro-puits, une lyse in situ, une électrophorèse sur gel, un transfert photo-initié pour immobiliser les protéines et les marquer avec un anticorps pour une détection par immunofluorescence (Figure 43).

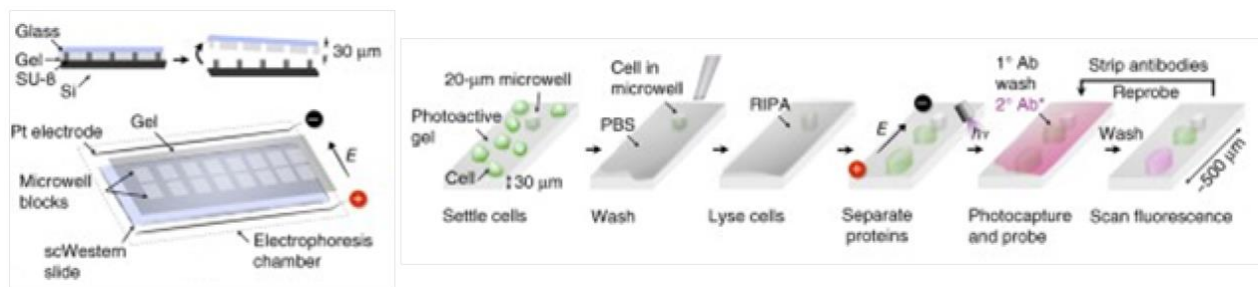


Figure 43 : Principe du Western Blot en cellule unique a) schéma du gel c) schéma du protocole.

Hughes *et al.* ont appliqué cette méthode pour suivre la différenciation unicellulaire des cellules souches neurales de rat et les réponses à la stimulation mitogène. En indiquant la masse moléculaire (par électrophorèse) ainsi que la détection par immunofluorescence (par liaison d'anticorps), le scWestern a pu identifier deux isoformes putatives d'une protéine et suggérer que l'une d'entre elles (NESt $\alpha$ ) reflète mieux la sortie des cellules souches neuronales de leur état immature <sup>166</sup>.

### (c) siMOA

Le siMOA (pour « *single molecule array* ») est un test basé sur un réseau de molécules uniques avec une limite de détection très faible. Ces tests combinent le format des tests ELISA (décrits dans la partie « physisorption ») avec un réseau de micro-puits de la taille d'un femtolitre pour isoler des billes magnétiques contenant un ou plusieurs immunocomplexes et leurs marqueurs enzymatiques associés. En immobilisant des anticorps à la surface des billes paramagnétiques puis en les mettant au contact des molécules d'intérêt, d'un anticorps secondaire biotinylé et d'une enzyme marquée à la streptavidine, un immunocomplexe se forme. Les puits sont ensuite scellés en présence du RDG (résorufine- $\beta$ -D-galactopyranoside) non fluorescent, un substrat enzymatique qui est converti en un produit fluorescent (résorufine) lorsqu'il est hydrolysé dans la réaction enzymatique. Le format multi-micro-puits permet un multiplexage à haut débit des analyses (jusqu'à 50000 micro-puits en une analyse). Une quantification de l'intensité de la fluorescence dans les puits actifs (où la réaction a eu lieu) permet d'obtenir la concentration en molécules d'intérêt.

Ces essais ont été appliqués pour déterminer les niveaux de cytokines dans les échantillons de sérum humain sain. Grâce à l'ultra-sensibilité de ces tests, les cytokines GM-CSF, TNF- $\alpha$ , IFN- $\gamma$ , IL-1 $\beta$ , IL-5, IL-6, IL-7 et IL-10 ont pu être détectées dans tous ou la plupart des échantillons humains sains, sachant que la plupart ne sont pas détectables à l'aide des kits ELISA classiques <sup>167</sup>.

### (d) CITE-Seq

Le CITE-Seq est une méthode de séquençage multimodale qui permet de détecter à la fois les transcriptomes et les épitopes cellulaires des cellules d'intérêt. Des anticorps sont marqués par des oligonucléotides et utilisés pour intégrer les mesures des protéines à celles des ARNm sur des cellules uniques. Ces oligonucléotides seront détectés par les amorces pour une amplification par PCR (voir la partie « transcriptomique en cellule unique »). Ces anticorps sont incubés afin de marquer des épitopes d'intérêt. Les oligonucléotides de marquage et l'ARNm sont séquencés par Drop-Seq, une technique de transcriptomique en cellule unique <sup>168</sup>. Le CITE-Seq est donc une méthode qui combine la détection de marqueurs protéiques hautement multiplexés avec le profilage du transcriptome pour des milliers de cellules uniques. L'analyse multimodale des données ainsi faite permet d'obtenir une caractérisation plus complète des phénotypes cellulaires que les mesures du transcriptome seules <sup>169</sup>.

### (e) Abseq

Abseq est une méthode utilisant des anticorps spécifiques pour détecter des épitopes d'intérêt. Ces anticorps sont marqués avec des étiquettes de séquence ADN qui peuvent être lues par un code-barres microfluidique ou séquençage de l'ADN. Shahi *et al.* ont caractérisé des protéines de surface de différents types cellulaires à l'échelle de la cellule unique. Ici aussi la technique bénéficie de l'amplification du signal car les marqueurs ADN spécifiques des protéines de faible concentration pourront être amplifiés pour les rendre détectables par séquençage. Cette méthode permet également un multiplexage des échantillons illimité, ce qui en fait une technologie de détection et quantification des protéines dans les cellules uniques à très haut débit <sup>159</sup>.

### 3.B.2. Approches basées sur la spectrométrie de masse

Des approches non ciblées ou *shotgun* tentant l'identification des protéomes totaux de cellules uniques ont été développées en se basant sur l'utilisation de la spectrométrie de masse. Ces approches non ciblées ne bénéficient d'aucune technique de marquage moléculaire ou d'amplification. Il est ici essentiel de travailler sur le rendement de transfert des molécules lors de la manipulation des échantillons. Au lieu d'utiliser des méthodes conventionnelles de préparation des échantillons qui ont une surface de contact très grande provoquant de l'adsorption, de larges volumes rendant les cinétiques de réaction inefficaces et de multiples transferts d'échantillons qui multiplient les pertes, un traitement dans des très faibles volumes est nécessaire.

#### (a) iPAD-1

L'iPAD-1 (pour « *integrated proteome analysis device* ») est un système de micro-valves permettant l'aspiration des cellules une par une dans un capillaire puis le traitement de la cellule directement dans ce capillaire.

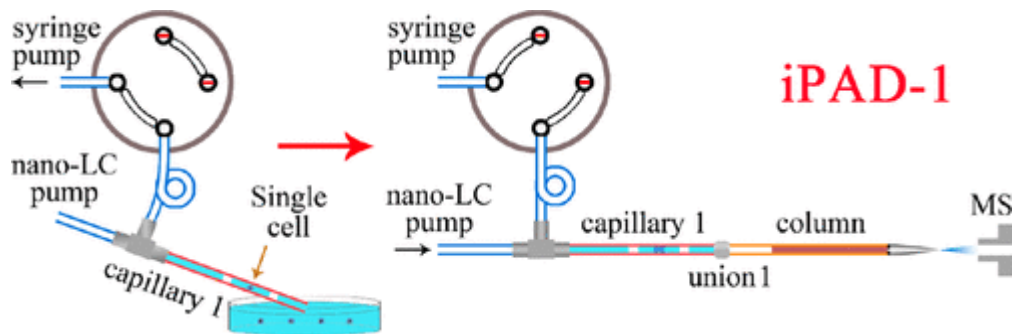


Figure 44 : Schéma du protocole dans le dispositif iPAD-1.

Le dispositif iPAD-1 permet de profiler le contenu en protéines dans une cellule en 1 heure. La cellule unique sélectionnée est directement aspirée dans un capillaire de 22  $\mu\text{m}$  de diamètre interne. Puis la lyse cellulaire et la digestion des protéines sont simultanément accomplies dans le capillaire dans un volume de 2 nL, ce qui permet d'éviter la perte de protéines par une dilution excessive. Une digestion accélérée est effectuée par l'utilisation d'une température élevée et d'ultrasons (Figure 44). Le traitement cellulaire dure 30 min. Un gradient de 30 min est appliqué pour l'éluion des peptides dans une colonne de 22  $\mu\text{m}$  de diamètre interne placée au bout du capillaire. Les analyses de 10 cellules uniques ont permis l'identification de 328 protéines. Une hétérogénéité cellulaire importante dans le profil d'expression des protéines avec un coefficient  $R^2$  entre 0.25 et 0.9 entre deux cellules HeLa uniques a été observée expliquée selon les auteurs par les différences entre les étapes du cycle cellulaire des cellules analysées<sup>170</sup>.

Ce système bien qu'il rassemble toutes les étapes de préparation d'échantillons et une interface avec la MS présente le désavantage de ne pas permettre le traitement en parallèle de plusieurs échantillons.

## INTRODUCTION : Les avancées en protéomique de la cellule unique

### (b) « Oil Air Droplets » (OAD)

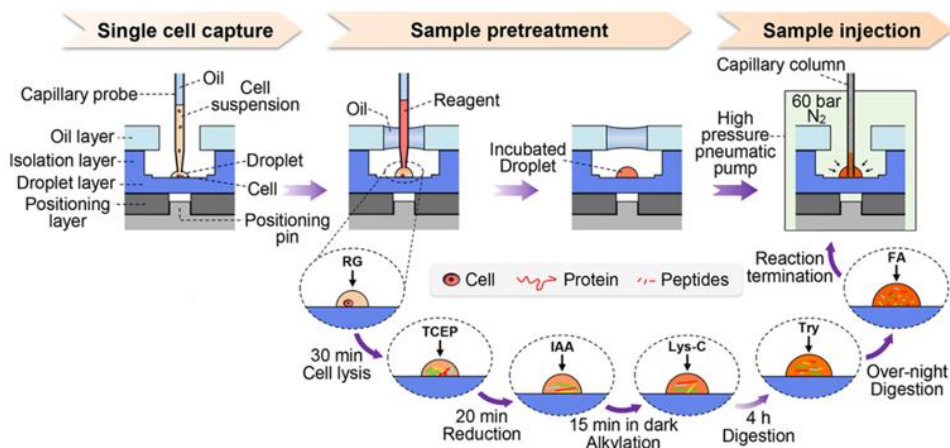


Figure 45 : Schéma du protocole de protéomique en cellule unique dans le dispositif OAD <sup>171</sup>.

Le dispositif OAD (pour « *oil-air droplets* ») publié en 2018, est une puce à l'échelle du nanolitre qui permet le traitement et l'injection de cellules uniques pour leur analyse en protéomique *shotgun*. En utilisant une chambre de micro-réaction et une aspiration sous vide pour la récupération des échantillons, ce système permet une perte minimale d'échantillon et une grande efficacité d'injection augmentant ainsi la sensibilité analytique (Figure 45). La chambre a été fabriquée en polypropylène ou en verre. La couche d'huile sert à prévenir toute évaporation pendant les réactions.

Il a été appliqué par Li *et al.* sur différentes quantités de cellules HeLa et pour 100 et 1 cellule, environ 1360 et 51 protéines ont été identifiées respectivement. Ce système a également été testé sur de plus grosses cellules, identifiant 355 protéines sur un seul ovocyte de souris, prouvant ses avantages dans l'augmentation de la couverture protéique, de l'identification des protéines hydrophobes et de l'efficacité enzymatique dans leur chambre de réaction par rapport aux réactions en tube <sup>171</sup>.

Cependant, une limitation majeure est la présence d'autres familles de molécules. Il faut être très attentif à la signature lipidique lors de l'utilisation d'huiles en spectrométrie de masse. Les lipides pouvant être séparés par des colonnes de phase inverse C18 ou C8 comme utilisées en protéomique, leur signature peut encrasser la colonne et empêcher une bonne séparation peptidique <sup>172,173,174</sup>.

### (c) SCOPE-MS

En 2018, l'équipe de Nikolai Slavov de la « *Northeastern University* » a développé la stratégie SCoPE-MS (pour « *single cell proteomics by mass spectrometry* »), qui a permis l'analyse protéomique sur cellule unique dans une plaque 96 multi-puits en isolant les cellules par FACS et en les marquant au TMT. SCoPE-MS réserve un des marqueurs TMT pour un échantillon de cellules appelées « *carriers* » qui contient entre 100 et 1000 cellules selon les exemples et qui est utilisé pour saturer les surfaces avec une grande quantité de protéines afin de limiter la perte des protéines et peptides provenant des cellules uniques (Figure 46) <sup>124,158</sup>.

Un second protocole, SCoPE2, augmente considérablement la précision quantitative et le débit en réduisant le coût et le temps de manipulation grâce à une préparation automatisée

## INTRODUCTION : Les avancées en protéomique de la cellule unique

et miniaturisée des échantillons avec un protocole en plaques 384 puits appelé mPOP (Figure 46)<sup>135</sup>. En analyse de cellules uniques, les niveaux de protéines étant réduits, le nombre de peptides séquencés et identifiés de façon robuste diminue. Un traitement bio-informatique a été intégré au protocole de SCoPE2, le DART-ID (pour « *data-driven alignment of retention times for identification* »)<sup>175</sup>, qui vient en complément de l'analyse MaxQuant. DART-ID met en application des modèles bayésiens pour l'alignement global des temps de rétention (RT) et pour l'incorporation des estimations des RT en vue d'améliorer la robustesse des correspondances peptide-spectre. DART-ID augmente le nombre de données de 30 à 50% à 1% FDR et réduit le nombre de données manquantes<sup>175</sup>. Ces avancées permettent d'analyser l'émergence de l'hétérogénéité cellulaire lors de la différenciation de monocytes en macrophages. SCoPE2 quantifie plus de 3042 protéines dans 1490 monocytes et macrophages uniques, et les protéines quantifiées permettent de discerner les différents types cellulaires<sup>135</sup>. Pour obtenir ce résultat 62 préparations ont été réalisées, chaque extrait analysé contient 10 cellules individuelles marquées différemment, 56 cellules références et 200 cellules du mélange « carrier ». Chaque extrait ainsi préparé a permis de quantifier 1000 protéines et 2500 peptides en moyenne, et 6 extraits n'ont pas donné de résultats.

En 2022, le protocole de traitement de données mPOP a été automatisé en nano-protéomique « nPOP » (Figure 46). nPOP utilise un tri cellulaire et une distribution piézo-acoustique grâce à l'instrument CellenOne (Scienion) pour isoler des cellules individuelles dans des volumes de 300 picolitres et effectue toutes les étapes de préparation ultérieures en petites gouttelettes sur une lame hydrophobe. Cela permet une préparation parallèle des échantillons, y compris la lyse, la digestion et le marquage des cellules individuelles dans des volumes inférieurs à 20 nl. L'analyse protéique unicellulaire utilisant nPOP a permis de classer les cellules par type cellulaire et par phase de cycle cellulaire<sup>176</sup>. Deux modes d'analyse ont été ici utilisés. Le premier utilise la stratégie en mode DDA avec TMT et 105 cellules carriers et permet d'identifier un total de 2848 protéines sur 1543 préparations avec en moyenne 338 par préparation, sachant que seules 880 (30%) protéines sont retrouvées dans plus de 50% des préparations. Ce chiffre tombe à 645 (23%) si seules les protéines retrouvées dans 75% des préparations nous intéressent. Le second utilise un marquage différent et travaille sur des sets de 3 cellules en mode DIA, et permet d'identifier un total de 795 protéines sur 106 préparations en se reposant sur une base de données spectrale préparée à partir d'un grand nombre de cellules du même modèle, avec en moyenne 386 protéines identifiées par préparation.

Enfin, l'utilisation du plexDIA permet d'augmenter encore le *multiplexing* en augmentant le débit sans réduire la couverture du protéome ou la précision quantitative. En utilisant des marqueurs de masse non isobares, plexDIA permet de quantifier plus de protéines. Appliquée à des cellules humaines uniques, plexDIA a quantifié environ 1 000 protéines par cellule et a atteint 98 % de complétude des données<sup>177</sup>.

La technologie émergente de la protéomique unicellulaire a été portée par l'équipe de Nikolai Slavov qui organise tous les ans, depuis 2017, le congrès SCP (« *single cell proteomics* ») à Boston réunissant de plus en plus de monde chaque année.



INTRODUCTION : Les avancées en protéomique de la cellule unique

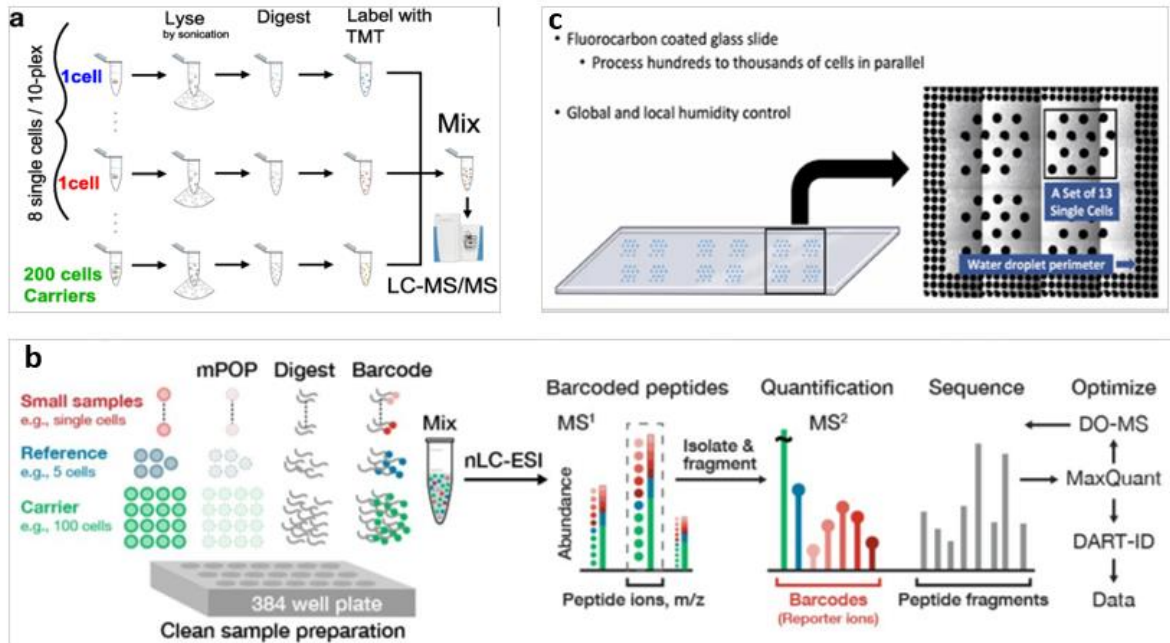


Figure 46 : Schéma des méthodologies Scope MS (a), Scope 2 (b) et nPOP (c).

(d) NanoPOTs

Les nanoPOTs (pour « *nanodroplet processing in one pot for trace samples* ») ont été développés par les équipes de Ryan Kelly et Ying Zhu au PNNL (*Pacific Northwest National Laboratory*) en 2018, pour une analyse protéomique des échantillons de trace dans des volumes minimaux <sup>136</sup>. Il s'agit d'une approche non biaisée utilisant la microdissection par capture laser (LCM pour « *laser capture microdissection* ») couplée à un système de préparation d'échantillons à l'échelle du nanolitre pour établir le profil des protéines dans des sections de tissus avec une haute résolution spatiale et d'identification.

Le traitement est automatisé et la surface de contact des protéines au contenant ou support est réduite grâce au faible volume d'une goutte, réacteur de 200nL, ce qui limite aussi la dilution. Les cellules individuelles sont triées par FACS et déposées sur une puce présentant des ancrs hydrophiles. De plus, tout le processus se fait dans la même goutte évitant la perte d'échantillon liée aux transferts (Figure 47). Pour éviter l'évaporation, le système était maintenu à 95% d'humidité en présence de tensio-actif clivables en conditions acides (Rapigest). Associés à une nano LC-MS ultrasensible utilisant des colonnes de 30µm de diamètre interne, les nanoPOTs permettent l'identification de 967 à 2167 protéines reproductibles de ~10 à ~140 cellules HeLa , respectivement <sup>136</sup>.

Par la suite, l'utilisation des marqueurs de masse en tandem (TMT pour « *tandem mass tags* ») permet un multiplexage des échantillons et une quantification reproductible. Entre 800 et 1000 protéines ont pu être identifiées grâce à cette méthodologie sur des cellules HeLa uniques.

INTRODUCTION : Les avancées en protéomique de la cellule unique

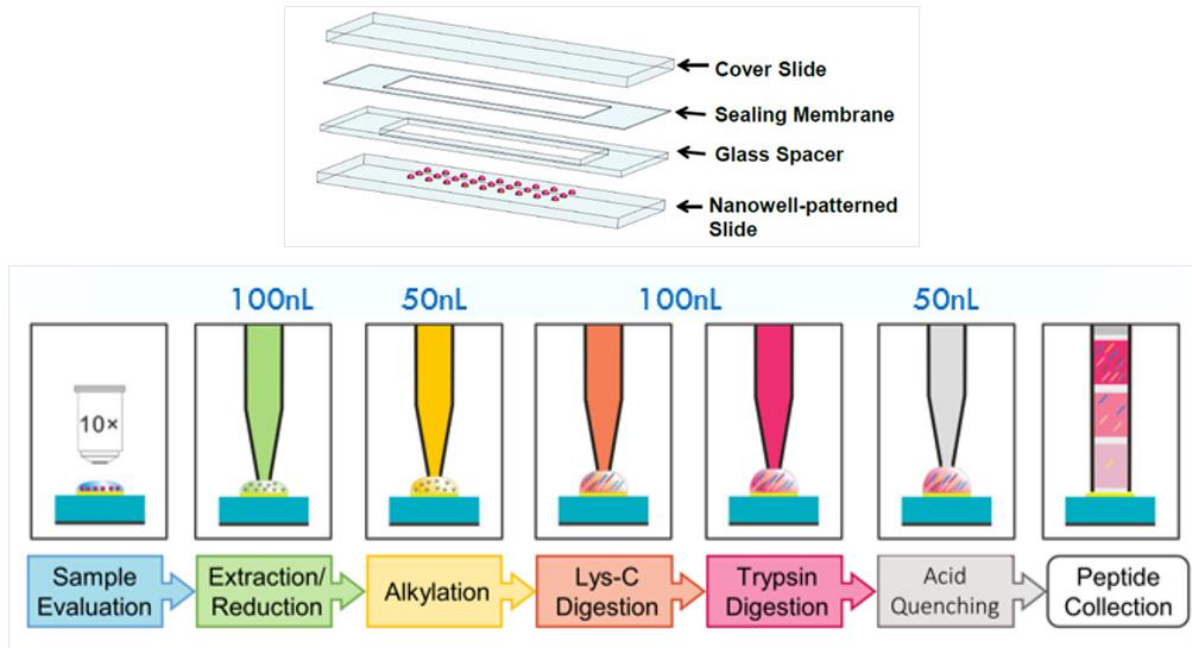


Figure 47 : Première génération de puces NanoPOTS (en haut). Schéma du protocole de préparation dans les NanoPOTS (en bas).

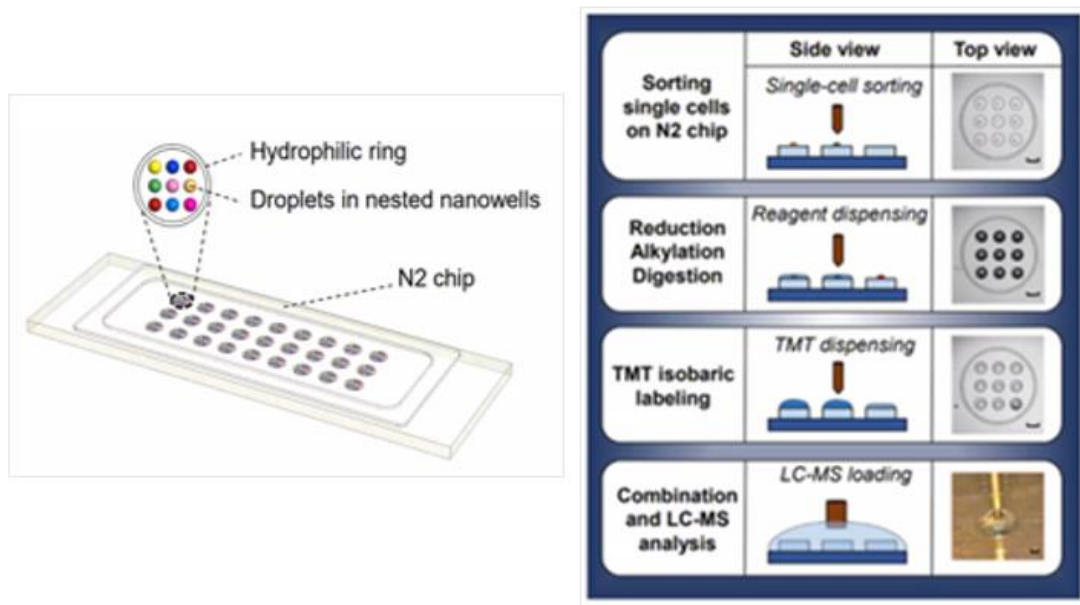


Figure 48 : Schéma de la puce N2 (à gauche) et du nouveau protocole amélioré (à droite).

Cette méthodologie a été améliorée avec une deuxième génération de puce nommée N2. Cette amélioration est basée sur un traitement automatisé dans un volume de réaction total de moins de 30nL et une récupération des échantillons marqués en TMT groupé sans mélange préalable directement dans le capillaire d'injection, permettant une parallélisation du traitement des échantillons (Figure 48). La présence de n-dodécyl- $\beta$ -d-maltoside (DDM) assure la lyse et la solubilisation des protéines qui sont ensuite réduites puis alkylées avant protéolyse. Les peptides sont alors marqués avec du TMT. Il faut noter que les auteurs utilisent un amplificateur de signal, à partir de préparations de 10 millions de cellules marquées au TMT : dans chaque goutte d'échantillon 10 ng de ce mélange protéolytique et

## INTRODUCTION : Les avancées en protéomique de la cellule unique

0.5 ng d'un peptide de référence sont ajoutés. L'analyse de 108 cellules individuelles provenant de trois lignées cellulaires différentes et mélangées avant analyse par groupe de 9 a permis de valider la présence de 1500 protéines quantifiées dans au moins une cellule et de détecter des marqueurs de protéines membranaires <sup>178</sup>.

### (e) ProteoCHIP

La proteoCHIP développée par l'équipe de Karl Metchler en 2021 représente une solution pour la préparation unicellulaire avec un débit et une sensibilité élevés. Le traitement est automatisé utilisant le système commercial CellenONE<sup>®</sup> pour l'isolement des cellules et la distribution de picolitres de réactifs (Figure 49). Il permet de réduire les volumes des échantillons à quelques nanolitres et évite la manipulation des échantillons. La particularité de ce système est qu'il évite l'évaporation en immergeant les échantillons dans une couche d'hexadécane, un hydrocarbure saturé de la famille des alcanes.

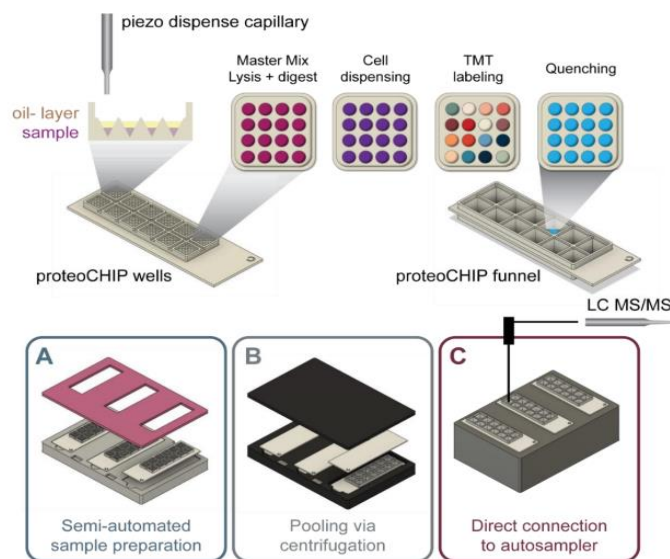


Figure 49 : Schéma de la proteoCHIP et du protocole appliqué. (a) 16 cellules uniques sont préparées dans le CellenOne<sup>®</sup>, (b) sont automatiquement mélangées par centrifugation et (c) sont directement mis à l'interface d'un échantillonneur automatique pour l'acquisition.

La conception de la proteoCHIP permet l'injection directe de cellules uniques non marquées via un passeur automatique standard, ce qui permet de récupérer 30 % de groupes protéiques en plus par rapport aux échantillons transférés dans des tubes LC couverts de PEG. Grâce à ce protocole, environ 2000 protéines dans 158 cellules uniques ont été identifiées et deux types cellulaires très similaires ont pu être séparées en fonction de leurs protéomes <sup>179</sup>. Dans cette stratégie les extraits sont complétés avec une quantité minimale de carrier (20 cellules), et l'ensemble des expériences est réalisé dans des puces en PTFE avec des nano puits de volume total 600nL.

Il est important de ne pas la confondre avec son ancêtre, la proteoCHIP créée en 2003 par Lee *et al.* <sup>180</sup>. Celle-ci permet une immobilisation efficace de protéines spécifiques sur des matrices solides et rend possible l'analyse à haut débit des interactions protéine-protéine.

### 3.B.3. Bilan des stratégies actuelles pour l'analyse protéomique de très petits nombres de cellules

Ces exemples ne forment pas une liste exhaustive mais donnent un aperçu des étapes clés du développement du domaine de la protéomique en cellule unique. Aujourd'hui, de nombreuses techniques de préparation d'échantillons pour traiter une faible quantité de cellules existent. L'objectif global est de collecter ces cellules sans perte, de minimiser la lyse des échantillons, de maximiser la cinétique de digestion des protéines en peptides détectables en spectrométrie de masse tout en limitant la perte d'échantillon <sup>181</sup>.

La préparation d'échantillons a été développée en réduisant les volumes de réaction et en rendant leur gestion automatique. Zhu *et al.* utilisent une plateforme robotique développée dans leur laboratoire pour la gestion des réactifs en très petits volumes <sup>136</sup>. Trois des équipes citées utilisent une solution commerciale comme l'instrument CellenONE®, un système de manipulation de liquides par distribution acoustique avec une collecte dédiée en aval pour l'isolement de cellules uniques et la réalisation de toutes les étapes de traitement dans un volume de l'ordre du nanolitres, ce qui limite les interactions de surface et donc les pertes de protéines dues à l'adsorption <sup>124,181,182,183</sup>. Mais ce genre d'instruments n'est pas disponible dans tous les laboratoires.

La préparation d'échantillons a également été développée pour l'amélioration de la récupération peptidique ou la diminution des pertes. Le protocole nPOP utilise l'ajout du DMSO pour aider à la lyse cellulaire et limiter l'évaporation <sup>176</sup>. Nie *et al.* ajoutent du n-dodécyl- $\beta$ -D-maltoside (DDM), un agent tensioactif non anionique, à une analyse de cellules HeLa uniques afin d'améliorer la récupération de peptides hydrophobes sans avoir d'impact négatif sur la récupération des peptides hydrophiles <sup>184</sup>. L'équipe de Karl Metchler a limité son protocole à deux étapes cruciales qui sont la lyse et la digestion en parlant de protocole « *less is more* » <sup>182</sup>.

L'amélioration de la séparation pour améliorer la couverture protéique s'est surtout basée sur l'utilisation de colonnes miniaturisées de RPLC <sup>183</sup>. En effet, les diamètres internes plus petits des colonnes permettent un débit plus faible et ainsi une amélioration de la sensibilité de la séparation peptidique. Ce débit plus faible permet la concentration du peptide élué et l'efficacité de l'ionisation <sup>124</sup>. Shen *et al.* ont détecté 190 fois plus de spectres de masse sur 100ng de digestat tryptique de levure en diminuant le diamètre interne (d.i.) de la colonne à 14.5  $\mu\text{m}$  au lieu des 74.5 $\mu\text{m}$  habituels <sup>185</sup>. L'équipe de Ryan Kelly a démontré une augmentation de 32% des identifications de peptides sur 10ng de peptides tryptiques en utilisant des colonnes de 30 $\mu\text{m}$  d.i. au lieu des 75 $\mu\text{m}$  d.i. <sup>186</sup>. Aussi, Cong *et al.* ont obtenu une augmentation de 41% en identifications protéiques en utilisant une colonne de 20 $\mu\text{m}$  d.i. au lieu de 30 $\mu\text{m}$  d.i. pour des cellules HeLa uniques préparées par nanoPOTS <sup>187</sup>.

Il existe aussi des développements en CE (électrophorèse capillaire) couplée à la spectrométrie de masse mais ils ont plus rarement été adaptés à des protocoles en cellule unique. Cette technique présente pourtant une séparation très efficace de volumes très faibles d'échantillons (quelques nl) avec une limite de détection très faible (zmol) <sup>183</sup>. Lors de son utilisation pour des analyses protéomiques en cellules uniques, les études portaient sur des cellules larges comme les neurones de souris ou des ovocytes de grenouilles et de poisson-zèbre <sup>188, 189</sup>.

## INTRODUCTION : Les avancées en protéomique de la cellule unique

L'acquisition des données par spectrométrie de masse a également évolué. DO-MS a été développée par l'équipe de Slavov pour optimiser l'acquisition des données MS. En DDA, 1000 protéines par cellules ont pu être identifiées voire 1500 protéines avec les nanoPOTS. En DIA, avec scope MS, jusqu'à 2000 protéines par cellule ont pu être identifiées.

De plus, la parallélisation du traitement des échantillons grâce au multiplexage a permis d'accélérer le temps des expériences. En effet, même si la méthode *label free* apporte une précision de quantification supérieure à celle des marquages, le multiplexage avec des marqueurs comme le TMT permettent d'identifier jusqu'à 1000 protéines par cellule et permettent l'analyse de plus de 100 cellules par jour<sup>124</sup>. Aujourd'hui, des études ont même couplé le multiplexage au TMT avec une analyse DIA, c'est le cas du plexDIA de l'équipe de Slavov par exemple<sup>177</sup>. Cependant, ces techniques de parallélisation utilisent souvent un champ pour marquer un échantillon de 100 ou 200 cellules qui sera mélangé par la suite aux cellules uniques pour booster l'identification. L'utilisation d'échantillons 100 fois plus abondants a suscité des inquiétudes quant aux biais de quantification qui en résultent<sup>124</sup>.

Les améliorations énoncées précédemment, surtout la diminution des volumes des échantillons de l'ordre du nanolitre apportent des avantages indéniables puisque les rapports signal/bruit améliorés permettent l'identification de protéines sur des cellules uniques sur des spectromètres de masse existant depuis plus de 10 ans<sup>124</sup>. Un exemple est l'utilisation d'une trappe à ions avec une méthode DIA sur de très faibles quantités d'échantillons<sup>190</sup>.

L'instrumentation s'est adaptée au domaine puisque des couplages avec la mobilité ionique piégée (TIMS, Bruker) ou de la spectrométrie de mobilité ionique asymétrique de champ (FAIMS, Thermo) ont été développés pour améliorer la sensibilité des analyses. D'ailleurs, à partir d'une très petite quantité d'échantillon, la plus haute couverture de protéome a été reportée sur un instrument TIMS-TOF (Bruker)<sup>124</sup>.

Pour l'analyse des données, il n'existe aujourd'hui que très peu de *packages* dédiés à la protéomique en cellule unique<sup>124</sup>. DART-ID (pour « *data-driven alignment of retention times for identification* ») complète l'analyse MaxQuant en améliorant l'alignement des temps de rétention par le calcul d'un score<sup>175</sup>. En langage R, il existe aussi les *packages* SCoPE2 et SCP et en langage Python, Sceptre a été implémenté<sup>191</sup>.

Ces solutions permettent aujourd'hui de dépasser le stade du développement technologique et commencent à répondre aux questions biologiques<sup>124</sup>.

Dans la recherche embryonnaire, les ovocytes de mammifères subissent une série de modifications à travers l'ovogenèse, la maturation, la fécondation et le développement embryonnaire précoce<sup>27</sup>. L'analyse en protéomique de la cellule unique est nécessaire pour identifier les produits protéiques fonctionnels à différents stades de croissance et de maturation. Virant-Klun *et al.* ont identifié 2154 protéines dans 100 ovocytes humains et ont montré que leur protéome favorise la fixation cellulaire, l'homéostasie et les interactions environnementales via des protéines secrétées<sup>192</sup>.

Dans la recherche sur le cancer, la protéomique en cellule unique a transformé la compréhension de l'hétérogénéité et des caractéristiques biologiques des tumeurs en fournissant de meilleurs diagnostics basés sur les caractéristiques moléculaires. Cette technologie peut être utilisée pour identifier des sous populations rares de cellules, des

## INTRODUCTION : Les avancées en protéomique de la cellule unique

cellules tumorales circulantes, pour caractériser le microenvironnement tumoral, les mécanismes de progression, de métastase ou de résistance aux traitements <sup>27</sup>.

Afin d'établir le profil de différents stades du cycle cellulaire au niveau de la cellule unique, Brunner *et al.* ont comparé les résultats d'une perturbation du cycle de cellules cancéreuses par des médicaments analysée par protéomique et transcriptomique <sup>193</sup>. Ils ont rapporté des regroupements cellulaires différents selon le transcriptome et le protéome, ce qui selon eux appuie la notion de modes de régulation différents entre les ARNm et les protéines. Cependant, ils ont observé une variation des protéines beaucoup plus faible que celle des ARNm correspondants et ont attribué ceci à la nature stochastique de l'expression des transcrits par rapport aux protéines <sup>181</sup>. Ils ont déduit que si la variation due au traitement était plus visible sur les ARNm, cela pourrait être grâce au débit et l'exhaustivité des technologies de séquençage des ARN qui sont beaucoup plus élevés que ceux de la protéomique aujourd'hui. Enfin, ils suggèrent l'utilisation des données de protéomique en cellule unique pour valider les résultats de transcriptomique ou de génomique au niveau des protéines <sup>181</sup>.

Les macrophages, selon leur polarisation, peuvent avoir des fonctions pro-inflammatoires ou anti-inflammatoires. Leur diversité phénotypique rendait très difficile leur analyse au niveau cellulaire <sup>27</sup>. L'exploration de la différenciation des monocytes en macrophages par le protocole SCOPE2 a mis en évidence un gradient d'états protéomiques qui a révélé l'émergence de l'hétérogénéité des macrophages en l'absence de cytokines polarisantes <sup>135</sup>. Cette recherche a également permis d'explorer les interactions entre la protéine p53 (suppresseur de tumeur), son transcrit et les gènes régulés par cette protéine.

En 2022, Osburn *et al.* ont décrit l'analyse de cellules uniques provenant du modèle KRAS<sup>G12C</sup> de cancer du poumon NCI-H-358 et ont constaté que les protéines quantifiées par une analyse parallélisée de plusieurs cellules individuelles grâce au TMT et mesurées par un TIMS-TOF Flex (Bruker) se regroupaient selon le stade du cycle cellulaire <sup>194</sup>. La grande abondance intracellulaire relative des protéines histones a permis d'identifier avec confiance les sites d'acétylation, de méthylation et de diméthylation des histones dans la majorité des cellules analysées. Le protocole par PASEF-RIQ (pour « reporter ion quantification ») utilisé fournit des informations quantitatives précises à des concentrations d'une cellule unique et est moins entravé par l'effet des cellules *carriers* que d'autres configurations. De plus, l'utilisation d'un instrument TIMS-TOF permet une couverture de séquence élevée pour chaque protéine identifiée. Les résultats obtenus sont concordants avec ce qui est trouvé dans la littérature sur l'inhibition de KRASG12C en *bulk*. Ils ont observé la suppression de KRAS ainsi que des protéines dans les voies VEGF et MAPK qui sont hyperactivées par la protéine mutante KRAS. Cependant, lorsque les cellules sont considérées individuellement, les auteurs ont montré des résultats moins clairs qui reflètent l'hétérogénéité de la réponse identifiée par une analyse en cellule unique.

La limite principale de la plupart de ces approches est l'utilisation d'appareils très spécialisés qui ne sont pas nécessairement faciles à utiliser par la communauté <sup>181</sup>. En revanche, la majorité des exemples cités utilisent des plaques standard 96 ou 384 puits pour le traitement des cellules. Seules quelques équipes se sont concentrées sur l'étude des supports (nanoPOTS, ProteoCHIP). Cependant, la sensibilité peut être affectée par l'écart important entre les contenants et les volumes de réactions conventionnels (10 à 100µl) qui ont été développés pour des échantillons contenant des millions de cellules et la dimension

## INTRODUCTION : Les avancées en protéomique de la cellule unique

de quelques cellules ( $1E-7$  à  $1E-3\mu\text{l}$ ), soit un rapport de 4 à 9 ordres de grandeur. L'amélioration de la préparation des échantillons nécessite principalement de réduire les volumes de réaction par la miniaturisation des dispositifs pour les adapter aux petites quantités de matériel et minimiser les surfaces de contact, par l'amélioration des techniques séparatives en réduisant par exemple le diamètre interne des colonnes de séparation ainsi que l'amélioration de la sensibilité d'analyse en spectrométrie de masse en adaptant les paramètres d'analyse. Il faut de surcroît éviter les pertes et l'évaporation, envisager une plus haute parallélisation pour conserver des temps d'analyse réalistes, soit au niveau de la séparation par multiplexage des échantillons par exemple avec l'utilisation de TMT, soit au niveau de la spectrométrie de masse par multiplexage des précurseurs séquencés par exemple par DIA <sup>195,196</sup>.

## II. OBJECTIFS DU PROJET DE RECHERCHE

### **Chapitre 4 Objectifs**



## OBJECTIFS DU PROJET DE RECHERCHE : Objectifs

Au niveau cellulaire, la méthodologie basée sur la spectrométrie de masse n'en était qu'à ses débuts en 2019 car confrontée à plusieurs défis tels que la miniaturisation, la sensibilité et le débit. Il fallait concevoir une préparation d'échantillons de traces compatible avec une analyse en MS, optimiser une méthode de séparation assez sensible, gérer le débit et éviter la perte d'échantillons. Dans le cadre de ce projet de thèse en collaboration avec Sanofi, l'objectif est d'améliorer la sensibilité et la préparation des échantillons afin d'optimiser la couverture du protéome sur cellules individuelles. Décrire un protéome total sur cellule unique n'est pas possible en spectrométrie de masse pour des raisons inhérentes aux principes liés à la détection par spectrométrie de masse pour la protéomique *bottom-up*. Ce travail a permis de développer quelques éléments du protocole de protéomique sur cellule unique afin d'identifier les signatures protéomiques de monocytes et de lymphocytes humains. Pour approfondir l'analyse du protéome pour une petite quantité de cellules, deux axes de travail ont été développés.

De l'isolement des cellules à l'analyse LC-MS/MS, chaque étape du processus de préparation des échantillons nécessite un développement minutieux afin d'augmenter la sensibilité de l'analyse et de réduire les pertes d'échantillons avant leur détection. Dans la miniaturisation de ce processus, outre la LC haute performance et un spectromètre de masse de dernière génération pour atteindre la sensibilité et résolution requises, la préparation de l'échantillon est une étape très importante pour une sensibilité optimisée. Pour atteindre la sensibilité requise, cette préparation doit être effectuée dans la même échelle de concentration afin de minimiser les pertes non spécifiques. La première partie du travail a donc été dédiée au développement et à l'optimisation de différents protocoles expérimentaux de protéomique. Les développements technologiques ont porté sur l'isolement des cellules, la miniaturisation de la préparation des échantillons (la lyse cellulaire, les différentes réactions, la digestion enzymatique), la minimalisation du protocole et l'amélioration de la séparation des échantillons en concentrant les efforts sur les composants impliqués en chromatographie liquide (colonnes, configurations). Enfin, nous avons utilisé des spectromètres de masse à analyseur à haute résolution tels que l'orbitrap Q Exactive HF (Thermo Fisher Scientific), ou le temps de vol Tims TOF Pro™ (Bruker). Les développements technologiques ont porté sur la récupération des peptides et des protéines pour améliorer la couverture du protéome pour un très faible nombre de cellules.

D'autre part, la sensibilité d'une analyse ne dépend pas seulement de la sensibilité de l'instrument utilisé pour l'analyse mais aussi du traitement et de la manipulation de l'échantillon impliqué. Ce dernier point dépend de la qualité des surfaces avec lesquelles les échantillons sont en contact à chaque étape du protocole. Le matériel utilisé est très important pour limiter l'adsorption des peptides et des protéines. Les tubes spécifiquement conçus pour minimiser l'adsorption des polypeptides, dits *low binding*, qui sont utilisés en protéomique ont beaucoup d'avantages, mais ces tubes commerciaux ont des surfaces de contact de plusieurs ordres de grandeur supérieurs à la surface d'une cellule. Il faut trouver un support avec le bon matériau et une surface suffisamment petite pour travailler avec une petite quantité de cellules. La seconde partie a donc été consacrée à l'étude des surfaces des supports sur lesquels sont déposées les cellules dès le début du protocole. Cela implique la miniaturisation complète de plusieurs stratégies de micro-fabrication pour le traitement cellulaire avec la conception d'un support dédié. Dans le présent travail, nous avons optimisé les supports de protéines en réduisant la taille de la surface de contact avec des micro-puits faits maison et en modifiant leurs propriétés physicochimiques avec un traitement de surface dédié. En réduisant la surface d'adsorption et le volume de réaction, l'adsorption non spécifique doit être réduite, tandis que la récupération et l'identification seront améliorées.

### III. ETUDES ET RESULTATS

## Chapitre 5 Adaptation du protocole à l'analyse de petites quantités d'échantillons

<b>5.A Manipulation de petites quantités de cellules .....</b>	<b>65</b>
5.A.1. Protocole adapté d'analyse des cellules en petites quantités .....	65
(a) Choix des cellules.....	65
(b) Choix des échantillons standard .....	65
5.A.2. Isolement des cellules .....	66
(a) Inspiration du smart Aliquotor .....	66
(b) FACS and cell sorter for high throughput .....	75
<b>5.B Étude de l'étape de réduction et d'alkylation .....</b>	<b>77</b>
<b>5.C Étude de l'étape de digestion.....</b>	<b>80</b>
5.C.1. Un temps de digestion adapté aux petites quantités .....	80
5.C.2. Un rapport enzyme/protéine mieux adapté aux petites quantités .....	81
5.C.1. Du DMSO pour limiter l'évaporation.....	83
<b>5.D Optimisation des conditions d'analyses en LC-MS .....</b>	<b>85</b>
5.D.1. Colonnes et configurations du système de chromatographie .....	85
5.D.2. Fabrication de colonnes "homemade" .....	87
5.D.3. Essais sur le Tims TOF Pro.....	91

ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

## 5.A Manipulation de petites quantités de cellules

### 5.A.1. Protocole adapté d'analyse des cellules en petites quantités

#### (a) Choix des cellules

Deux modèles cellulaires ont été mis en culture tout au long de ce projet pour tester les protocoles.

La cellule THP-1 est un monocyte isolé du sang périphérique d'un patient atteint de leucémie monocyttaire aiguë. C'est un modèle de monocyte humain très utilisé dans la recherche en immunologie et toxicologie. Elles sont souvent utilisées dans la recherche contre l'inflammation et même en protéomique de la cellule unique<sup>197,198</sup>. Les cellules THP-1 font entre 16 et 23  $\mu\text{m}$  de diamètre en fonction de leur état de croissance, ce qui les rapproche en taille et en quantité protéique des cellules HeLa et en fait des bons modèles cellulaires<sup>94</sup>.

Le clone Jurkat E6-1 est un lymphocyte T de la lignée cellulaire Jurkat-FHCRC, dérivé de la lignée cellulaire Jurkat, établie à partir du sang périphérique d'un patient de 14 ans, atteint de leucémie aiguë à prédominance de cellules T. C'est un modèle de lymphocyte T humain utilisé dans la recherche en immunologie, immuno-oncologie ou même contre le COVID<sup>199,200,201</sup>. La particularité de ces cellules est leur petite taille (entre 7 et 10  $\mu\text{m}$  de diamètre en fonction du cycle cellulaire), ce qui en fait un bon modèle de petites concentrations protéiques pour les protocoles testés dans cette étude<sup>94</sup>.

#### (b) Choix des échantillons standard

En plus des échantillons cellulaires, des standards commerciaux utilisés classiquement en protéomique comme des digestats commerciaux de HeLa (Pierce, Thermo) ou de levure *Saccharomyces cerevisiae* (Promega) et des extraits protéiques de levure (*Saccharomyces cerevisiae*) et humains (macrophages K562 humains) (Promega) ont été utilisés dans cette étude.

ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

### 5.A.2. Isolement des cellules

L'étape d'isolement des cellules est obligatoire pour travailler sur une seule cellule ou même sur une petite quantité de cellules dès le début d'un protocole. Il existe trois méthodes pour l'isolement des cellules uniques : la dilution en série, la micromanipulation et la cytométrie de flux.

#### (a) Inspiration du smart Aliquotor

Le design du Smart Aliquotor (TebuBio) utilisé pour l'isolement des cellules en cellules uniques nous a inspiré pour créer notre propre design en PDMS. Il s'agit d'un isolement à la pipette depuis un puits central vers 100 micro-puits à travers des canaux (Figure 50).

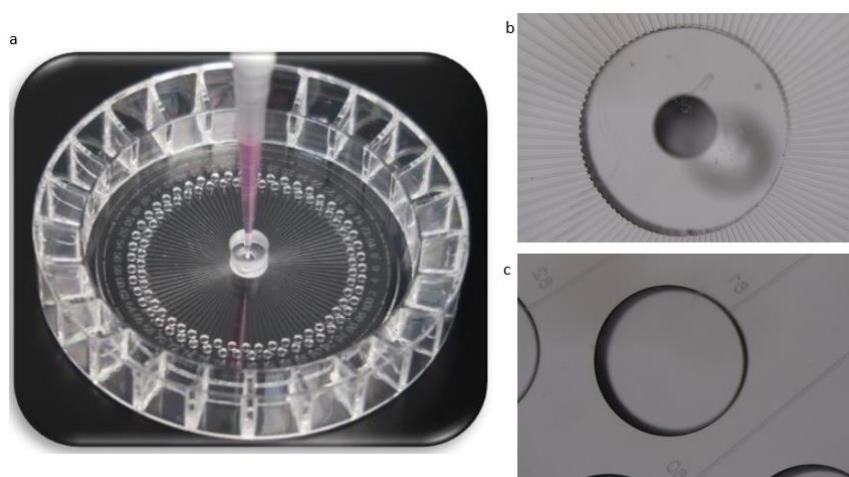


Figure 50 : (a) Photographie du Smart Aliquotor (TebuBio). (b) Zoom sur le puits central et les 100 canaux qui l'entourent. (c) Zoom sur le canal qui arrive au puits récepteur n°61.

#### Méthode

Le smart Aliquotor a d'abord été photographié et mesuré à l'aide d'une loupe binoculaire. Pour cela, un marqueur d'1mm de long permet d'établir une courbe de correspondance des tailles observées en fonction du grossissement du microscope choisi (Figure 51). Par exemple, avec un grossissement 100X, le marqueur de taille totale d'1mm mesure 8cm. Cela permet de remplir le tableau avec les tailles observées pour les différents composants du smart Aliquotor et le grossissement choisi pour chaque objet, puis de faire la correspondance avec la taille réelle en mm et en  $\mu\text{m}$  (Tableau 2). Par exemple, avec un grossissement 100X, le canal fait 0.3cm de largeur observée, il mesure donc  $60\mu\text{m}$  en taille réelle. Cela permet de redessiner un schéma du smart Aliquotor avec le puits central dont le diamètre externe est de 2.77mm et le diamètre interne est de 0.7mm puis un canal de 0.6mm de long et un puits d'arrivée de 1.8mm de diamètre (Figure 52).

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

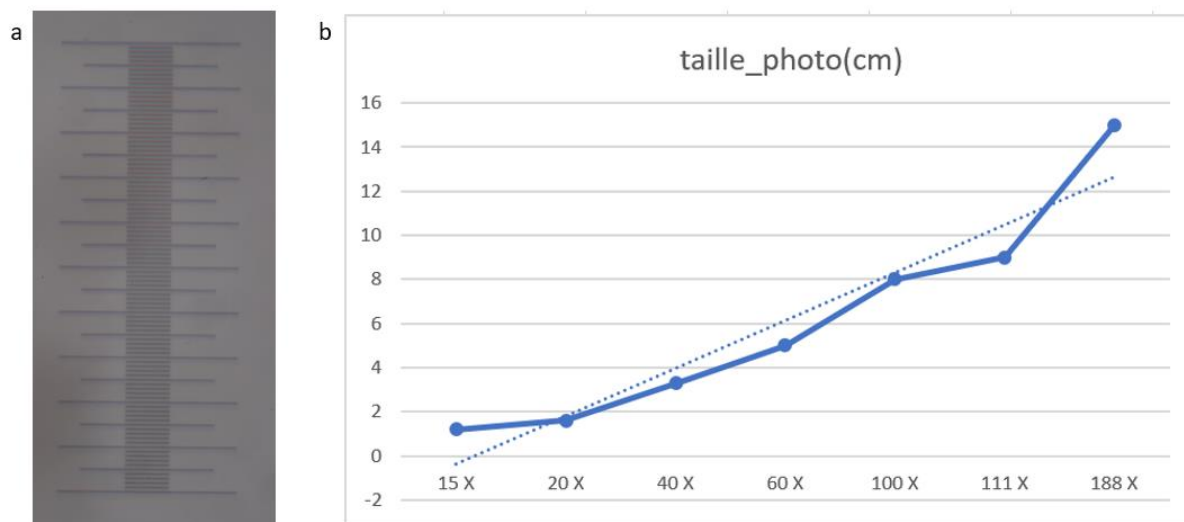


Figure 51 : (a) Photographie du marqueur de 1mm de taille à mesurer sous le microscope (grossissement 100x) (1mm taille totale, 100 $\mu$ m taille entre deux graduations longues, 10 $\mu$ m taille entre deux graduations courtes) (b) courbe des correspondances de la taille (en cm) sur la photographie du marqueur à 1mm en fonction du grossissement choisi.

Objet	Grossissement	Taille observée (cm)	Taille réelle (mm)	Taille réelle ( $\mu$ m)
Diamètre de l'injecteur	50 x	3.5	0.77	770
Diamètre du puits central	50 x	12.5	2.77	2770
Diamètre des puits	60 x	9	1.8	1800
Section du canal	60 x	0.3	0.06	60
	100 x	0.5	0.06	60

Tableau 2 : Mesure des canaux du smart Aliquoter photographiés par le microscope confocal.

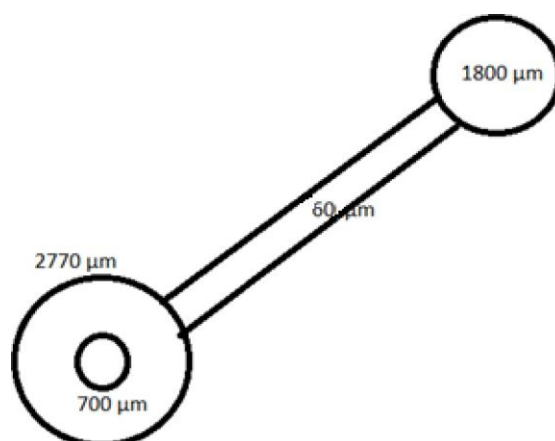


Figure 52 : Schéma du smart Aliquoter avec le puits central dont le diamètre externe est de 2.77mm et le diamètre interne est de 0.7mm puis un canal de 0.06mm de large et un puits d'arrivée de 1.8mm de diamètre.

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

Le processus de création et de fabrication des puces pour l'isolement cellulaire commence par la conception d'un design puis une technique de micro-fabrication suivie d'un test avec des colorants et des billes fluorescentes (Figure 53). La technique de micro-fabrication choisie est la photolithographie douce qui permet de créer un motif, de l'imprimer en masque et de l'appliquer à un *wafer* en métal contenant une résine par-dessus.

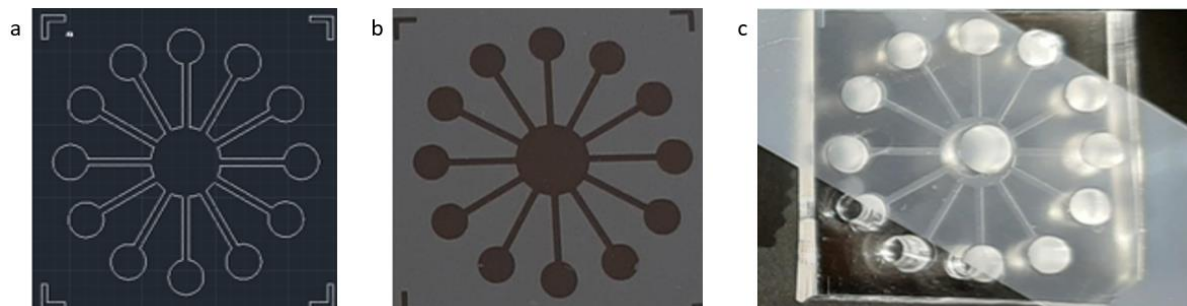


Figure 53 : Différentes étapes de création d'une puce de séparation microfluidique pour l'isolement des cellules. (a) dessin du modèle et choix des dimensions. (b) micro-fabrication commence par l'impression d'un masque pour la photolithographie et processus de photolithographie sur un *wafer*. (c) coulage du PDMS pour former la puce et poinçonnage des puits puis collage sur une lame de verre pour l'étanchéité.

Le design du Smart Aliquoter nous a inspiré un design en flocon de neige proche de celui-ci. L'idée de base était de mettre les cellules dans un puits principal à la pipette puis d'avoir des canaux pour transporter ces cellules à des puits secondaires depuis lesquelles les cellules séparées pourraient être récupérées. La technique de micro-fabrication choisie étant la photolithographie douce avec une résine en film sec, la hauteur des canaux est limitée à l'épaisseur du film choisi (ici  $50\mu\text{m}$ ).

Différentes idées de conception ont été dessinées sur le logiciel AutoCAD (*student* version, 2020) en variant la largeur des canaux, leur longueur, leur nombre ou encore le diamètre des puits (Figure 54). Une de ces conceptions (design F) était également un dessin linéaire avec un puits principal et un canal principal dont sortent des puits récepteurs. Le protocole de photolithographie est décrit dans la partie matériel et méthodes. Les dimensions des différents puits et canaux sont indiquées grâce au logiciel AutoCAD sur la Figure 126. Les diamètres des puits centraux et récepteurs sont indiqués sur le tableau (Tableau 3). Le *wafer* prêt est montré dans la Figure 56. Les puces sont montrées sur la Figure 56.

ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

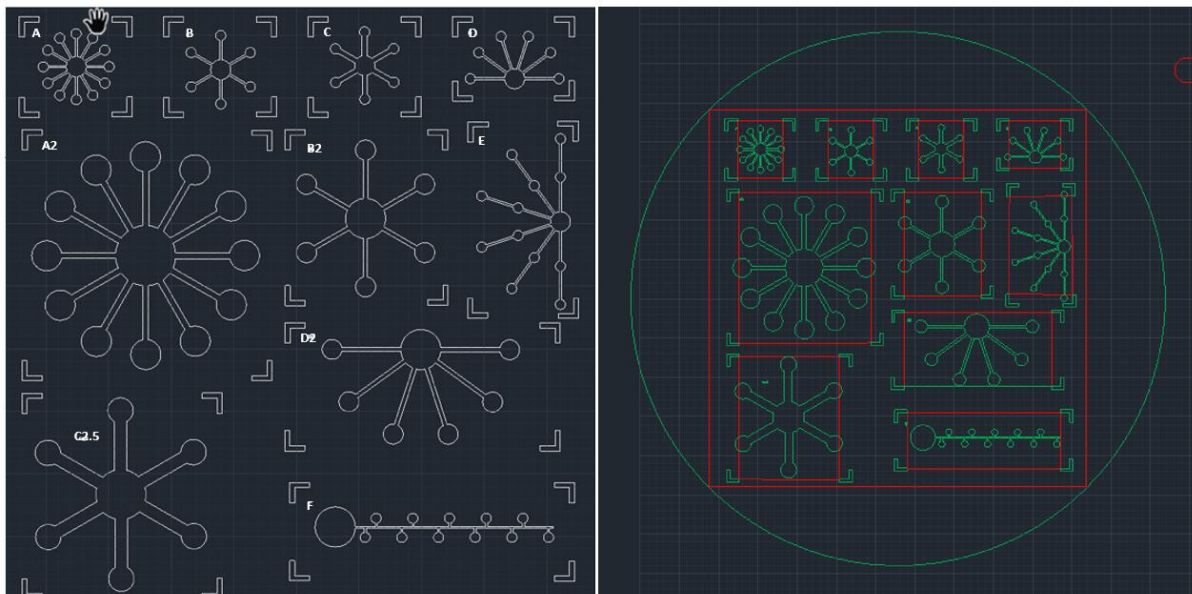


Figure 54 : Designs sur AutoCAD de différentes puces (de A à F) avec différentes formes et dimensions pour l'isolement cellulaire.

<i>Puce</i>	<i>Diamètre puits récepteur</i>	<i>Diamètre puits central</i>
C2.5	2500	5000
F	1000	4000
D2	2000	4000
A3	3000	6000
B2	2000	4000
E	1000	2000
D	1000	2000
C	1000	2000
B	1000	2000
A	1000	2000

Tableau 3: Diamètres des puits centraux et récepteurs en  $\mu\text{m}$ .



## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

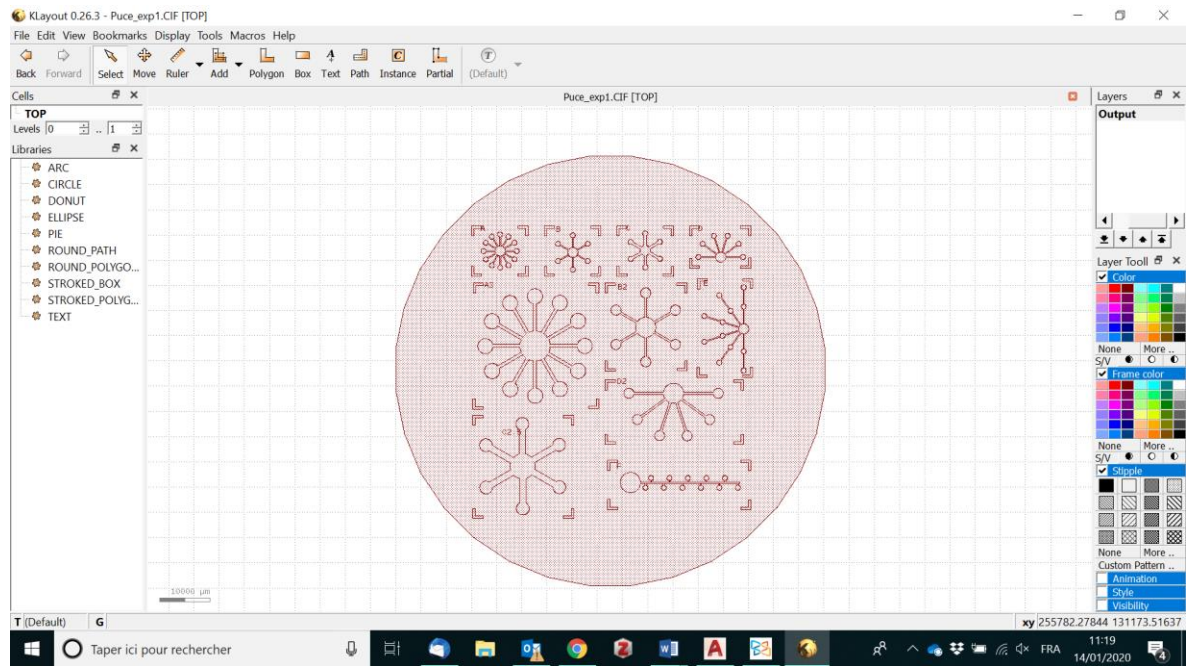


Figure 55 : Ouverture du fichier avec le logiciel Klayout et délimitation des designs de puces.

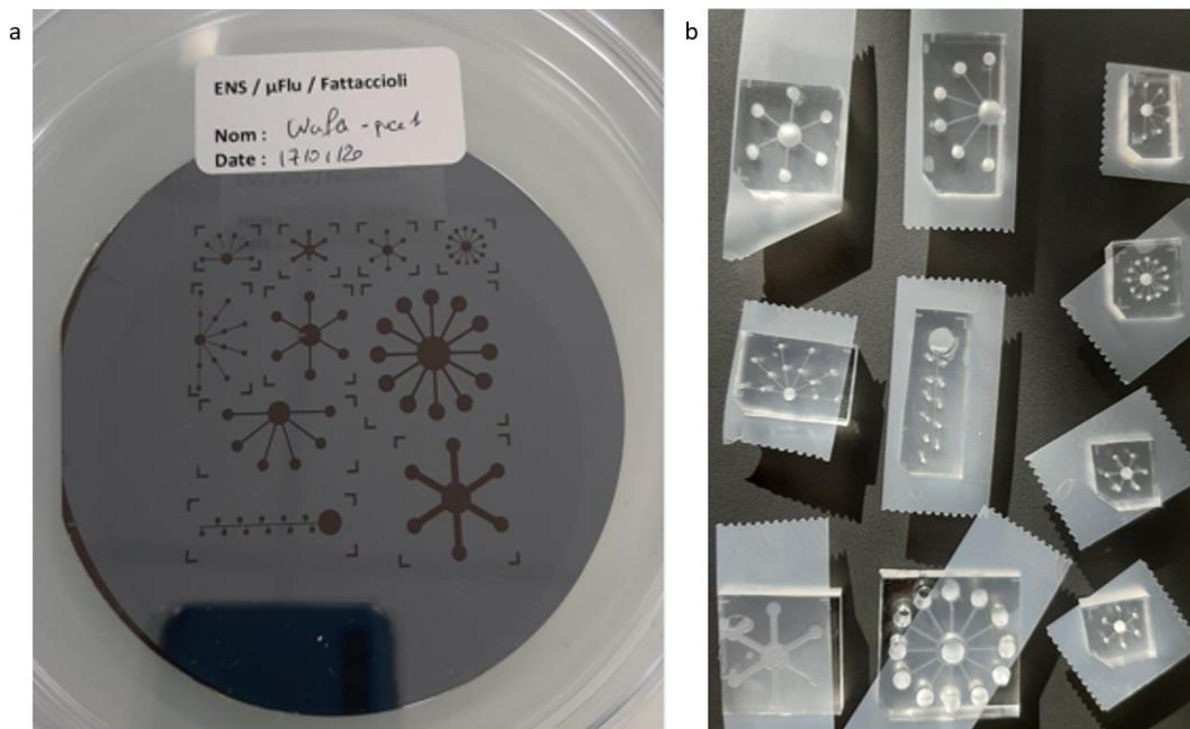


Figure 56 : (a) *Wafer* avec les différents modèles. (b) Différentes puces en PDMS fabriquées à partir du *wafer*.

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

### Résultats

Le *wafer* obtenu par photolithographie a permis la fabrication des puces selon les différents modèles dessinés. Chacun de ces modèles a été testé avec des colorants alimentaires pour vérifier le mouvement du liquide par capillarité depuis le puits central à travers les canaux jusqu'aux puits récepteurs.

Pour faciliter la visualisation, différents volumes (en fonction de la taille du puits central de chaque modèle) de colorants alimentaires (rouge ou bleu) ont été injectés et le mouvement du liquide par capillarité a été vérifié (Figure 57 et Figure 58).

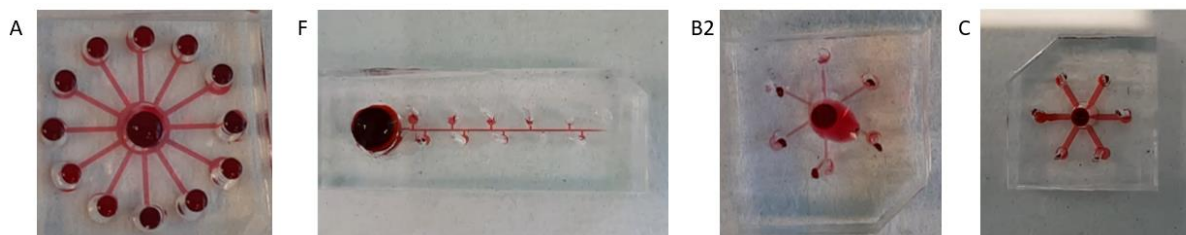


Figure 57 : Modèles A, F et C où le liquide (colorant rouge) s'écoule par capillarité du puits central vers tous les puits récepteurs.

Les 4 modèles (A, F, B2 et C) en flocons de neige ou linéaire sur la figure montrent une bonne distribution du colorant rouge après injection dans le puits central vers tous les canaux du design et tous les puits récepteurs (Figure 57).

Cependant, sur d'autres modèles (D, D2 et E) en éventail avec des canaux fins ou éventail avec une succession de deux puits récepteurs, cette distribution n'était plus homogène et montrait des canaux non irrigués et des puits non atteints par le colorant bleu.

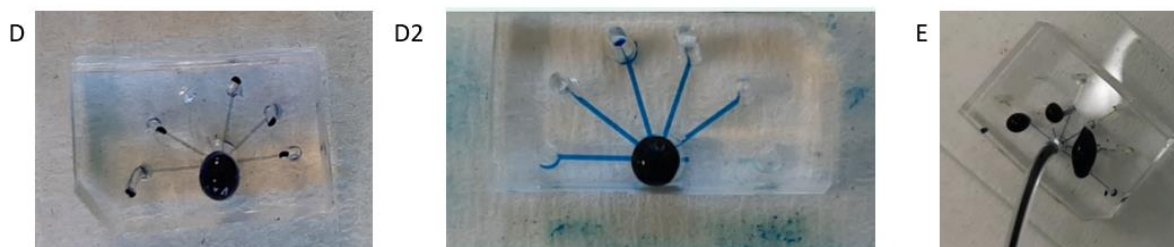


Figure 58 : Modèles D, D2 et E où le liquide (colorant bleu) ne s'écoule pas dans tous les canaux et n'atteint pas tous les puits récepteurs.

Dans la figure, sur les 3 modèles, le volume de colorant dans le puits central a été doublé mais l'équilibre de l'écoulement par capillarité semblait atteint après 10 minutes d'attente et certains puits n'étaient toujours pas irrigués par le colorant (Figure 58). Sur le modèle E, même une injection continue amenée par une seringue n'a pas réussi à atteindre les seconds puits récepteurs. Le temps d'attente pour atteindre un équilibre et vider le puits central par capillarité était de 10 minutes lorsque le puits central était rempli de 62 $\mu$ l et de 5 minutes lorsque le puits central était rempli de 16 $\mu$ l. Or, cela peut causer une sédimentation des cellules dans le puits central et les canaux avant leur arrivée dans les puits récepteurs. Les designs A2 et C2.5 n'ont pas été montrés ici car leurs dimensions étaient beaucoup plus grandes et le temps d'attente pour vider le puits central par effet de capillarité était très long.

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

Un test avec des billes en polystyrène fluorescentes de  $5.7\mu\text{m}$  de diamètre (Fluoresbrite<sup>®</sup> plain YG micro-spheres) diluées au  $1/100^{\text{ème}}$  a été effectué pour mimer la séparation de cellules. Sous microscope *Leica*, un grossissement 4X a été utilisé pour observer les puces. Selon le protocole, les billes ont été excitées à  $441\text{nm}$  et émettent à  $486\text{nm}$ . La fluorescence nous a permis d'observer le mouvement des billes et de suivre le remplissage des canaux et des puits. Les billes ont été injectées dans deux modèles (A et F) (Figure 59). Par fluorescence à transmission, les canaux des puces apparaissent transparents, les puits apparaissent en noir et les billes sont bleues. Elles sont visibles dans le puits central du modèle A sur la figure traversant les canaux vers les puits récepteurs (Figure 59a). Elles peuvent également être observées dans les puits récepteurs après leur migration due à l'effet de capillarité dans le modèle F (Figure 59b).

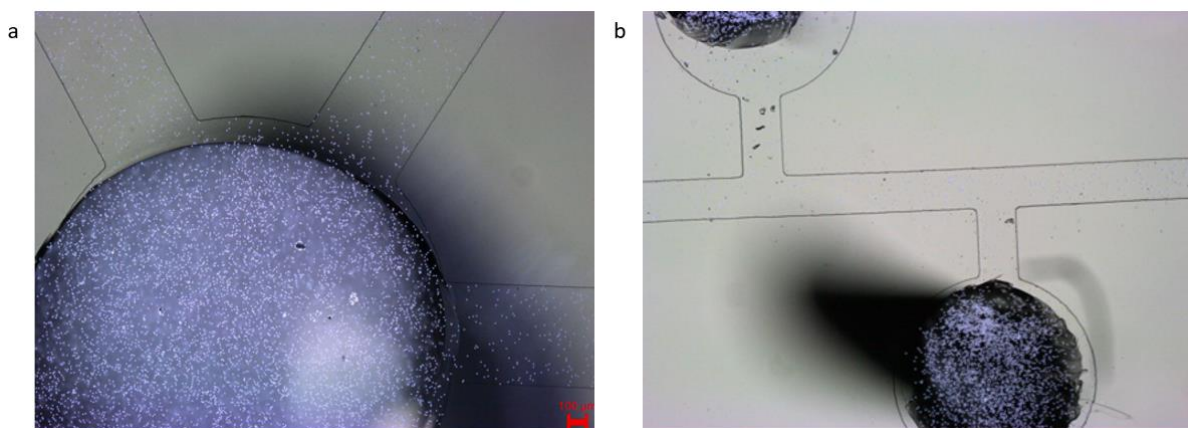


Figure 59 : Test sur deux modèles avec des billes fluorescentes. (a) Zoom sur le puits central dans le modèle A. (b) Zoom sur un puits récepteur dans le modèle F. La barre en rouge indique un marqueur de taille de  $100\mu\text{m}$ .

Enfin, un test avec des cellules Jurkat wt ( $7$  à  $10\mu\text{m}$  de diamètre) cultivées en laboratoire a été effectué dans les modèles A, B, C et F. Après préparation de nouvelles puces, des cellules à différentes concentrations ont été déposées dans les puits centraux des modèles. Grâce à un microscope optique, elles ont été observées avec deux grossissements différents 4x et 20x avec des flèches rouges indiquant les cellules lorsqu'elles sont difficiles à distinguer sur la Figure 60. Avec 3 concentrations de cellules différentes, la visualisation de la distribution cellulaire a fonctionné dans certains modèles mieux que d'autres.

Pour le modèle A, toutes les cellules sont restées bloquées dans le puits central avec quelques cellules arrivées jusqu'à un puits récepteur (Figure 60). Pour le modèle B, des centaines de cellules se sont dispersées dans les 3 puits inférieurs seulement. Pour les modèles C et F, une dispersion correcte a été observée à des concentrations de  $10^5$  et  $10^4$  cellules/ml dans tous les puits récepteurs.

La concentration de cellules a ensuite été diluée jusqu'à  $2000$  cellules/ml et  $100\mu\text{l}$  ont été injectés dans le modèle F (l'équivalent de  $200$  cellules). Les cellules ont migré du puits central vers les puits récepteurs par effet de capillarité mais plusieurs d'entre elles sont restées dans les canaux une fois l'équilibre atteint (Figure 60).

Plusieurs problèmes ont été rencontrés durant ces tests. La visualisation des puits qui apparaissent noirs en microscopie optique rendait impossible un comptage exact des cellules dispersées dans les puits récepteurs (Figure 61). La quantification de la dispersion et de

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

l'isolement cellulaire est donc impossible lorsque les cellules ne fluorescent pas, ce qui est le cas des Jurkat wt utilisées dans cet exemple. D'autre part, sans pression externe, la plupart des cellules à disperser restaient dans le puits central ou sédimentaient dans les canaux avant d'arriver aux puits récepteurs. De plus, le PDMS étant un matériau souple, le risque de cassure lors de l'injection n'était pas négligeable (Figure 61). Enfin, la récupération des cellules depuis les puits allait être très compliquée à mettre en place. Il fallait donc imaginer un nouveau design impliquant des puits récepteurs avec des diamètres plus grands que le puits central, des canaux qui se resserrent pour favoriser le flux dans un sens et des valves qui puissent se fermer afin de pouvoir faire la préparation des échantillons directement dans ces puits.

Toutefois, à cause de l'épidémie de COVID, l'ensemble de ces essais n'a pas pu être poursuivi. Nous avons alors décidé de changer de système d'isolement cellulaire et d'utiliser un trieur cellulaire commercial nouvellement disponible au laboratoire.

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

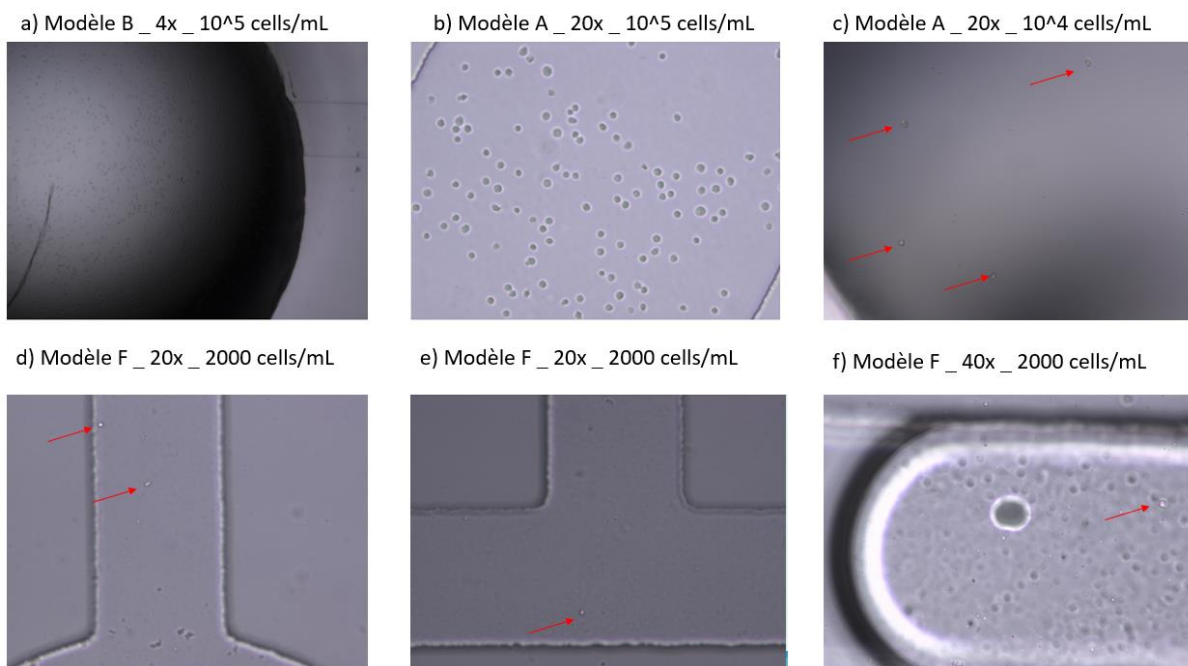


Figure 60 : Photographies de l'isolement cellulaire à différentes concentrations par les puces de différents modèles. (a) puits central du modèle B à  $10^5$  cellules/ml. (b) canal du modèle A à  $10^5$  cellules/ml. (c) puits récepteur du modèle A à  $10^4$  cellules /ml. (d) canal du modèle F à 2000 cellules/ml. (e) intersection des canaux du modèle F à 2000 cellules/ml. (f) avancée du liquide dans un canal après injection dans modèle F à 2000 cellules/ml. Les flèches rouges indiquent la présence de cellules.

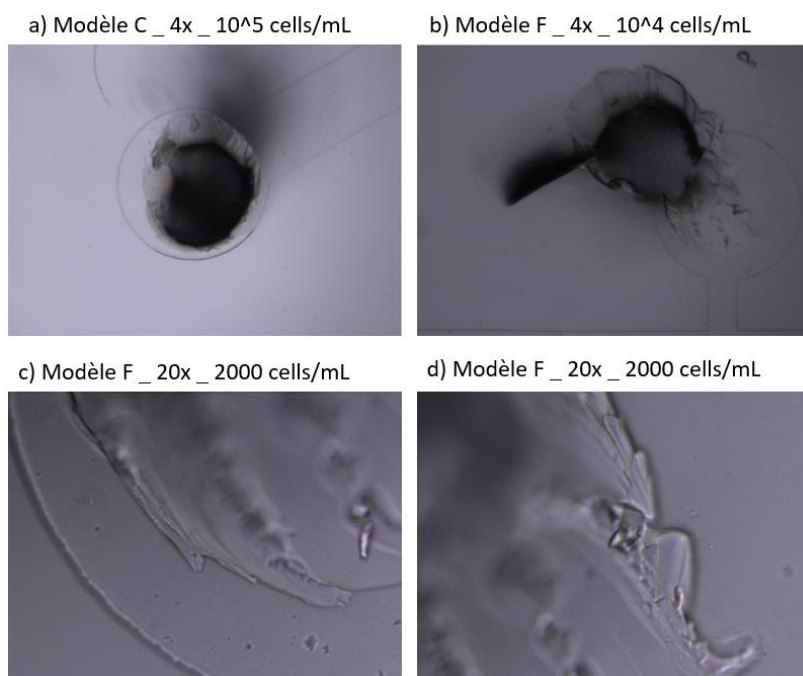


Figure 61 : Photographies de puits récepteurs dans deux modèles (C et F). (a) puits récepteur apparaît noir au microscope optique. (b) déchirure de la surface du puits due à la manipulation. (c) imperfections dans le poinçonnage des trous rend difficile la visualisation.

ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

### (b) FACS and cell sorter for high throughput

Le tri cellulaire activé par fluorescence aussi appelée FACS (pour « *fluorescence activated cell sorting* ») est un type de cytométrie de flux permettant de trier une cellule à la fois, un mélange hétérogène de cellules dans différents récipients, sur la base de caractéristiques spécifiques de diffusion de la lumière et de fluorescence de chaque cellule. C'est la technique d'isolement cellulaire la plus utilisée en protéomique de la cellule unique <sup>202</sup>. Brièvement, la suspension cellulaire est entraînée au centre d'un flux étroit de liquide (souvent, du PBS). Une vibration sépare le flux en gouttelettes contenant une cellule (le flux est fait de telle sorte à créer de la distance entre les cellules). Une station de mesure détecte la fluorescence de chaque cellule grâce à un laser. Les gouttelettes sont ensuite chargées en fonction de la détection de fluorescence et déviées en fonction de leur charge dans une position ou l'autre dans le cas d'un tri en tube (Figure 62).

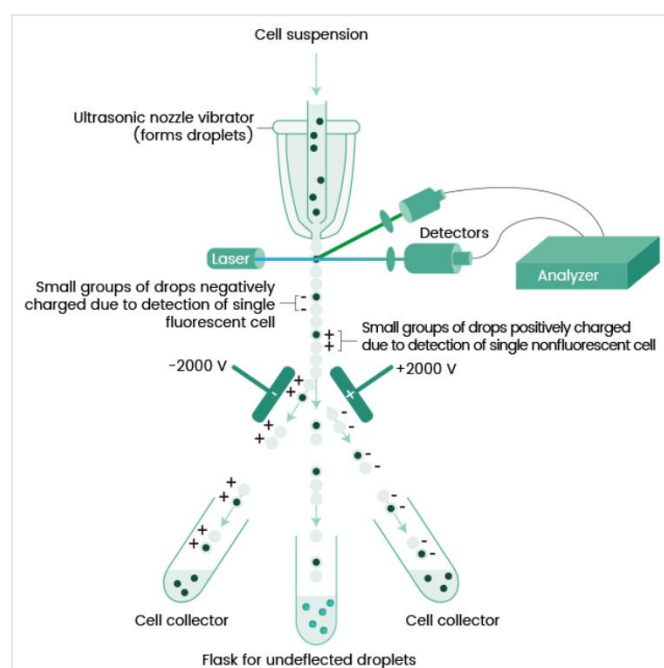


Figure 62 : Principe du tri cellulaire activé par la fluorescence (FACS).

Dans ce projet, le trieur de cellules SH800 (Sony) est utilisé (Figure 63). La séparation peut se faire en FACS basée sur la fluorescence de marqueurs spécifiques ou en tri cellulaire en série basé sur les marqueurs de tailles des cellules <sup>203</sup>. Les cellules étudiées dans ce projet ne sont pas marquées par des anticorps fluorescents. La principale caractéristique de ce trieur est l'utilisation d'une puce microfluidique avec un embout de 100µm permettant la création de gouttelettes de 3.6nL en flux constant (Figure 64). De cette façon, le volume et l'espacement des gouttes sont fixes et les cellules sont « positionnées » de façon aléatoire dans les gouttes. Ensuite, un mode de tri permet de choisir la pureté du tri en fonction de la résolution souhaitée. En mode « cellule unique », cette pureté est fixée à 100% ce qui signifie que seules les gouttelettes contenant une seule cellule cible dans leur région centrale sont triées. Les cellules peuvent être récupérées en tubes ou en plaques. Un exemple de tri est montré dans la Figure 65. Cette méthode allie un haut débit, un comptage exact des cellules, une rapidité d'exécution et une très bonne reproductibilité.

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

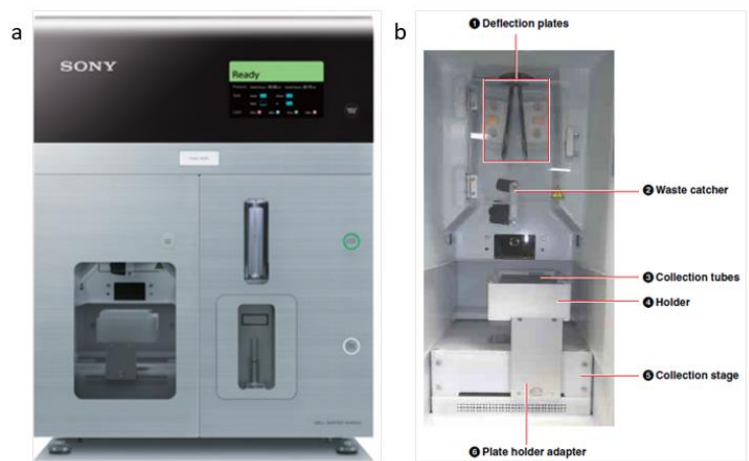


Figure 63 : Photographie du trieur de cellule SH-800 (Sony) (a) l'appareil (b) zoom sur l'emplacement où se fait la collection des cellules avec les plaques de diffraction (1), la poubelle (2), l'emplacement de la plaque (3), un porte-plaques (4) et un adaptateur (6).

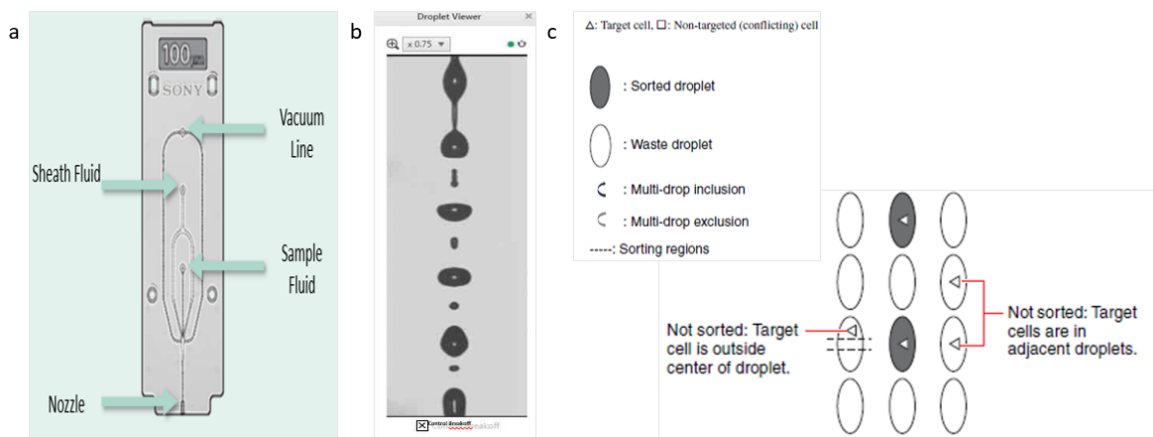


Figure 64 : La puce microfluidique et le mode de tri. (a) la puce microfluidique utilisée pour créer (b) un flux de gouttelettes permanent. (c) mode de tri en cellule unique. Seules les gouttelettes contenant des cellules cibles en leur centre sont triées (en gris), les autres sont jetées à la poubelle (en blanc).

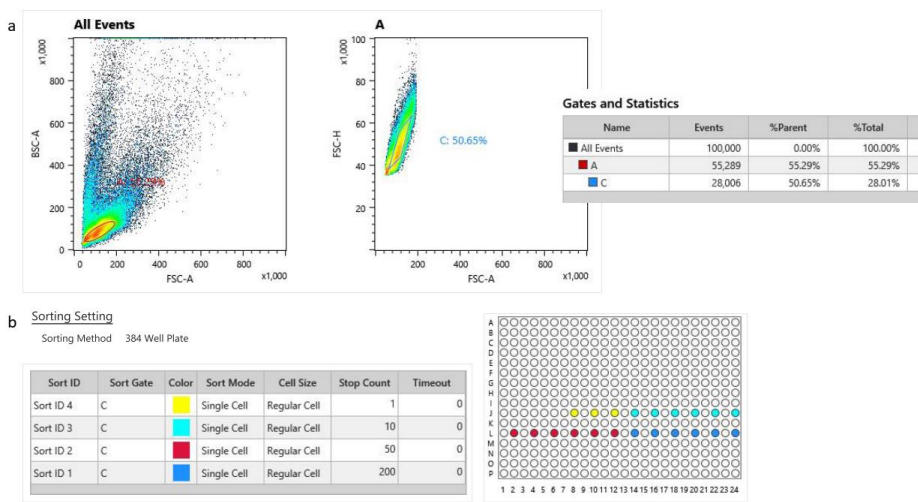


Figure 65 : (a) Stratégie de tri en fonction de la taille des cellules (sélection A) et élimination des doublets en faisant une seconde sélection (sélection C). (b) Tri sur une plaque 384 standard avec un tri en mode cellule unique de 200, 50, 10 et 1 cellules.

### 5.B Étude de l'étape de réduction et d'alkylation

Dans un protocole de protéomique *bottom-up* classique, après la distribution des cellules, viennent les étapes de dénaturation des protéines incluant la réduction et l'alkylation puis la digestion des protéines en peptides avant leur analyse en LC-MS. La nécessité de ces étapes est théoriquement indispensable. Cependant, leur compatibilité et efficacité sur des petites quantités de protéines doit être vérifiée.

#### Méthode

Des aliquots d'1µl d'extraits protéiques de levure (50ng) ont été déposés au fond des puits des plaques micro-puits PDMS 10µl (support présenté en seconde partie du manuscrit). Ces extraits protéiques ont été soit alkylés, réduits puis digérés, c'est le protocole ERAD (extraits réduits, alkylés, digérés), soit digérés uniquement, c'est le protocole ED (extraits digérés). Les protocoles sont décrits dans la partie matériel et méthodes. Par la suite, l'équivalent de 10ng a été analysé sur un Q Exactive HF. La médiane des 3 réplicats est annotée au-dessus des barres d'intervalles de confiance à 95% de chaque condition sur la Figure 66.

#### Résultats

Avec le protocole ERAD, 251 protéines, 496 peptides et 527 PSM. Avec le protocole ED, 849 protéines, 3010 peptides et 3415 PSM ont été identifiés. D'après le test statistique effectué, le nombre d'identifications avec le protocole ED est significativement supérieur qu'avec le protocole ERAD. Dans les deux designs de plaques micro-puits PDMS 10µl, le protocole ED a démontré une meilleure identification pour les quantités de matériel analysées (10ng analysés). L'hypothèse émise est que le gain théorique apporté par les étapes de réduction et d'alkylation était entravé par la perte due à l'adsorption des peptides et protéines à la surface à cause de la durée supérieure de prolongement du contact entre l'échantillon et le support lors de l'ajout des étapes de réduction et d'alkylation.

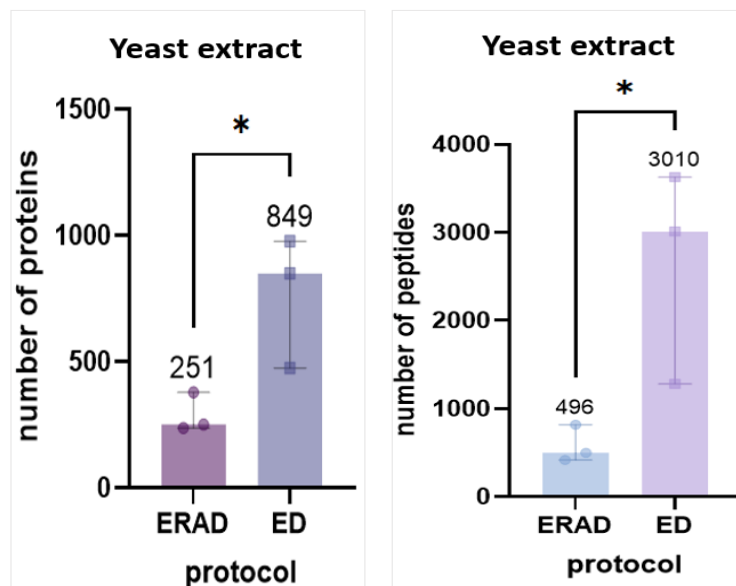


Figure 66 : Nombre de protéines et de peptides identifiés à partir de 10ng d'extraits de levure *Saccharomyces cerevisiae* dans les puces micro-puits PDMS 10µl, avec deux protocoles de préparation des échantillons. ERAD : extrait réduit alkylé et digéré, ED : extrait digéré. Les médianes des 3 réplicats est annoté au-dessus des barres d'intervalles de confiance à 95%. Les \* indiquent une p-value < 0.05 obtenue par un t-test non apparié (avec \*<0.05, \*\*<0.005, \*\*\*<0.0005).



## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

### *Méthode*

Trois protocoles ont également été comparés avec deux quantités de cellules Jurkat (100 et 1000 cellules). Pour rappel, selon nos estimations, 100 cellules Jurkat correspondent à 5 ng et 1000 cellules Jurkat correspondent à 50 ng de protéines (Tableau 1). Ces cellules ont été obtenues à partir d'un culot de  $10^5$  cellules et diluées en série en tubes. Une lyse thermique (cycle de congélation/température ambiante) a été réalisée. Le premier protocole appelé « protocole A » commence par une dénaturation des protéines par la chaleur à 70°C puis une réduction par du DTT à 10mM, alkylation par de l'IAA à 30mM et digestion par de la trypsine – Lys C de celles-ci. Le deuxième protocole appelé « protocole B » commence par une dénaturation des protéines par la chaleur puis une digestion enzymatique. Enfin, le troisième protocole appelé « protocole C » consiste en une digestion enzymatique des protéines. Les peptides ont été analysés sur un Q Exactive HF. La médiane des 3 ou 6 réplicats est annotée au-dessus des barres d'intervalles de confiance à 95% de chaque condition sur la Figure 67.

### *Résultats*

Pour les deux quantités de cellules, le protocole A qui inclut les étapes de réduction et d'alkylation a identifié très peu de protéines, peptides et PSM avec seulement 24 protéines identifiées pour 100 cellules et 382 protéines identifiées pour 1000 cellules. Ensuite, avec le protocole B qui inclut une dénaturation suivie d'une digestion, en moyenne 313 protéines, 806 peptides et 1031 PSM ont été identifiés pour 100 cellules et 2010 protéines, 12243 peptides et 20157 PSM ont été identifiés pour 1000 cellules. Enfin, avec le protocole C qui consiste en une digestion enzymatique directement après la lyse, en moyenne 228 protéines, 578 peptides et 768 PSM ont été identifiés pour 100 cellules et 1928 protéines, 12237 peptides et 21853 PSM ont été identifiés pour 1000 cellules.

Les protocoles B et C permettent une meilleure identification des protéines issues des cellules Jurkat que le protocole A. D'après ces résultats, les étapes de réduction et d'alkylation qui théoriquement améliorent la digestion enzymatique et donc l'identification, ont un effet contraire pour des petites quantités de cellules. L'hypothèse est que le temps prévu pour ces étapes augmente l'évaporation et l'adsorption des analytes sur le support. D'autre part, les protocoles B et C présentent des identifications comparables pour 1000 cellules suggérant que la dénaturation protéique apporte une légère amélioration en nombre de protéines identifiées mais non significative d'après le test statistique effectué. Cependant, pour 100 cellules Jurkat analysées, cette différence devient significative d'après le test statistique effectué. En effet, le protocole incluant une dénaturation donne des identifications protéiques et peptidiques significativement supérieures au protocole avec une simple digestion.

Néanmoins, ces protocoles ont été étudiés en tube et la dénaturation par la chaleur des protéines dans les plaques micro-puits utilisées dans ce projet (présentées en seconde partie du manuscrit) favorise l'évaporation et la perte des échantillons. Le protocole avec une digestion seule est donc préféré.

ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

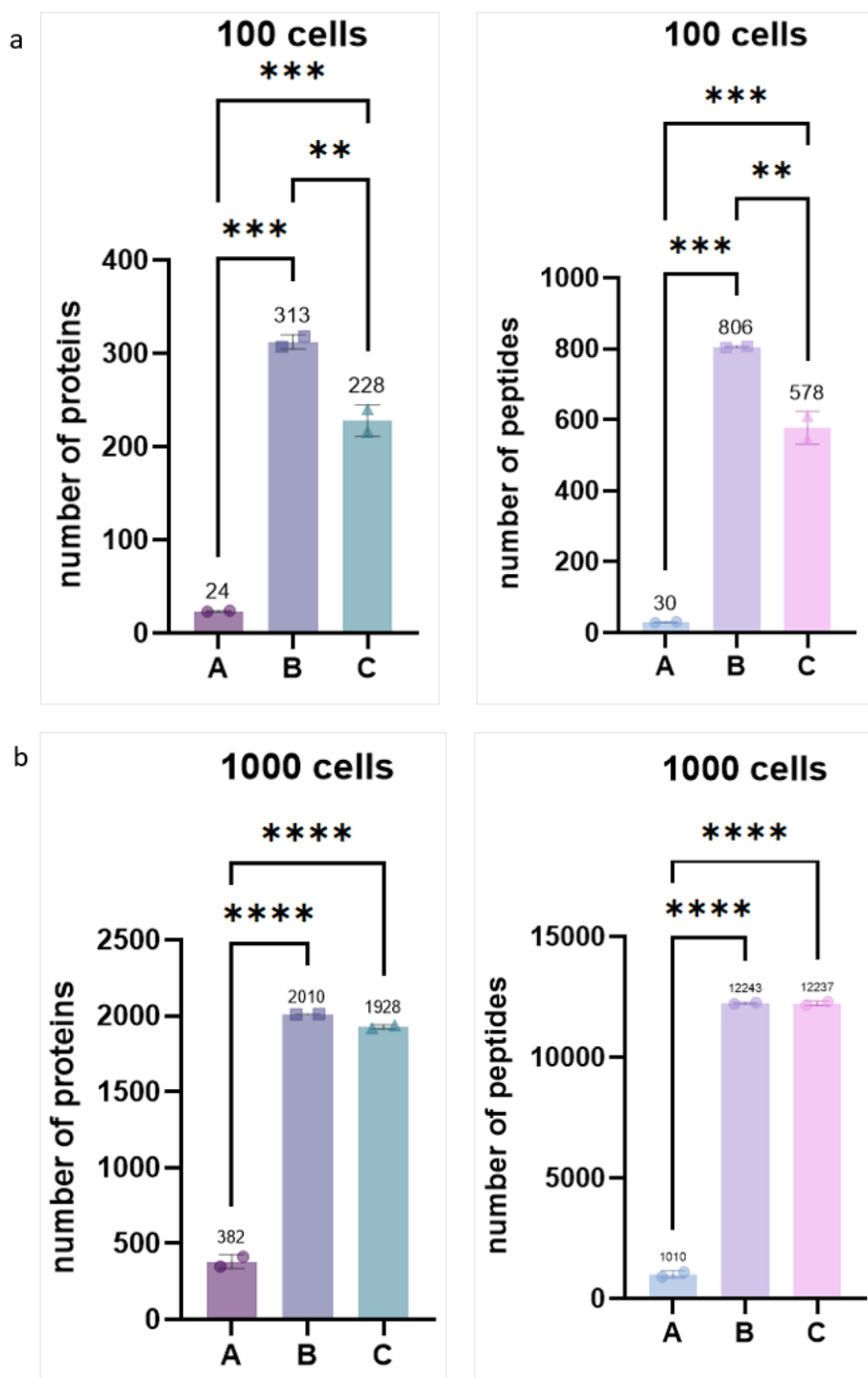


Figure 67 : Nombre de protéines et peptides identifiés dans des échantillons de 100 ou 1000 cellules Jurkat après une lyse thermique selon 3 protocoles : A = dénaturation + réduction + alkylation + digestion, B = dénaturation + digestion, C = digestion. Les médianes des 2 réplicats sont annotées au-dessus des barres d'intervalles de confiance à 95%. Les \* indiquent une p-value < 0.05 obtenue par un Anova à un facteur (avec \* < 0.05, \*\* < 0.005, \*\*\* < 0.0005).

ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

## 5.C Étude de l'étape de digestion

### 5.C.1. Un temps de digestion adapté aux petites quantités

La rapidité et l'efficacité de digestion sont d'un intérêt primordial pour améliorer le rendement et le débit des analyses en protéomique. L'étape de digestion à la trypsine, nécessaire pour créer des peptides, est généralement réalisée pendant plusieurs heures (entre 12 à 16h) sur des échantillons classiques (des millions de cellules digérées par trypsine, réaction dans un tube). Dans la communauté scientifique, les chercheurs essaient de simplifier les protocoles de digestion. Les kits de préparation commerciaux comme le kit iST pour la préparation des protéines (Preomics) et l'addition de nouveaux surfactants anioniques comme le Rapigest SF (Waters) pour faciliter la dénaturation protéique ont réussi à diminuer ce temps à 1 à 3 heures seulement. Il existe aujourd'hui des systèmes microfluidiques qui permettent grâce à des molécules d'endoprotéases attachées sur une puce microfluidique de diminuer le temps de cette réaction à quelques minutes<sup>204</sup>. D'autres systèmes utilisent des cycles d'ultra sonication afin d'accélérer ce processus en solution<sup>205</sup>.

Dans ce projet, l'étude du temps de digestion avec le cocktail de trypsine – Lys C a été effectuée en solution et l'efficacité de la digestion a été vérifiée par le nombre d'identifications obtenues entre 15 minutes et 3 heures de réaction (Figure 68).

#### *Méthode*

Cette étude a été faite pour deux quantités (1 ng ou 20 ng) d'extraits protéiques issus de cellules humaines K562 (Promega) ou de levure *Saccharomyces cerevisiae* (Promega). Les protéines ont été digérées avec un ratio 2:1 (E/S). Les peptides ont été analysés sur un Tims TOF Pro et les données sont présentées sur la Figure 68.

#### *Résultats*

Pour 20 ng de protéines humaines et de levure, environ 1500 protéines et 7000 peptides humains ont été identifiés et environ 1300 protéines et 8000 peptides de levure ont été identifiés, respectivement. Un gain de 100 protéines et 1000 peptides a été observé lorsque le temps de digestion passait de 15min à 3 heures.

Pour 1ng de protéines humaines, environ 300 protéines, 500 peptides ont été identifiés. Un gain de 50 protéines et 200 peptides a été observé lorsque le temps de digestion passait de 15min à 3 heures. Pour 1ng de protéines de levure, la différence était encore plus visible avec l'augmentation du temps de digestion. En 15 min de digestion, 319 protéines et 1575 peptides ont été identifiés. Ces nombres ont augmenté après 3 heures de digestion à 512 protéines et 1724 peptides, observant un gain de 200 protéines et 200 peptides grâce à l'augmentation du temps de digestion.

D'après ces résultats, l'efficacité de digestion augmente avec le temps de digestion jusqu'à 3 heures. Dans ce projet, nous avons conservé un temps de 3 heures et pas un temps classique de 16 heures afin de limiter l'évaporation et l'adsorption due à un long temps de contact entre les protéines/peptides et le support.

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

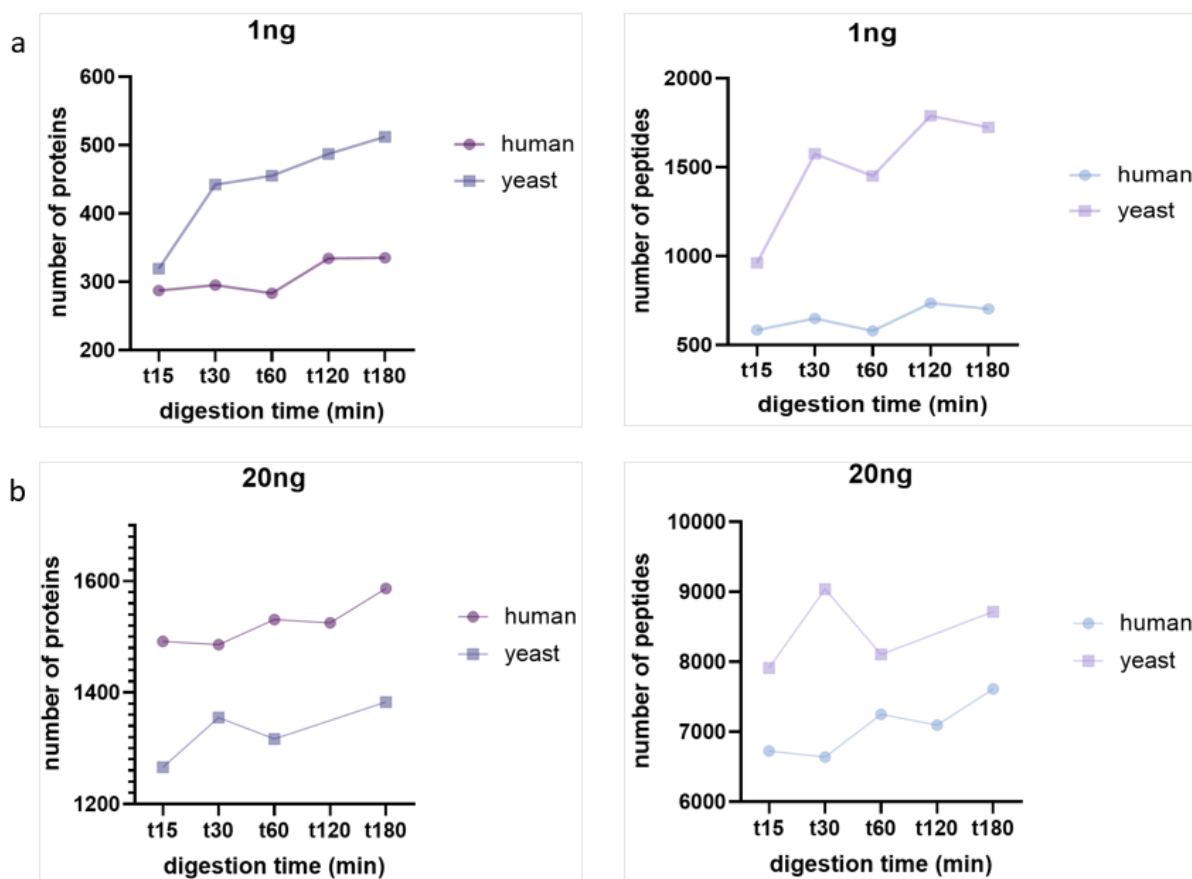


Figure 68 : Nombre de protéines et peptides identifiés à partir d'extraits humains ou de levure à différents temps de digestion en minutes (avec t15 = 15minutes, t30 = 30minutes, etc.).

### 5.C.2. Un rapport enzyme/protéine mieux adapté aux petites quantités

Deux ratios massiques d'enzyme/substrat ont été comparés et étudiés pour différents échantillons et différentes concentrations. Un ratio de trypsine-Lys C/substrat (1:20 (E/S)) est comparé à un ratio où la concentration de trypsine-Lys C est doublée par rapport à la quantité de protéines (2:1 (E/S)).

#### Méthode

Sur une plaque micro-puits en PDMS (support présenté dans la seconde partie du manuscrit), deux quantités (50ng et 10ng) de protéines extraites de cellules de levure (*Saccharomyces cerevisiae*) ont été digérées par la trypsine-Lys C avec par les deux ratios enzyme/protéines décrits précédemment. L'équivalent de 10ng et 2ng est analysé, respectivement sur un Tims TOF Pro. La médiane des 3 réplicats est annotée au-dessus des intervalles de confiance à 95% pour chaque condition (Figure 69).

#### Résultats

Pour 10ng de protéines analysées, les deux rapports enzyme/substrat utilisés ne montrent aucune différence significative selon le test statistique effectué. Cependant, il est à noter une augmentation du nombre moyen d'identifications protéiques, peptidiques et de PSM lorsque le ratio 2:1 (E/S) est utilisé.

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

Pour 2 ng de protéines analysées, seulement 37 protéines, 51 peptides et 53 PSM sont identifiés avec un ratio 1:20 (E/S) alors que 395 protéines, 1114 peptides et 1192 PSM sont identifiés avec un ratio 2:1 (E/S). En effet, le test statistique montre une différence significative entre les deux ratios d'enzyme/substrat en faveur du ratio 2:1 (E/S).

Un ratio d'enzyme/substrat où la concentration d'enzyme est supérieure à la concentration de protéines à digérer est plus efficace pour la digestion de petites quantités de protéines digérées.

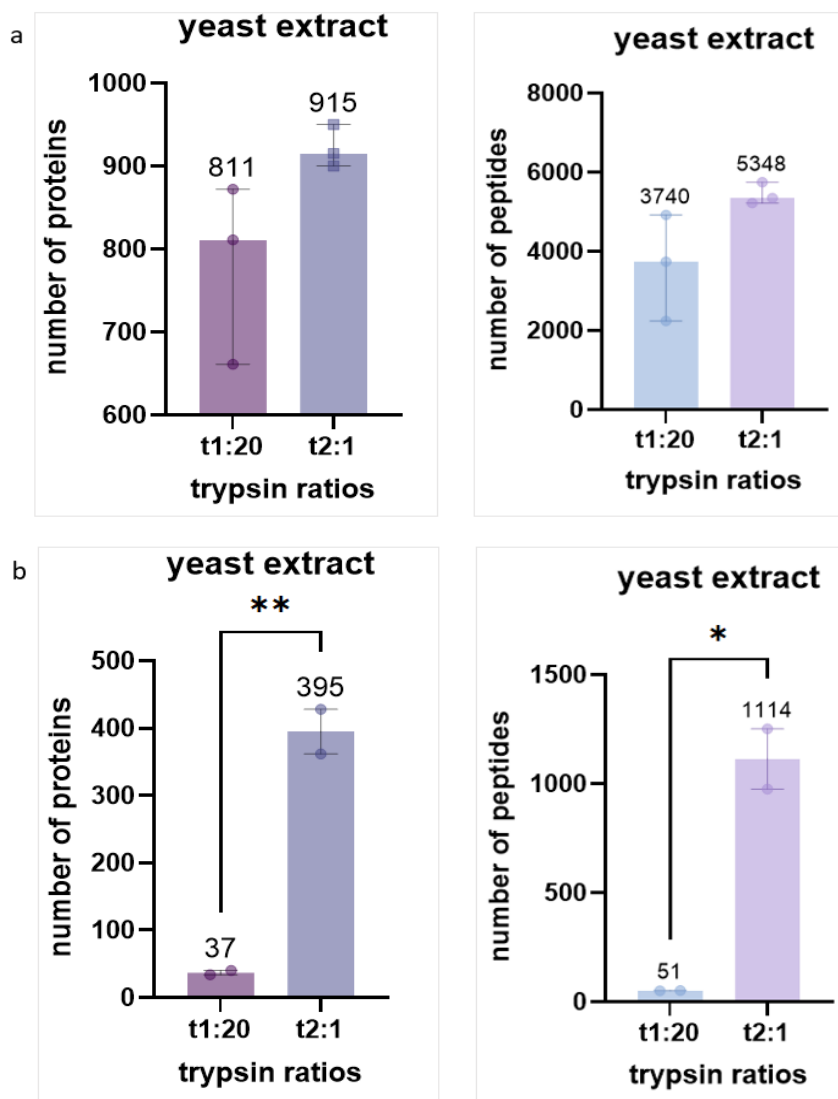


Figure 69 : Nombre de protéines et peptides identifiés dans (a) 10ng et (b) 2ng analysés d'extraits protéiques de levure digérés par deux ratios de trypsine – Lys C/substrat. La médiane est annotée au-dessus des barres d'intervalles de confiance à 95%. Les \* indiquent une p-value < 0.05 selon un t-test non apparié (avec \* < 0.05, \*\* < 0.005, \*\*\* < 0.0005).

ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

### 5.C.1. Du DMSO pour limiter l'évaporation

Le DMSO ou diméthylsulfoxyde est un solvant polaire aprotique, de formule  $C_2H_6OS$  et miscible dans l'eau (Figure 70). Dilué de 5 à 20 %, il est utilisé comme agent cryoprotecteur lors de la congélation de cellules. Pendant longtemps, le DMSO était considéré comme non compatible à la spectrométrie de masse jusqu'à récemment où son ajout a été évalué en tant qu'additif à la phase mobile. Il a été trouvé que l'ajout de petites concentrations (3%) de DMSO améliorerait l'identification peptidique et la quantification protéique<sup>206</sup>. Aujourd'hui, il est utilisé en protéomique de la cellule unique pour la lyse cellulaire<sup>176</sup>. Dans ce projet, il est utilisé pour ses propriétés visqueuses qui le rendent plus stable pendant les phases de chaleur et résistant à l'évaporation. De plus, il permet d'augmenter l'efficacité de la désolvatation lors de l'électro nébulisation<sup>181</sup>.

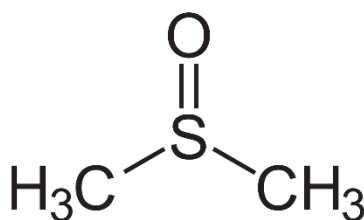


Figure 70 : Formule brute du DMSO.

#### Méthode

Des aliquots d'1µl de digestats peptidiques ou d'extraits protéiques de levure *Saccharomyces cerevisiae* à 50ng/µl ont été dilués dans de l'eau (condition contrôle = ctrl), dans 10% de DMSO initial (condition DMSO 10, l'équivalent de 2% en concentration finale) ou dans 30% de DMSO initiale (condition DMSO 30, l'équivalent de 5% en concentration finale). Les échantillons ont été déposés sur une plaque micro-puits PDMS 10µl traitée au PVP (support et traitement présentés en seconde partie du manuscrit). Les protéines ont été digérées avec 500nl de trypsine-Lys C avec un ratio 2:1 (E/S). Les peptides ont été analysés sur un Q Exactive HF. La médiane des 3 réplicats est annotée au-dessus des barres d'intervalles de confiance à 95% de chaque condition sur la Figure 71.

#### Résultats

Pour les mélanges de peptides de levure déjà digérés, environ 1850 protéines ont été identifiées dans toutes les conditions. Aucune différence n'a été observée après un test statistique entre les conditions de dilution. L'ajout de DMSO ne fait aucune différence pour la récupération peptidique et donc l'identification des protéines pour ces extraits peptidiques. D'autre part, un Anova à un facteur montre des différences significatives entre les 3 conditions de dilution des échantillons. En effet, après 3 heures de digestion, l'apport du DMSO à 10% ou à 30% permet une identification protéique meilleure qu'en l'absence de DMSO. L'ajout du DMSO à 30% améliore plus le nombre de protéines identifiées que l'ajout de DMSO à 10% ou l'absence de celui-ci.

L'apport du DMSO dans l'analyse de petites quantités d'échantillons permet de limiter l'évaporation pendant les étapes de digestion (3h à 37°C) et d'améliorer les identifications peptidiques et protéiques.

ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

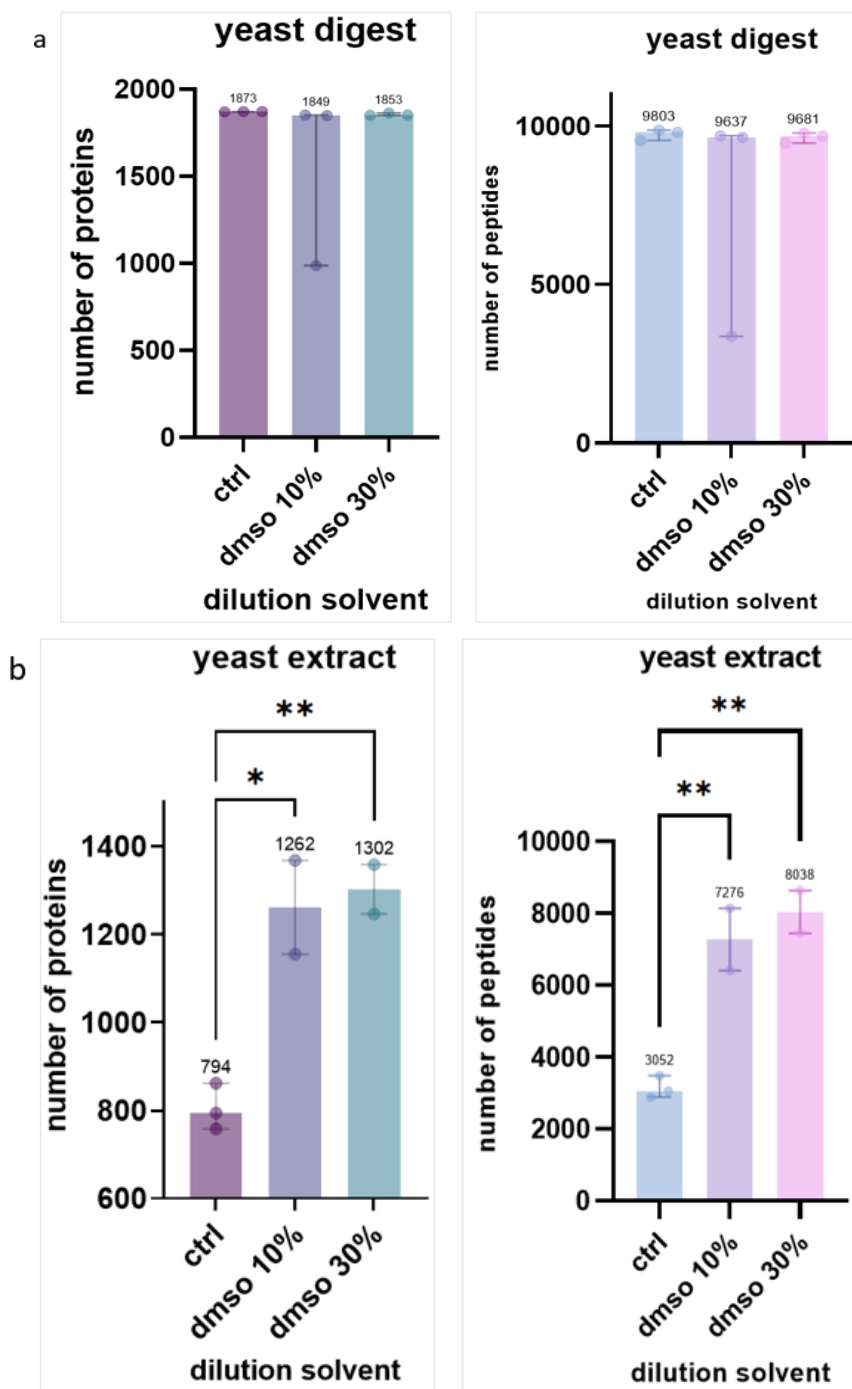


Figure 71 : Nombre de protéines et peptides identifiés (a) depuis un digest de levures, ou (b) après digestion d'un extrait de levure, tous deux dilués dans différentes conditions (avec dilution dans le DMSO). Les médianes des 3 réplicats sont annotées au-dessus des barres d'intervalles de confiance à 95%. Les \* indiquent une p-value < 0.05 obtenue par un ANOVA à un facteur (avec \* < 0.05, \*\* < 0.005, \*\*\* < 0.0005).

ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

## 5.D Optimisation des conditions d'analyses en LC-MS

### 5.D.1. Colonnes et configurations du système de chromatographie

Comme discuté dans l'introduction, les paramètres d'une colonne (diamètre interne, colonne, débit) sont très importants pour améliorer la résolution chromatographique et donc la séparation avant la détection en spectrométrie de masse de mélanges complexes. Une étude des configurations existantes a été menée pour l'amélioration de l'identification protéique.

Sur un instrument Orbitrap (Q Exactive HF), une comparaison de colonnes avec deux dimensions de diamètres internes a été effectuée. Deux colonnes ont été comparées :

- Une colonne nanoEase (Waters) en phase BEH+C18 (avec des particules de 1.7µm de diamètre) de 75µm d.i. x 25cm, une dimension standard en nano LC-MS (volume total : 1.10µl, débit standard utilisé 300nl/min)
- Une colonne Acclaim PepMap (Thermo) en phase C18 (avec des particules de 2µm de diamètre) de 50µm d.i. x 50cm (volume total : 0.98µl, débit adapté utilisé 115nl/min).

Deux configurations sont testées, une configuration en préconcentration (PC) avec une pré-colonne, utilisée pour le dessalage des échantillons, reliée à la colonne de séparation et une configuration en directe injection (DI) dans laquelle la colonne de séparation est directement liée à l'injecteur sans pré-colonne de dessalage (voir introduction).

La colonne de 75µm d.i. est testée en conditions avec une pré-colonne de 100µm de diamètre interne x 2cm. La colonne de 50µm de d.i. est testée dans les deux configurations, en PC avec une pré-colonne de 75µm de diamètre interne x 2cm ou en DI sans pré-colonne.

La moyenne du nombre de protéines identifiées pour 3 ou 4 réplicats est représentée sur une courbe pour des quantités croissantes de digestats peptidiques de cellules HeLa (Figure 72). Plus la quantité d'échantillon injectée au départ augmente, plus le nombre de protéines identifiées augmente. Un plateau semble être atteint après 50ng de peptides analysés avec une légère augmentation entre 100 et 200ng pour toutes les conditions.

Ainsi, sur un Q Exactive HF, pour l'exemple standard de la colonne de 75µm d.i. en préconcentration (courbe bleue, symboles ronds), en moyenne 1200 protéines sont identifiées pour 1.6ng de digestat issu de cellules HeLa alors que près de 2500 protéines sont identifiées pour 100ng du même échantillon.

Quand la colonne de 50µm d.i. est branchée en préconcentration (courbe rouge, symboles carrés), les identifications sont moins bonnes que pour une colonne standard. En effet, pour 1.6ng de digestat peptidique, 1000 protéines sont identifiées et pour 100ng ou 200ng, 2000 protéines sont identifiées. Cela peut s'expliquer par la différence de phase stationnaire qui est une C18 dans la colonne de 50µm d.i. et BEH+C18 dans la colonne de 75µm d.i. L'utilisation de particules BEH (pour « *Bridged Ethylene Hybrid* ») liées à une phase C18 dans les colonnes nanoEase offre une plus grande stabilité sur une large gamme de pH et une meilleure rétention des peptides et ainsi une meilleure séparation chromatographique qu'une phase simple en C18.

Cependant, si la colonne de 50µm de diamètre interne est branchée en directe injection (courbe verte, symboles triangles), le nombre d'identifications augmente par rapport aux conditions standard surtout pour de très petites quantités d'échantillons analysés. Plus de



ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

2000 protéines sont identifiées pour 1.6ng de digestat et près de 3000 protéines sont identifiées pour 200ng. Cela suggère que la configuration en préconcentration induit une perte de matériel.

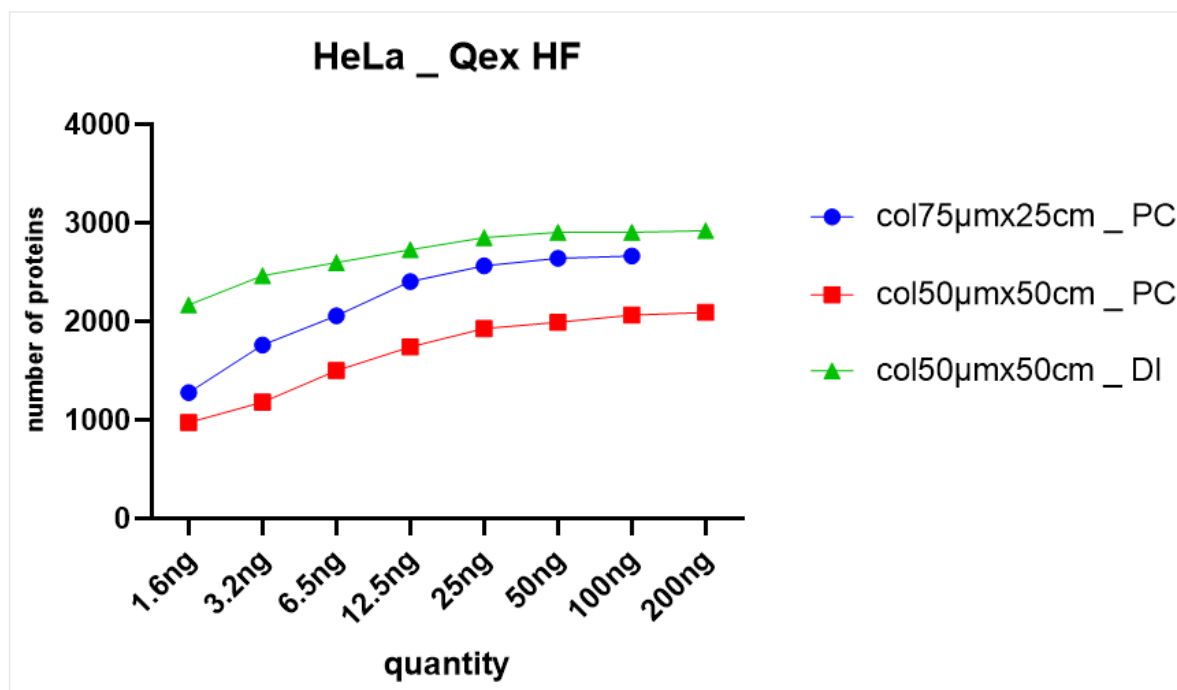


Figure 72 : Représentation graphique du nombre d'identifications protéiques à différentes concentrations de digests de HeLa avec différentes colonnes et configurations.

Nous avons aussi comparé la largeur des pics chromatographiques sur l'analyse d'un digestat de la protéine BSA (Figure 73). L'extraction de l'ion 588.3190 de la BSA a été réalisée pour deux colonnes en configurations différentes. La comparaison des deux pics montre une largeur de pic à mi-hauteur plus importante avec un effet de *tailing* pour l'analyse effectuée avec la colonne de 75µm d.i. en configuration PC que celle obtenue avec la colonne de 50 µm d.i. en configuration DI indiquant une meilleure séparation qui était attendue et surtout une meilleure résolution des pics avec la colonne de plus petit diamètre interne.

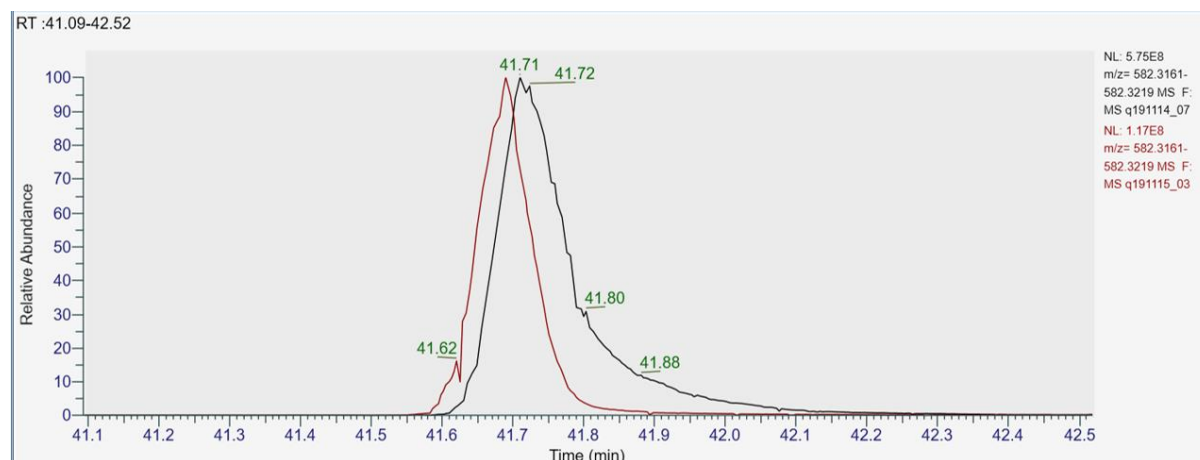


Figure 73 : Comparaison de la largeur des pics après extraction de l'ion « 588.3190 » de la BSA en pic de base après séparation sur une colonne de 75µm de diamètre interne (noir) ou sur une colonne de 50µm de diamètre interne (rouge).

ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

### 5.D.2. Fabrication de colonnes "homemade"

Le diamètre de colonne est un paramètre crucial pour augmenter la sensibilité de l'analyse LC-MS. Nous avons donc cherché à fabriquer des colonnes de plus petit diamètre interne que celles trouvées dans le commerce (<50 µm).

Dans un premier temps, nous avons cherché à fabriquer des colonnes avec des dimensions classiques (75µm d.i. x 15.7cm, volume : 0.69µl, débit de 300nl/min), appelées colonnes "homemade". La fabrication de la colonne est décrite dans la partie matériel et méthodes. La colonne *homemade* a été connectée comme les colonnes commerciales au Q Exactive HF avec un capillaire et un tip externe.

#### Résultats

Une mesure de la pression de la colonne en changeant le débit avec un gradient en fonction du temps a été effectuée (Tableau 4). Une augmentation stable de la pression de la pompe quand le débit passe de 50 nl/min à 300nl/min en 20min et un maintien de celle-ci quand le débit est maintenu à 300nl/min sont attendus. Cette observation sur deux réplicats indique une phase homogène dans la colonne (Figure 74). Cependant, un léger épaulement peut être observé à 12min sur le deuxième réplikat indiquant un reconditionnement de la phase à l'intérieur de la colonne.

TIME (MIN)	DEBIT (NL/MIN)
0	50
5	100
10	150
15	200
20	300
45	300

Tableau 4 : Gradient d'augmentation du débit en fonction du temps.

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

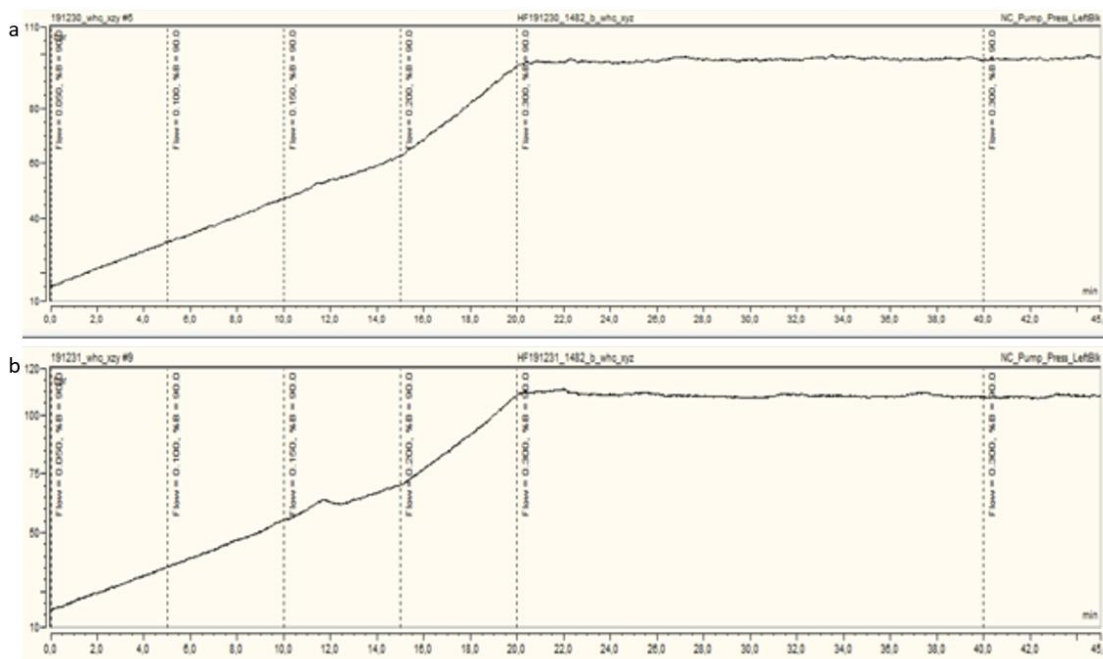


Figure 74 : Représentation du changement de pression de la pompe NC reflétant la pression dans la colonne lors de l'augmentation croissante du débit dans deux répliqués de blancs d'échantillons (a et b). Les marques en pointillés indiquent les changements de débit décrits dans le tableau précédent.

Il est important de surveiller la pression de la colonne. Si la pression est trop élevée, la phase peut s'échapper de la colonne par dégradation physique du fritté. Par ailleurs, une phase mal conditionnée va induire des changements de pression et des épaulements lors de l'augmentation du débit et la pression va mettre plus de temps avant de se stabiliser (Figure 75).

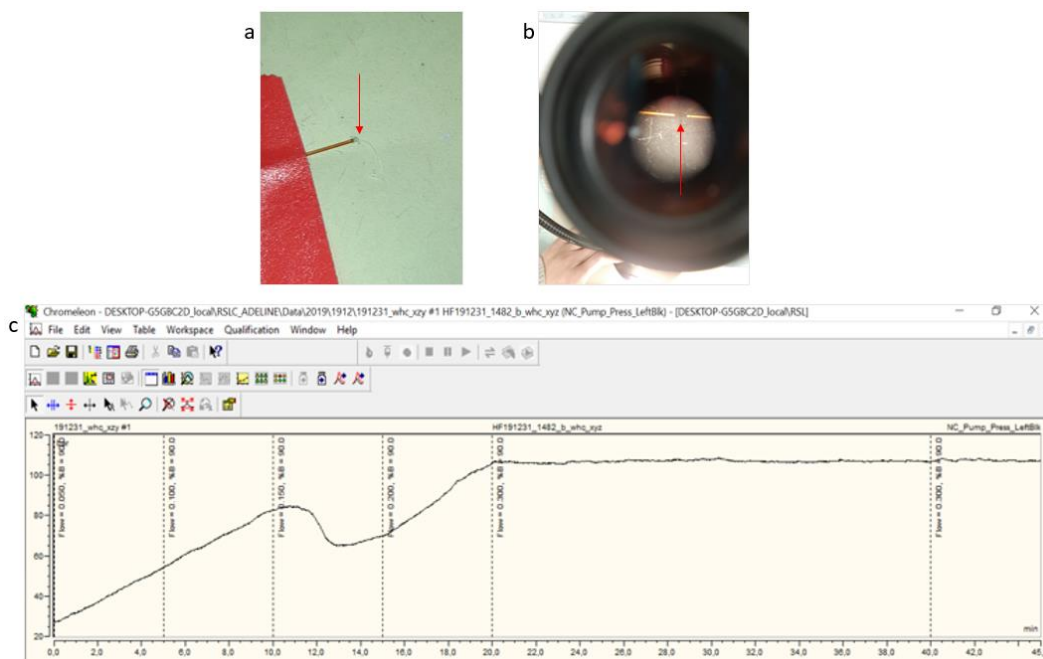


Figure 75 : Problèmes rencontrés avec les colonnes *homemade* (a) phase s'échappant d'une colonne fabriquée en laboratoire. (b) observation de « trous » dans la phase créant (c) des épaulements dans la

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

pression lors de la montée de débit qu'il est important de surveiller. Les flèches rouges indiquent les problèmes discutés.

Ensuite, un débit à 220nl/min a été maintenu et un gradient d'élution à l'acétonitrile a été appliqué (Tableau 5). Le profil de pression observé est le même que celui observé avec les colonnes commerciales. La pression diminue dramatiquement lorsque le gradient d'élution atteint 90% d'acétonitrile et remonte lorsque la quantité d'acétonitrile baisse pendant l'équilibrage de la colonne (Figure 76).

TIME (MIN)	% ACETONITRILE
0	1
6	2
65	40
66	90
83	90
84	1
112	1

Tableau 5 : Gradient d'élution de 112 minutes et % d'acétonitrile utilisée pour l'élution.

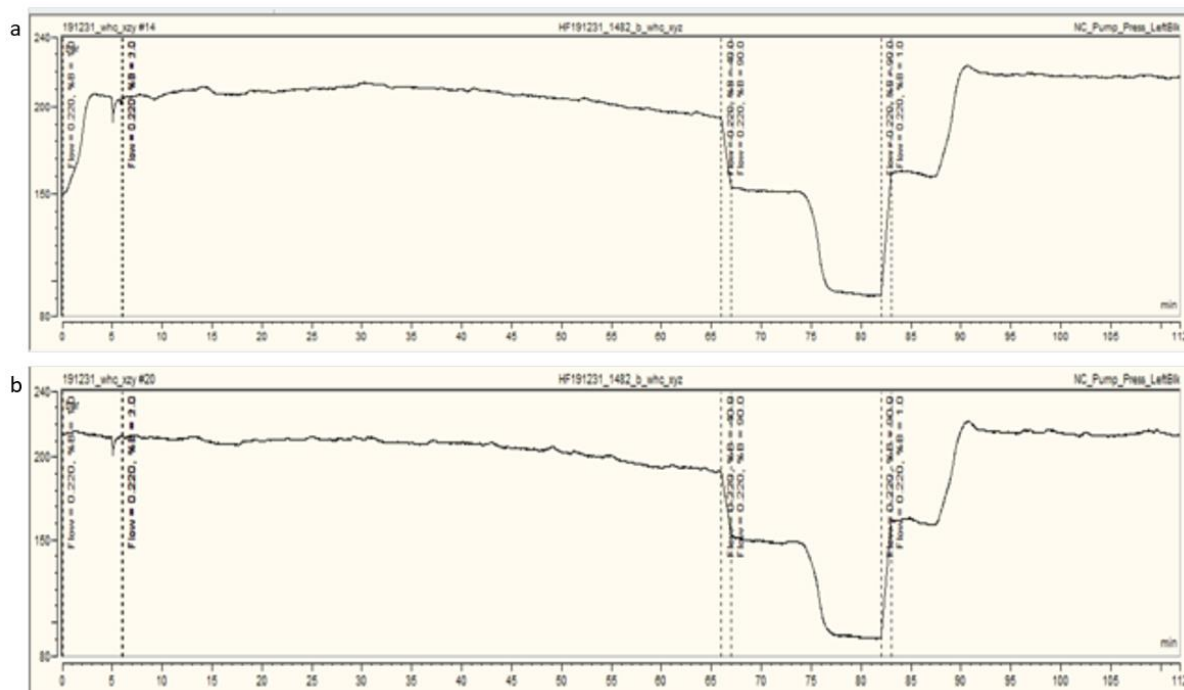


Figure 76 : Représentation du profil de pression de colonne de séparation lors d'un gradient de 112 minutes dans deux répliqués d'échantillons témoins (a et b). Les marques en pointillés indiquent les changements de gradient décrits dans le tableau précédent.

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

Une gamme de concentration de digestats extraits de cellules HeLa a été injectée dans la colonne *homemade*. Cette colonne a été testée dans plusieurs configurations en PC et en DI et sur deux spectromètres de masse, le Q Exactive HF (Q Ex HF) et l'Eclipse (en PC uniquement). La différence entre les deux configurations en PC sur le Q Exactive HF et l'Eclipse est l'emplacement de la colonne par rapport à la source nano ESI. Sur le Q Exactive HF, le four étant dans l'HPLC, la sortie de la colonne est reliée au tip par un capillaire de 30 $\mu$ m de diamètre interne et 20 cm de longueur. Sur l'Eclipse, grâce à un four connecté à la source (Sonation), la sortie de colonne est directement reliée au tip ce qui évite les volumes morts post-colonnes. La moyenne du nombre de protéines identifiées pour 2 réplicats est représentée sur une courbe pour des quantités croissantes de digestats peptidiques de cellules HeLa (Figure 77).

La colonne homemade branchée en PC sur le Q Exactive HF permet l'identification en moyenne de 500 protéines pour 1.6ng de digestat de HeLa et 3200 protéines pour 200ng de digestat de HeLa (courbe bleue, symboles ronds). Pour les grandes quantités d'échantillons (>50ng), elle identifie plus de protéines que la colonne commerciale en PC (courbe grise) mais pour les petites quantités d'échantillons, cette colonne branchée en PC identifie deux fois moins de protéines que la colonne commerciale dans la même configuration et sur le même appareil.

La colonne homemade branchée en PC sur l'Eclipse permet l'identification en moyenne de 1100 protéines pour 1.6ng de digestat de HeLa et 3300 protéines pour 100ng de digestat de HeLa (courbe verte, symboles triangles). Le changement d'appareil et la suppression du volume mort post colonne permettent une amélioration des identifications protéiques pour toutes les quantités de HeLa analysées.

Enfin, la colonne homemade branchée en DI sur le Q Exactive HF permet l'identification en moyenne de 2000 protéines pour 1.6ng de digestat et 3200-3300 protéines pour 100 à 200 ng de digestat de HeLa (courbe rouge, symboles carrés). La configuration en DI permet un gain d'identifications sur toute la gamme de digestats analysés comme pour les colonnes standard. D'ailleurs, en comparant avec une colonne commerciale, grâce à cette configuration en DI à diamètre interne égal, la colonne homemade permet un nombre d'identifications supérieur aux identifications obtenues dans la colonne commerciale. Elle se compare à la colonne commerciale de 50 $\mu$ m de diamètre interne en termes de nombre de protéines identifiées pour les petites quantités analysées.

En conclusion, la colonne homemade permet d'identifier plusieurs centaines de protéines pour de très petites quantités de digestats de HeLa analysées (de l'ordre du nanogramme). Avec une configuration en DI, elle permet des identifications comparables à une colonne commerciale.

Ces premiers essais de fabrication de colonnes se sont poursuivis par des essais de fabrication de colonne de diamètre de 20 $\mu$ m afin d'augmenter la sensibilité. Les données ne seront pas présentées ici mais le manque de reproductibilité constaté de ces colonnes *homemade* de très petit diamètre interne est discuté dans la discussion.

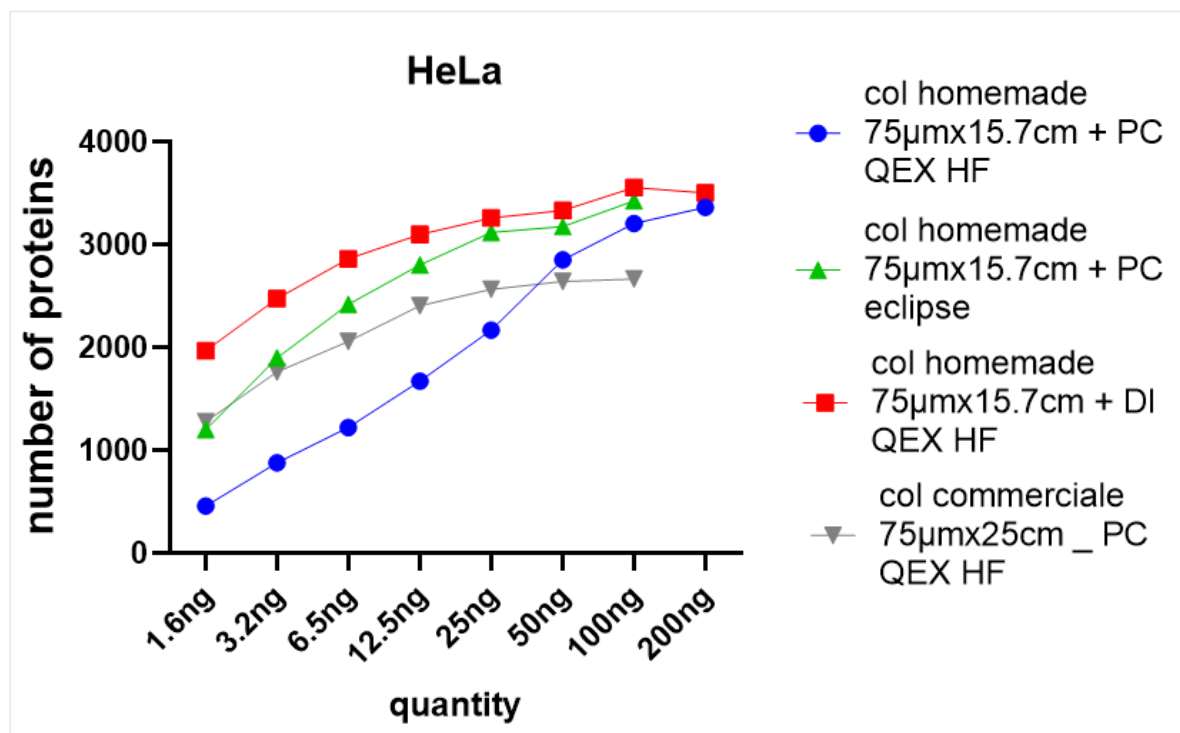


Figure 77 : Représentation graphique du nombre d'identifications protéiques à différentes concentrations de digests de HeLa avec une colonne « *homemade* » et une colonne commerciale en différentes configurations avec deux appareils.

### 5.D.3. Essais sur le Tims TOF Pro

Le Tims TOF Pro (Bruker) est un spectromètre de masse de dernière génération dont la spécificité est de faire une analyse en 2 dimensions en séparant les ions en fonction de la mobilité ionique en plus des masses sur charges des ions (voir introduction).

#### Méthode

Depuis sa réception au laboratoire, il a été utilisé pour les expériences dans ce projet. Les colonnes compatibles avec la nano ESI du Tims TOF Pro sont les colonnes Aurora (Ion Opticks). Elles sont disponibles en phase C18 avec 75µm de diamètre interne et 25cm de long et possèdent un tip intégré qui permet d'éviter les volumes morts après la séparation (Figure 78). Un four externe permet de maintenir la colonne à une température stable de 50°C. Selon les résultats obtenus précédemment, ces colonnes ont été branchées en configuration DI directement.



Figure 78 : Photographie d'une colonne Aurora (Ion Opticks) de 75µm d.i. et 25cm de long, avec à gauche la tête de la colonne (adaptée aux vis nano Viper) et à droite le tip intégré protégé par une capuche qui s'intègre directement à la source nano ESI du Tims TOF Pro (Bruker).

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

### Résultats

Différentes méthodes d'acquisition spécifiques du spectromètre de masse Tims TOF Pro ont été comparées.

Une méthode *low amount* dédiée à l'analyse des petites quantités a été comparée à la méthode standard. La méthode *low amount* (appelée *low* par la suite) est caractérisée par des paramètres Tims modifiés, c'est-à-dire, une rampe entre 0.7V et 1.1V et un temps d'accumulation plus long que la méthode standard (appelée *std*) à 188ms au lieu de 100ms. Les paramètres de la méthode PASEF restent inchangés avec une gamme de m/z entre 100 et 1700 en mode positif.

En parallèle de la comparaison des méthodes de MS, deux gradients d'élution ont été testés, un gradient de 30min (run = 1h) et un gradient de 100min (run = 2h) (Tableau 6). Pour les grandes quantités de protéines (200ng et plus), il est habituel d'utiliser des gradients de 2 ou 3 heures pour mieux séparer les mélanges complexes en répartissant les espèces sur des fenêtres plus larges et gagner en nombre d'identification protéique. Il est donc important de vérifier l'influence du temps de séparation pour de petites quantités d'échantillons (1 et 5ng de digestat de HeLa).

TIME (MIN)	% ACETONITRILE	TIME (MIN)	% ACETONITRILE
0	5	0	5
5	5	5	5
33	30	90	25
34	90	105	40
39	90	106	90
40	5	111	90
60	5	112	5
		130	5

Tableau 6 : Gradient d'élution de 30 ou 100 minutes (runs 60min ou 130min respectivement) et % d'acétonitrile utilisée pour l'élution.

## ETUDES ET RESULTATS : Adaptation du protocole à l'analyse de petites quantités d'échantillons

Le nombre de protéines identifiées pour 1ng et 5ng de digestat analysés, est montré sur la Figure 79.

Avec la méthode standard (std), le temps de gradient allongé diminue le nombre d'identifications protéiques de 489 à 388 pour 1ng de digestat et de 1517 à 1258 pour 5ng de digestat. Avec la méthode conçue pour les petites quantités (*low*), le temps de gradient allongé augmente légèrement le nombre d'identifications protéiques de 561 à 640 pour 1ng de digestat et de 1292 à 1428 pour 5ng de digestat. Pour une durée de gradient donnée, la méthode *low* donne de meilleures identifications pour 1ng d'échantillon analysé.

Cependant, d'après un Anova à un facteur avec de multiples comparaisons, aucune différence significative n'est observée entre toutes ces méthodes et tous ces gradients. La méthode *low amount* a donc été adoptée ainsi qu'un gradient de 30min (1h de run) afin d'augmenter le débit des analyses.

Avec une configuration en DI et la méthode *low amount* spécifique des petites quantités, une gamme de digestat a été analysée sur plusieurs mois afin de vérifier la sensibilité du spectromètre de masse. Le débit a été diminué à 220nl/min pour un gradient de 30min (run = 1heure), afin d'améliorer les identifications protéiques (données non présentées).

La figure montre une gamme s'étendant de 40pg à 200ng de digestat de cellules HeLa obtenue à partir de 50 réplicats (Figure 80). Une bonne reproductibilité est observée entre les réplicats analysés à des temps différents sur plusieurs mois. Le nombre d'identifications obtenu est similaire à ceux observés pour les configurations présentées dans les paragraphes précédents une colonne de 75µm de diamètre interne et 15cm de longueur (comme la colonne *homemade* qui avait les mêmes dimensions).

### *Bilan*

En fonction de nos observations, une configuration en directe injection avec une colonne de petit diamètre interne (50 µm) et un gradient de 30min s'est avérée la plus performante pour l'analyse de petites quantités d'échantillons.

Cependant, le Tims TOF Pro avec une colonne Aurora 75µm d.i. x 25cm en configuration en directe injection a été utilisé pour l'optimisation de la plupart des expériences décrites dans ce projet.



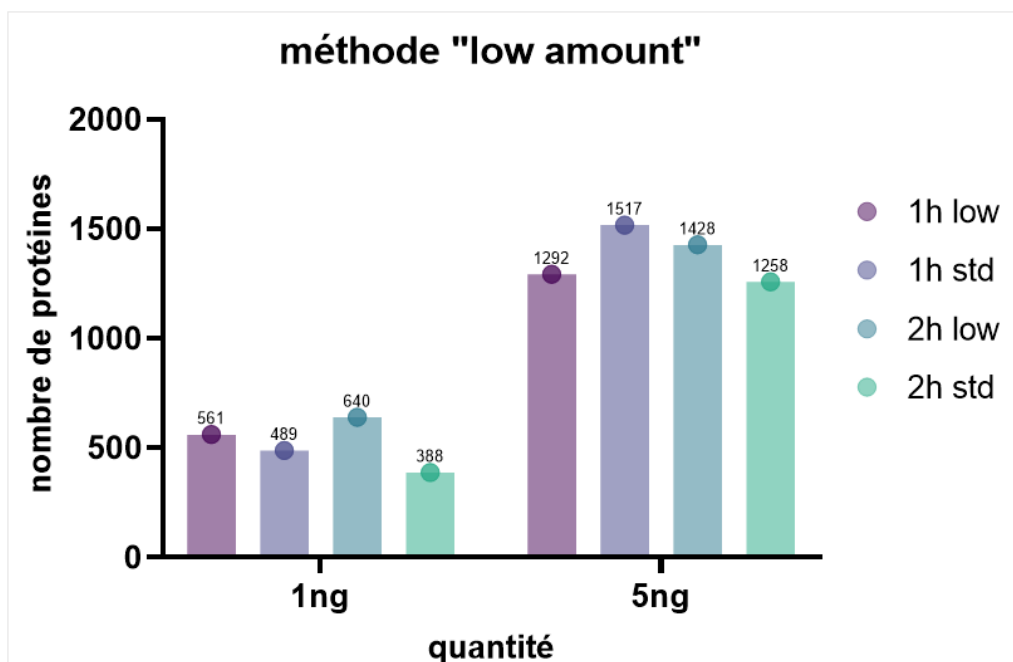


Figure 79 : Représentation graphique du nombre d'identifications protéiques pour deux quantités d'un digestat peptidique issu de cellules HeLa analysé en LC-MS/MS sur un Tims TOF Pro avec une colonne Aurora 75µm d.i. x25cm par 2 méthodes MS qui diffèrent par leur temps d'accumulation en TIMS (std = 100ms, low = 188ms).

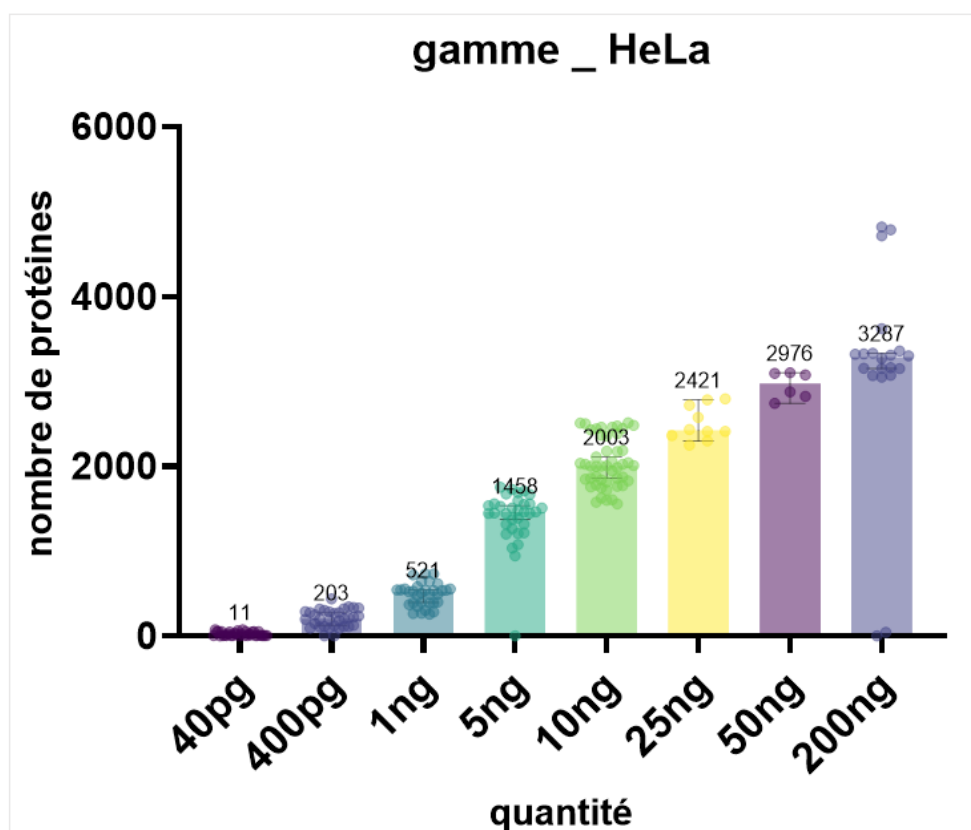


Figure 80 : Gamme de digestats extraits de cellules HeLa analysée en LC-MS/MS sur le Tims-TOF Pro avec une colonne Aurora de 75µm d.i. x 25cm, avec un gradient de séparation de 30min sur plusieurs mois.

## Chapitre 6 Importance du support

<b>Chapitre 6 Importance du support</b> .....	<b>95</b>
<b>6.A Influence de la géométrie et du matériau des supports</b> .....	<b>96</b>
6.A.1. Miniaturiser la surface de contact sur géométrie plane : évaluation de plaques MALDI	96
(a) Application sur une protéine : l'hémoglobine .....	96
(b) Application sur des extraits protéiques .....	98
(c) Évaluation des plaques MALDI modifiées : les <i>AnchorChip</i> .....	98
6.A.2. Miniaturisation de la surface de contact sur géométrie en puits .....	101
(a) Micro-fabrication de plaques micro-puits en PDMS : 3 designs .....	101
(b) Impression 3D de plaques micro-puits en polypropylène .....	104
<b>6.B Comparaison du support micro-puits en PDMS avec les plaques commerciales</b> .....	<b>106</b>
<b>6.C Compatibilité de la plaque micro-puits avec le trieur de cellules</b> .....	<b>108</b>
<b>6.D Apport d'un traitement de surface</b> .....	<b>109</b>
6.D.1. Traitement de surface avec du polyvinylpyrrolidone (PVP) .....	109
(a) Qu'est-ce que le Polyvinylpyrrolidone (PVP) ? .....	109
(b) Mise en place du traitement de surface .....	110
(c) Compatibilité du PVP avec la spectrométrie de masse .....	112
(d) Application du traitement de surface au PVP sur un mélange complexe	113
6.D.2. Le PVP est-il le meilleur traitement de surface ? .....	118
(a) Optimisation du setup expérimental .....	118
(b) Évaluation des différents traitements de surface .....	122
Polydopamine tamponnée (pH = 8.8) sans oxydation .....	124
Polydopamine tamponnée (pH = 8.8) avec oxydation préliminaire .....	124
Polydopamine tamponnée (pH = 5) avec oxydation .....	125
Traitement par le PVA .....	126
Traitement par le PVP .....	126
Traitement par l'APTES .....	127
Traitement par le $\beta$ - octyl/ .....	127
Traitements par l'acide hyaluronique (HA) .....	128
Bilan .....	128
Cas particulier du PVP .....	129
6.D.3. Le PVP est-il compatible avec la protéolyse des échantillons ? .....	132
6.D.4. Le PVP est-il compatible avec de plus petites quantités de matériel ? .....	133

### 6.A Influence de la géométrie et du matériau des supports

Le support utilisé pour déposer les échantillons est aussi important que le protocole de préparation de ces échantillons. Comme décrit précédemment, les défis de gestion des échantillons en protéomique sont nombreux et le principal est l'adsorption aux surfaces de contact et donc le choix des supports utilisés. En protéomique de la cellule unique, les chercheurs utilisent des plaques multi-puits ou des lames de verre. Chacune de ces solutions a ses avantages et ses inconvénients. Les plaques multi-puits possèdent des puits en profondeur et présentent donc une grande surface de contact favorisant l'adsorption mais l'évaporation peut être plus facilement gérée par un couvercle. Dans le cas des lames en verre, la surface de contact avec l'échantillon est minimale mais il faut un environnement contrôlé afin d'éviter l'évaporation.

#### 6.A.1. Miniaturiser la surface de contact sur géométrie plane : évaluation de plaques MALDI

Tout d'abord, un support commercial sur plaques MALDI a été testé pour l'identification des peptides et des protéines à partir d'extraits de protéines. Les plaques MALDI (comme les plaques OptiTOF utilisées) sont des plaques en métal ou en matériau conducteur inertes, présentant une surface plane. Par défaut, le métal ne présente pas une surface hydrophobe, ce qui favorise l'étalement des gouttes. D'autres plaques de type *AnchorChip* ont reçu un traitement de surface particulier. Leur surface est traitée pour être hydrophobe et les emplacements des échantillons sont identifiés par des ancrs hydrophiles. Elles ont donc été choisies car elles présentent une surface de contact minimale avec l'échantillon.

L'utilisation de ces plaques est comparée à celle d'un tube *low binding* pour la récupération d'échantillons plus ou moins complexes.

##### (a) Application sur une protéine : l'hémoglobine

###### *Méthode*

Après dépôt de deux quantités (200ng et 50ng) d'un digestat d'hémoglobine en aliquots de 1µl dans un tube et sur une plaque MALDI (Opti-TOF™, AB Sciex), conditions « tube » et « plaque », les échantillons sont analysés sur un Q Exactive HF. Les différentes sous-unités identifiées sont représentées par leur nombre de peptides (Tableau 7) et leur couverture de séquence dans la Figure 81.

###### *Résultats*

Sur les 8 sous-unités théoriques de l'hémoglobine humaine, 7 sous-unités ont été identifiées avec entre 2 et 24 peptides dans l'ensemble des conditions. Tout d'abord, des peptides théoriques ont été obtenus avec le logiciel MassLynx en digérant *in silico* les séquences des sous-unités de l'hémoglobine humaine par la trypsine en autorisant 2 *miscleavages* et en fixant les limites de m/z détectables par spectrométrie de masse entre 500 et 4000m/z. Deux des sous-unités sont très faiblement détectées : la Zeta et la Theta-1 avec environ 3 et 4 peptides sur 28 et 36 théoriques respectivement. Les autres sous-unités sont détectées avec un nombre plus important de peptides. Le nombre de peptides identifiés entre les conditions « tube » et « plaque » est le même pour 200ng de digestat déposé sauf pour les sous-unités Beta et Alpha pour lesquelles 1 peptide de plus est détecté dans la condition « plaque ». Pour 50ng de digestat déposé, le nombre de peptides varie entre les conditions « tube » et « plaque » selon les sous-unités. Pour la sous-unité Zeta, 1 peptide de

## ETUDES ET RESULTATS : Importance du support

plus est détecté sur la plaque, pour les sous-unités Delta et Theta-1, le même nombre de peptides est détecté et enfin pour les 4 autres sous-unités, 1 peptide de moins est détecté sur la plaque. Par ailleurs, la couverture de séquence est significativement équivalente avec une quantité de matériel déposée élevée mais reste inchangée pour les différents supports utilisés (Figure 81). Pour la sous unité alpha, elle est sensiblement plus élevée pour les gouttes contenant 200ng. En conclusion, l'utilisation de la plaque MALDI pour la récupération immédiate des peptides de l'hémoglobine permet bien la récupération de l'échantillon dans les conditions du test aussi bien qu'un tube *low binding* pour les quantités considérées.

Sous unité de l'hémoglobine humaine (identifiant UniProt)	Nombre de peptides théoriques par la trypsine (2 <i>miscleavages</i> )	Nombre de peptides identifiés dans les conditions expérimentales			
		Tube 200ng	Plaque 200ng	Tube 50ng	Plaque 50ng
Zeta (P02008)	36	4	4	3	4
Delta (P02042)	37	13	13	13	13
Theta-1 (P09105)	28	3	3	2	2
Beta (P68871)	32	20	21	22	21
Gamma-1 (P69891)	36	11	11	10	9
Gamma-2 (P69892)	35	11	11	10	9
Alpha (P69905)	28	21	24	22	21

Tableau 7 : Nombre de peptides identifiés en nano LC-MS/MS après dépôt du digestat d'hémoglobine humaine (200ng ou 50ng) sur une plaque MALDI comparé au nombre de peptides théoriques obtenus après digestion par la trypsine avec 2 *miscleavages* autorisés et compris entre 500 et 4000 m/z.

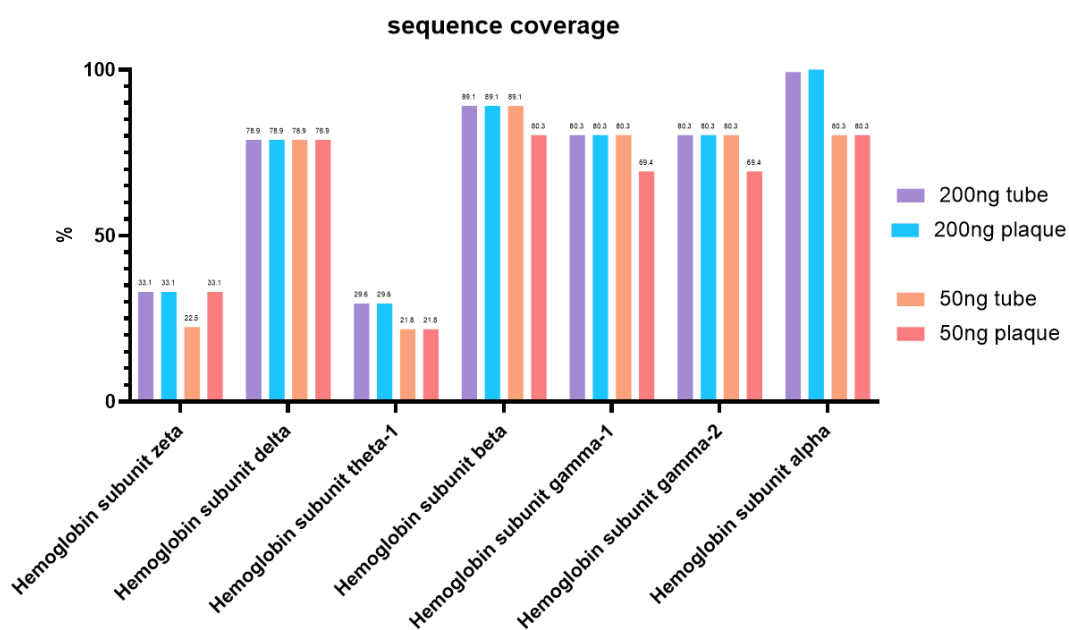


Figure 81 : Couverture de séquence pour chaque sous unité de l'hémoglobine après dépôt de 200 ng ou 50 ng sur une plaque MALDI.

### (b) Application sur des extraits protéiques

Pour que les protéines extraites de cellules soient analysées en *bottom-up*, il faut les digérer en peptides grâce à des endoprotéases, par défaut la trypsine couplée à la Lys-C afin d'obtenir une digestion plus efficace (voir introduction). Les conditions idéales pour une protéolyse impliquent une incubation à 37°C pendant plusieurs heures de la solution protéique. Or, sur une surface plane non couverte, chauffée à 37°C, le dépôt de gouttes s'évapore au bout de quelques minutes seulement. N'ayant pas d'environnement contrôlé en température et en humidité pour éviter l'évaporation, des récipients contenant de l'eau à 100°C ont été placés en présence de la plaque MALDI dans une boîte fermée afin de saturer l'atmosphère d'eau et de limiter l'évaporation (Figure 122).

#### Méthode

Des aliquots d'1µl de 50ng de protéines de levure *Saccharomyces cerevisiae* (Promega) ont été déposées dans un tube et sur une plaque MALDI (Opti-TOF™, AB Sciex). Une digestion à 37°C pendant 2h sous agitation thermostatée (Thermocycler, Eppendorf) et sur une plaque chauffante respectivement a été effectuée par ajout d'1µl de trypsine-Lys C (ratio 1:20 (E/S)). Les échantillons sont analysés sur un Q Exactive HF. Le nombre médian des 3 réplicats de préparation est annoté au-dessus des barres d'intervalles de confiance à 95% sur les graphiques (Figure 82).

#### Résultats

Il faut noter qu'une solution de référence d'un digestat de 50ng de levure *Saccharomyces cerevisiae* analysée par LC-MS/MS sur un Q Exactive HF permet d'identifier environ 1400 protéines, 6300 peptides et 7200 PSM (Figure 123). Cependant, les résultats provenant d'extraits de protéines digérés en petite concentration en laboratoire ne seront pas directement comparables à ceux obtenus pour un digestat commercial.

Pendant les 2h de digestion à 37°C des protéines extraites, une évaporation complète a été constatée sur la plaque MALDI dès les 30 premières minutes de chauffage (Figure 122). Le nombre médian de protéines identifiées est de 664 protéines pour les échantillons déposés dans le tube et 582 protéines pour les échantillons déposés sur la plaque (Figure 82a). De plus, 2571 peptides sont identifiés pour les échantillons dans le tube et 1999 peptides pour les échantillons sur la plaque. Enfin, 2945 PSM sont identifiées pour les échantillons dans le tube contre 2263 pour ceux sur la plaque (données non montrées). Une plus grande variabilité entre les réplicats de préparation est observée dans les tubes. Le nombre médian de protéines, peptides et PSM identifiés n'est pas significativement différent pour les échantillons déposés dans les tubes et ceux déposés sur la plaque MALDI chauffée.

D'après ces résultats, le dépôt des gouttes sur une surface plane égale leur dépôt au fond d'un tube malgré une évaporation importante dans la plaque. L'hypothèse principale pour expliquer la légère diminution des identifications dans la plaque est que l'évaporation entrave la digestion par la trypsine et la rend incomplète ce qui explique une moins bonne identification dans les gouttes déposées sur la plaque à 37°C.

### (c) Évaluation des plaques MALDI modifiées : les *AnchorChip*

Une amélioration de ce protocole est indispensable. Nous avons évalué l'apport du traitement de surface des plaques appelées *AnchorChip* sur lesquelles des ancrs hydrophiles ont été déposées sur une surface hydrophobe afin d'aider à la concentration des échantillons.

## ETUDES ET RESULTATS : Importance du support

En plus de l'utilisation de ces plaques, il a été décidé de tester une digestion à température ambiante (20°C) afin d'éviter une évaporation complète des gouttes sur la plaque.

### *Méthode*

Des aliquots d'1µl de 50ng de protéines de levure *Saccharomyces cerevisiae* ont été déposées dans un tube et sur une plaque *AnchorChip* (Bruker). Une digestion à 20°C pendant 2h a été effectuée par ajout d'1µl de trypsine-Lys C (ratio 1:20 (E/S)). Les échantillons sont analysés sur un Q Exactive HF. Les résultats des 3 réplicats sont représentés sur les graphiques (Figure 82b).

### *Résultats*

À 20°C, sur les plaques *AnchorChip*, l'évaporation des gouttes devenait visible après 2 heures de digestion et non à 30minutes comme pour la condition précédente à 37°C. Le nombre médian de protéines identifiées pour les échantillons déposés dans le tube est de 543 contre 690 protéines pour ceux déposés sur *AnchorChip* (Figure 82b). Le nombre médian de peptides est de 1826 pour les échantillons déposés dans le tube contre 2418 peptides identifiées pour ceux déposés sur *AnchorChip*. Enfin, 2974 PSM ont été identifiés pour les échantillons déposés dans les tubes contre 2756 PSM pour ceux déposés sur *AnchorChip*. Une très grande variabilité est observée dans les tubes.

En diminuant la température, le nombre médian de protéines, peptides et PSM identifiés dans les tubes a diminué de 100 protéines, 700 peptides et 900 PSM comparé à la condition tubes à 37°C. En tube, à 20°C, la digestion n'étant pas faite à une température optimale, le résultat attendu est une digestion incomplète et donc une moins bonne identification.

Pour les échantillons déposés sur les plaques *AnchorChip* à 20°C, les identifications ont augmenté par rapport à la condition à 37°C avec la plaque MALDI. Cette augmentation confirme que plus l'évaporation est retardée et donc plus la trypsine reste en solution, plus la digestion est efficace et donne de meilleurs résultats d'identification. Cette augmentation confirme également qu'un traitement de surface est très important pour améliorer les identifications. Cela signifie que même dans des conditions suboptimales (digestion à 20°C et évaporation), ces identifications sont maintenant supérieures aux identifications en tube ce qui confirme également que moins les gouttes sont en contact avec une surface, meilleures sont les identifications obtenues.

Il faudrait répliquer ces résultats pour les confirmer. La conclusion de cette expérience est que la minimisation de la surface de contact et un traitement de surface par ancrages hydrophiles sur une plaque en métal lisse, alors qu'elle est difficile à manipuler et qu'elle n'interagit pas bien avec les peptides, améliorent les identifications peptidiques et protéiques. Le défi est donc de gérer l'évaporation pour permettre une digestion efficace tout en limitant la surface de contact avec l'échantillon. N'ayant pas d'environnement contrôlé, le projet a continué dans la recherche de supports ayant de surfaces de contact miniaturisées en gardant un format de plaque multi-puits car il est plus facile d'y gérer l'évaporation.

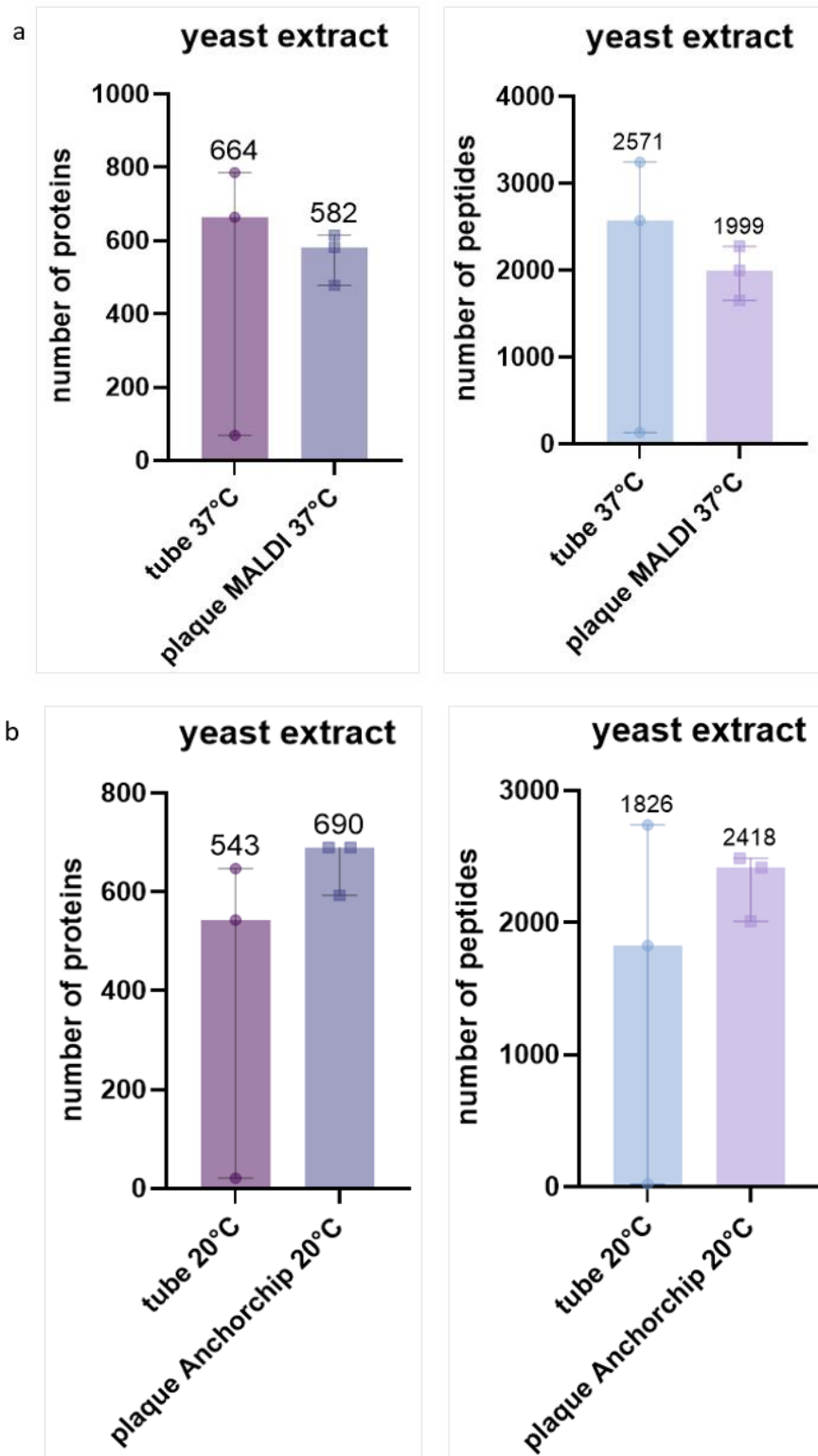


Figure 82 : Nombre de protéines et peptides d'extraits de levure *Saccharomyces cerevisiae* à 50ng/ $\mu$ l digérés pendant 2h (a) à 37°C dans un tube ou après dépôt sur une plaque MALDI OptiTOF (b) à 20°C dans un tube ou après dépôt sur une plaque *AnchorChip*. Les médianes des répliquats sont annotées au-dessus des barres d'intervalles de confiance à 95%.

### 6.A.2. Miniaturisation de la surface de contact sur géométrie en puits

La recherche d'un support en format de plaque multi-puits avec une surface de contact minimale a mené à la fabrication de différents prototypes de supports selon deux techniques de micro-fabrication citées dans l'introduction. Un prototype fabriqué en PDMS a été préféré en première approche pour ses propriétés (coût, facilité de fabrication...) évoquées dans l'introduction. Puis un second prototype a été fabriqué en polypropylène car c'est le matériau le plus répandu en protéomique. Les techniques de photolithographie et de micro-usinage ont été utilisées pour la mise au point des différents designs.

#### (a) Micro-fabrication de plaques micro-puits en PDMS : 3 designs

Pour pallier le problème d'évaporation, la conception d'une plaque micro-puits en PDMS a été réalisée (Figure 83). L'idée de départ était un format de plaque (ayant les mêmes dimensions qu'un réceptacle de tubes LC de 40 positions) contenant des micro-puits dont le volume ne dépasse pas 10 $\mu$ l. Pour comparaison, les plaques commerciales utilisées possèdent des puits de 120 $\mu$ l de volume.

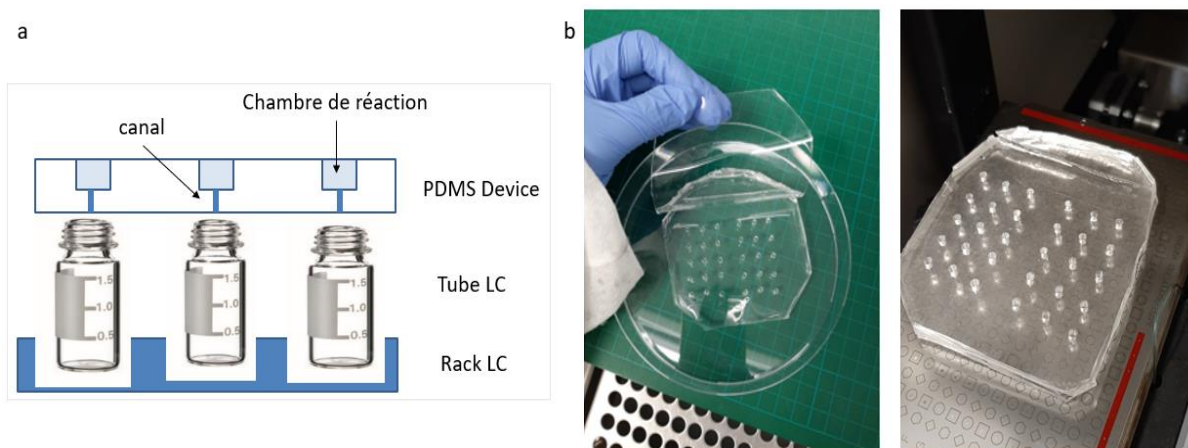


Figure 83 : 1<sup>er</sup> design du support miniaturisé (a) schéma du design (b) photographies du design poinçonné

#### Méthode

Brièvement, des cercles de 875 $\mu$ m de rayon (1.7mm de diamètre) espacés de 7.5mm ont été dessinés pour être reproduits sur un masque puis reproduits sur un *wafer* par photolithographie comme discuté dans l'introduction (voir « la micro-fabrication dans ce projet »). Une fois le moule obtenu, ces cercles ont été utilisés comme indicateurs de positions. Ensuite, le PDMS a été coulé sur le *wafer* et un poinçon de 2.5mm a été utilisé pour créer les chambres de réaction sur la première couche de PDMS. D'autre part, une seconde couche de PDMS est coulée sur un *wafer* vierge. Par la suite, les deux couches ont été traitées au plasma à oxygène et collées entre elles après activation de leur surface par le plasma à oxygène. Un poinçon de 750 $\mu$ m a été utilisé pour créer les canaux dans la deuxième couche. Ce canal doit rester assez petit pour ne pas permettre un écoulement sans qu'une force (pression ou aspiration) ne soit appliquée au liquide. Ce protocole est décrit en détails dans le matériel et méthodes.



Résultats

Le design consiste en une chambre de réactions de 10 $\mu$ l de volume puis un canal en dessous qui relie chaque chambre à un tube LC en verre. L'objectif était de faire la préparation d'échantillons dans la chambre à réactions puis de pousser les peptides dans les tubes LC à travers le canal à l'aide de la pression de l'air. En effet, après la digestion, l'élution peut se faire directement dans les tubes LC par pousse-seringue. L'utilisation de la seringue permet de créer l'écoulement du contenu de chaque chambre de réaction à travers le canal. Pour limiter l'évaporation durant les réactions, la chambre de réaction peut être fermée à l'aide d'un film comme les plaques multi-puits classiques ou d'une couche de PDMS.

Grâce à ce design, deux puces de 18 positions chacune peuvent être fabriquées en parallèle sur un seul *wafer* (Figure 84). Ces différentes étapes ont donc permis d'avoir des chambres de réaction, qui, une fois fabriquées en PDMS sont parfaitement alignées à des tubes LC disposés sur un réceptacle Dionex par exemple. Les gouttes peuvent donc être déposées directement dans les chambres de réaction pour effectuer les différentes étapes de traitement des protéines extraites ou des cellules.

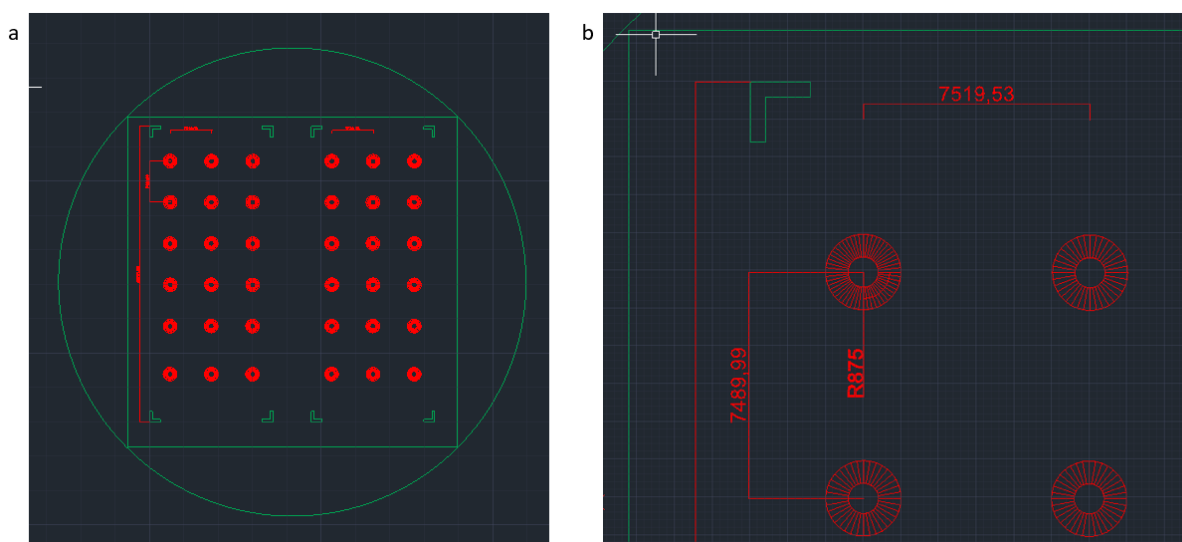


Figure 84 : (a) Dessin sur AutoCAD du 1<sup>er</sup> design de puce selon les dimensions d'un *wafer* (b) dimensions des puits et espacements entre les puits en  $\mu$ m.

Cependant, les trous étant faits à la main, ce premier design présentait plusieurs problèmes. Les coordonnées des trous présentaient un manque de régularité. Il était difficile d'envisager le couplage avec un appareil automatique. L'étanchéité des deux couches n'était pas garantie et plusieurs échantillons se sont vidés entre les couches après une pression de la seringue. Mais le plus souvent, la pression de la seringue n'était pas assez forte pour que l'échantillon complet sorte par le canal et il restait donc dans la chambre de réaction. Un prélèvement des échantillons par aspiration depuis une chambre de réaction fermée a donc été préféré pour les futurs designs par souci de robustesse.

Pour une meilleure reproductibilité des coordonnées et une plus grande facilité de fabrication, la conception d'un moule en laiton (qui est le reflet négatif de la plaque voulue) a été mise en œuvre sur la plateforme technique de l'Institut Pierre Gilles de Gennes (IPGG). La micro-fabrication du moule en laiton a été faite par l'utilisation d'une micro-usineuse par une technique de micro-fabrication appelée micro-usinage par perçage (Figure 85). Ensuite,

## ETUDES ET RESULTATS : Importance du support

du PDMS est coulé sur le moule afin d'avoir une plaque micro-puits en PDMS avec des micro-puits de 2.5mm de diamètre et 3mm de hauteur, ce qui fait 10 $\mu$ l de volume. Deux plaques de 27 puits peuvent être fabriquées en parallèle sur le même moule. Ce 2<sup>ème</sup> design présente plusieurs avantages car le temps de fabrication est réduit, les coordonnées sont reproductibles et le volume des puits est plus petit que celui d'une plaque commerciale.

Cependant, les dimensions des puits ne correspondaient pas parfaitement à ceux d'une plaque commerciale et l'interface avec des appareils qui utilisent des coordonnées de plaques standards (comme le trieur de cellules) était encore difficile à réaliser.

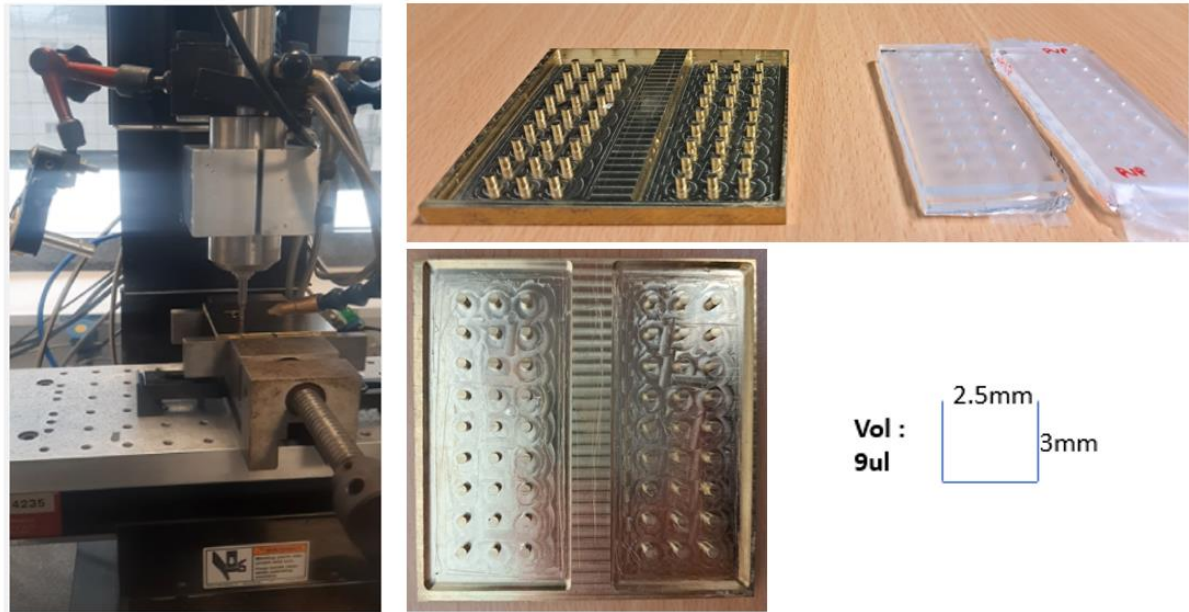


Figure 85 : 2<sup>ème</sup> design du support miniaturisé (à gauche) photographie de l'usineuse en plein fabrication du moule en laiton (à droite en haut) photographie du moule en laiton et d'une puce en PDMS fabriquée à partir du moule (à droite en bas) photographie du moule en laiton vu de haut et dimensions d'un puits.

La conception d'un nouveau moule avec un fond en V afin de limiter l'adsorption a été dessinée sur le logiciel de design Fusion 360 (Figure 86). L'usinage d'un nouveau moule en laiton sur la plateforme technique à l'IPGG, constitue le 3<sup>ème</sup> design de plaque micro-puits de ce projet.

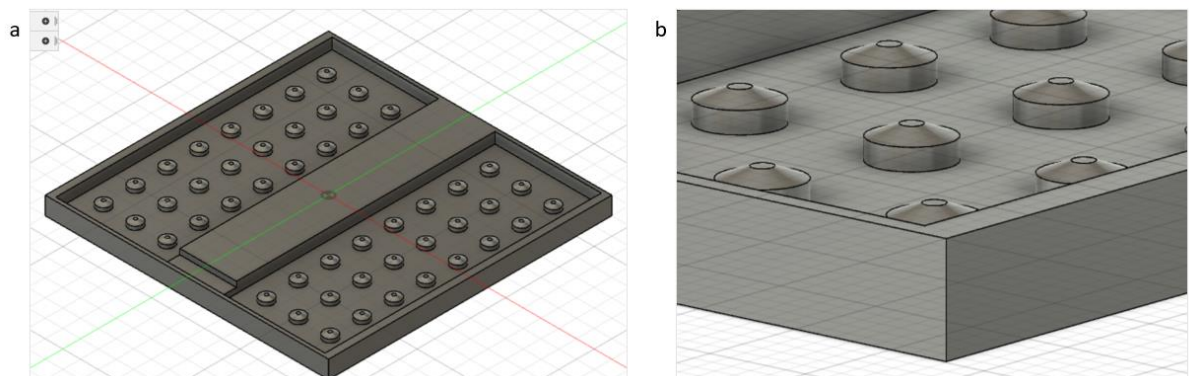


Figure 86 : 3<sup>ème</sup> design de support miniaturisé (a) Dessin du design de moule en laiton à fond V (b) zoom sur la forme des puits (reflet négatif de la plaque PDMS).

## ETUDES ET RESULTATS : Importance du support

Dans ce design, les diamètres des puits ont été élargis à 3.8 mm pour correspondre au diamètre des puits d'une plaque commerciale standard (384 multi-puits). La hauteur des puits a été rétrécie à 1mm afin de conserver un volume de 10 $\mu$ l par puits (Figure 87). Le fond en V fait aussi 1mm de profondeur. Grâce à ces nouvelles dimensions, les plaques sont maintenant adaptées à l'utilisation avec des équipements demandant des coordonnées standards comme le trieur de cellules. Ce design a donc été choisi pour l'ensemble des expériences qui suivent.

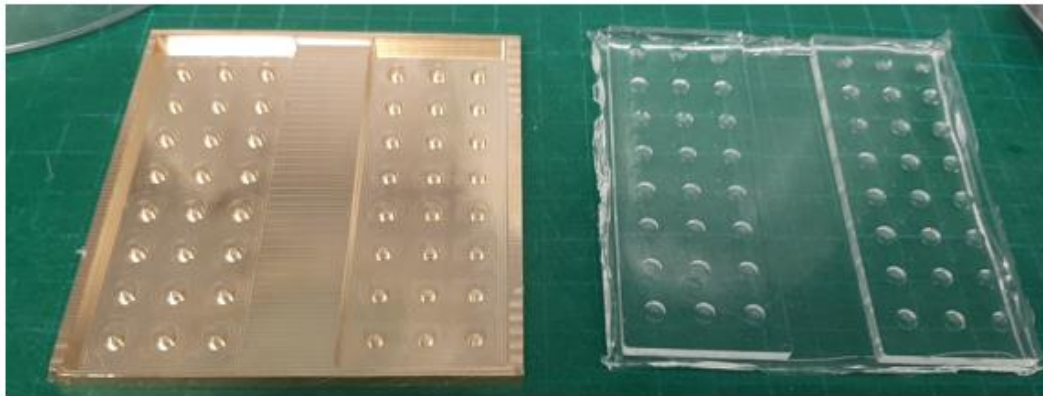


Figure 87 : 3<sup>ème</sup> design du support miniaturisé (en haut) photographie du moule en laiton et d'une plaque micro-puits en PDMS fabriquée à partir de ce moule (en bas) dimensions d'un puits de la plaque.

### (b) Impression 3D de plaques micro-puits en polypropylène

Dans l'industrie, les plaques commerciales en polypropylène sont fabriquées par moulage de polypropylène pur. Notre prototype en polypropylène a été réalisé en impression 3D.

#### *Méthode*

Les dimensions du 3<sup>ème</sup> design de plaques micro-puits ont été conservées afin d'allier le petit volume de 10 $\mu$ l et l'interface avec les instruments automatiques comme le trieur de cellules (Figure 88). Un design a été effectué sur le logiciel Fusion (Autodesk) puis ce design est adapté à l'imprimante 3D par un logiciel de « *slicing* » qui découpe le motif en couches d'impressions et une jupe (20 couches de large et 5 couches de haut) est ajoutée au motif afin d'assurer son ancrage. Ensuite, du fil en polypropylène blanc (PolyPro Ultra, Volumic) est chargé sur l'imprimante 3D Volumic®. L'impression est lancée avec des paramètres de vitesse « standard » et de précision « haute » (100 $\mu$ m) et un remplissage de 100%.

#### *Résultats*

La plaque micro-puits de 96 puits est obtenue après quelques heures de fabrication (10 heures par plaque). Ces plaques ont été fabriquées à une étape avancée du projet et n'ont pu être testées que pour quelques expériences.

## ETUDES ET RESULTATS : Importance du support

Cependant, la résolution de la fabrication est limitée à la taille de la buse utilisée (ici 100 $\mu$ m) et au temps de fabrication nécessaire. L'impression 3D utilisée pour le prototypage n'est pas une solution pérenne pour la fabrication de plaques car le temps de fabrication demandé pour atteindre une meilleure résolution serait démultiplié. D'autre part, cette impression n'est pas compatible avec l'obtention d'une surface lisse. Pour cela, il aurait fallu utiliser une technique de fabrication par moulage.

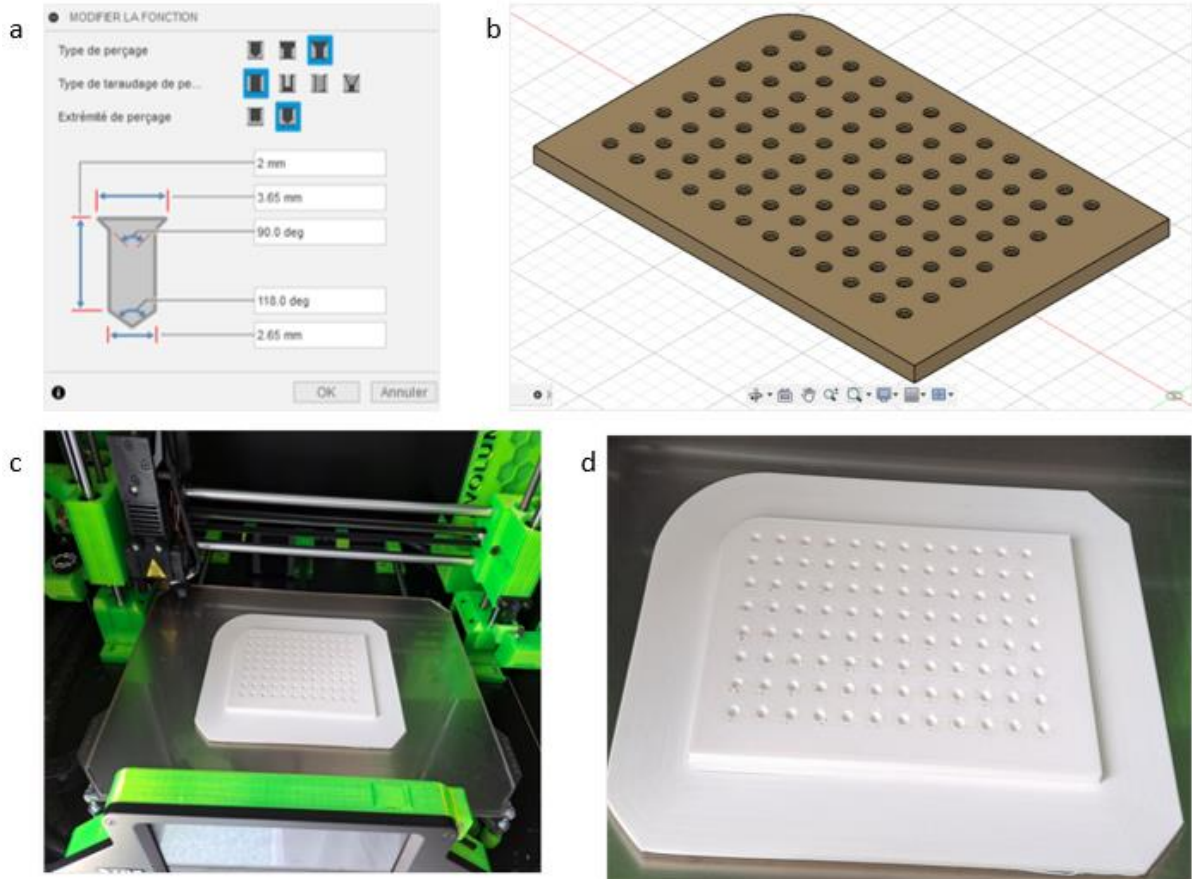


Figure 88 : (a) Dimensions du puits dans la plaque micro-puits en PP (b) design de la plaque sur le logiciel Fusion (Autodesk) (c) photographie d'une plaque micro-puits en PP à la fin de l'impression 3D (d) zoom sur une plaque micro-puits en PP (PP 10 $\mu$ l) et sa jupe d'ancrage.

## 6.B Comparaison du support micro-puits en PDMS avec les plaques commerciales

Dans les figures de ce manuscrit, les plaques PDMS micro-puits de 10 $\mu$ l par puits sont appelées « PDMS 10 $\mu$ l » et les plaques commerciales standard 384 puits de 120 $\mu$ l par puits sont appelées « PP 120 $\mu$ l ».

### *Méthode*

Un aliquot de 1 $\mu$ l de 50ng de deux extraits protéiques de macrophages humains K562 (Promega) et de levure (*Saccharomyces cerevisiae*, Promega) ont été déposés sur la plaque standard (PP 120 $\mu$ l) et la plaque micro-puits (PDMS 10 $\mu$ l). Puis, ils ont été digérés par 1 $\mu$ l de trypsine-Lys C (ratio 1:20 (E/S)). L'équivalent de 10ng ont été analysés sur un Tims TOF Pro. Le nombre médian des protéines identifiées à partir des 6 ou 12 réplicats de préparation est annoté au-dessus des barres d'intervalles de confiance à 95% de chaque condition sur la Figure 89.

Dans cette expérience, deux plaques commerciales standard de 384 puits (PP 120 $\mu$ l) (12 réplicats) et 1 plaque micro-puits en PDMS (PDMS 10 $\mu$ l) (6 réplicats) sont comparées.

### *Résultats*

Pour les échantillons humains, 1160 protéines ont été identifiées dans la plaque micro-puits alors que 1004 protéines seulement ont été identifiées dans la plaque standard PP 120 $\mu$ l. Pour les échantillons de levure, la même tendance est retrouvée avec 1149 protéines identifiées dans la plaque micro-puits et seulement 821 protéines identifiées dans la plaque standard. Pour ces deux exemples, le nombre médian de protéines et de peptides est plus élevé dans la plaque micro-puits en PDMS (PDMS 10 $\mu$ l) que dans la plaque commerciale standard (PP 120 $\mu$ l). De plus, cette différence est significative pour l'échantillon de levure d'après le test statistique effectué.

L'utilisation de la plaque micro-puits en PDMS comparée aux plaques commerciales et donc la minimisation du support permet une meilleure identification en nombre de protéines et peptides en induisant très probablement une réduction de l'adsorption des peptides et protéines.

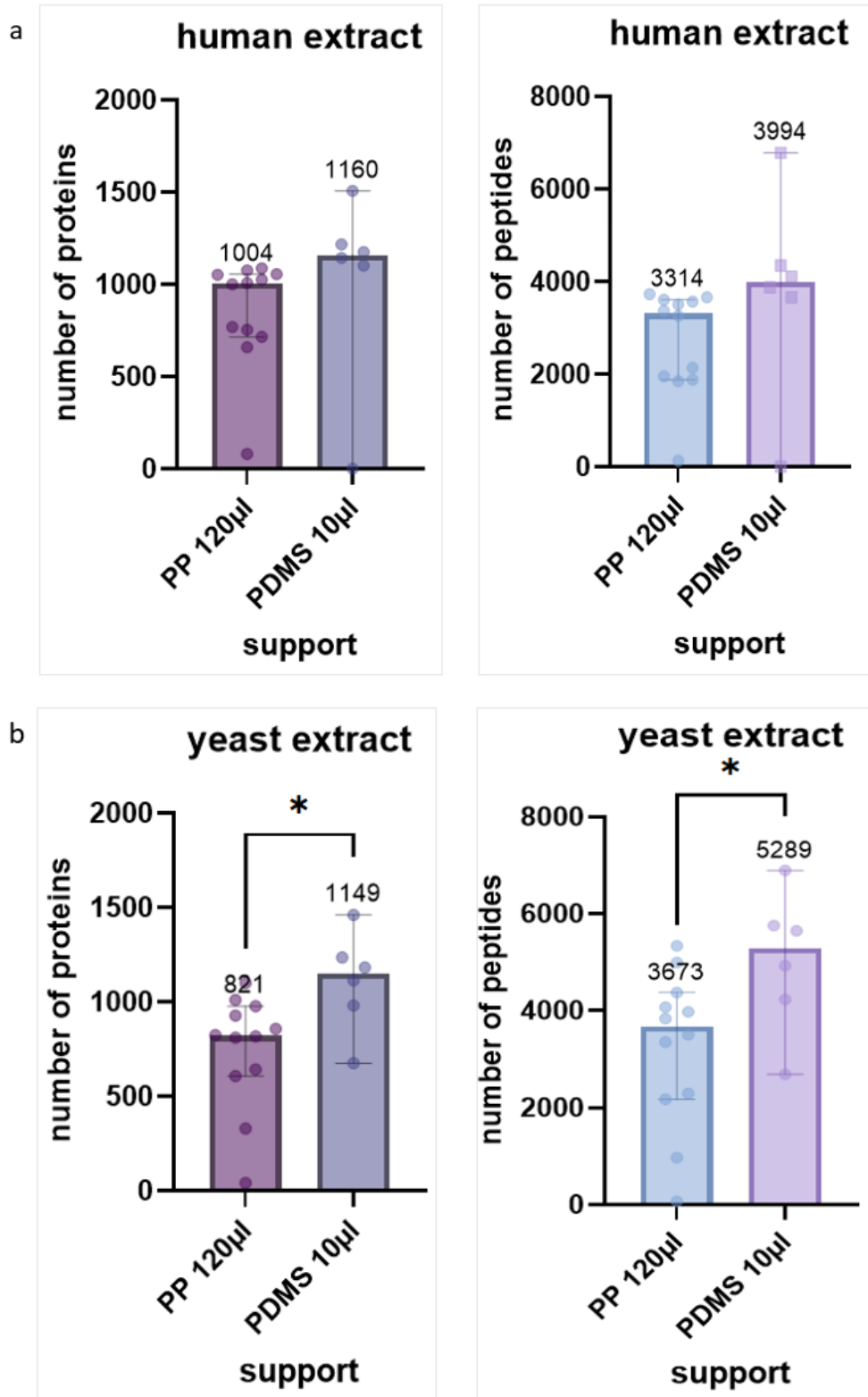


Figure 89 : Nombre de protéines et peptides identifiés à partir d'extraits de protéines humaines ou de levure, digérées en plaque commerciale 384-puits standard (PP120) ou dans une plaque micro-puits en PDMS (PDMS10). Les valeurs au-dessus des barres d'intervalles de confiance à 95% représentent les médianes des 12 réplicats. Les \* représentent une p-value <0.05 obtenue par un t-test non apparié (avec \*<0.05, \*\*<0.005, \*\*\*<0.0005).

### 6.C Compatibilité de la plaque micro-puits avec le trieur de cellules

La compatibilité de la plaque micro-puits avec le trieur de cellules a été vérifiée. Après les essais sur des échantillons commerciaux d'extraits protéiques, un essai sur des cellules Jurkat a été effectué afin de contrôler la récupération des protéines obtenues avec 2000 et 400 cellules.

#### Méthode

Selon nos estimations, 2000 cellules Jurkat représentent 100ng de protéines et 400 cellules Jurkat représentent environ 20ng (Tableau 1). Un protocole minimaliste discuté dans la première partie du projet (lyse thermique suivie d'une digestion enzymatique) a été appliqué à ces cellules (voir « amélioration du protocole »). Brièvement, les cellules ont été digérées par de la trypsine-Lys C (ratio 2:1 (E/S)). Ensuite, l'équivalent d'un cinquième de l'échantillon a été analysé, ce qui équivaut à 400 cellules et 80 cellules respectivement, sur un Tims TOF Pro. Le nombre médian des 13 réplicats de préparation est annoté au-dessus des barres d'intervalles de confiance à 95% de chaque condition sur la Figure 90.

#### Résultats

Un nombre médian de 1848 et 701 protéines a été identifié pour 400 et 80 cellules, respectivement. La plaque micro-puits permet la réception des cellules et la récupération protéique et peptidique à partir de cellules triées avec le trieur de cellules.

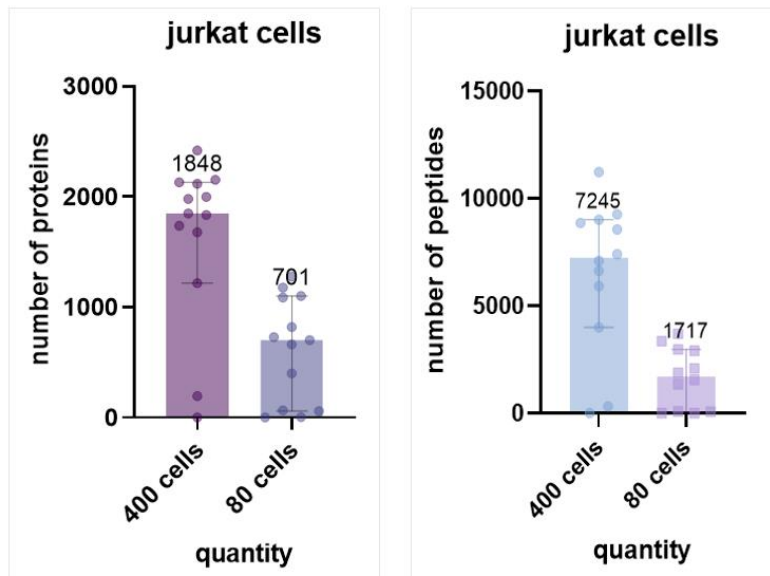


Figure 90 : Nombre de protéines et peptides obtenus à partir de cellules Jurkat triées sur une plaque micro-puits (PDMS 10µl). Les médianes des 13 réplicats de préparation sont annotées au-dessus des barres d'intervalles de confiance à 95%.

## 6.D Apport d'un traitement de surface

Avec les expériences sur les plaques MALDI et nos supports en micro-puits, la miniaturisation des supports a démontré une amélioration des identifications peptidiques et protéiques. Aussi, l'expérience sur les plaques *AnchorChip* a démontré qu'un traitement de surface avec des ancrs hydrophiles apporte un gain d'identification malgré des conditions suboptimales. Nous avons donc décidé de créer des ancrs hydrophiles sur nos supports.

### 6.D.1. Traitement de surface avec du polyvinylpyrrolidone (PVP)

Pour cela, un traitement du PDMS simple par le polyvinylpyrrolidone (PVP) a été trouvé dans la littérature et testé dans un premier temps sur un prototype linéaire très simple en PDMS.

#### (a) Qu'est-ce que le Polyvinylpyrrolidone (PVP) ?

Le polyvinylpyrrolidone est un polymère organique hydrosoluble synthétisé par polymérisation du monomère N-vinylpyrrolidone<sup>207</sup>. De formule brute  $(C_6H_9NO)_n$ ,  $n$  étant le nombre de monomères, sa structure est un noyau aliphatique (Figure 91). Le PVP se lie exceptionnellement bien aux molécules polaires, en raison de sa polarité. Cela a conduit à son application dans les revêtements de surface.

Comme discuté dans l'introduction, plusieurs traitements de surface peuvent être utilisés pour diminuer l'hydrophobicité des surfaces utilisées en protéomique pour la manipulation des échantillons. Dans la littérature, Hemmila *et al.* ont testé 3 traitements de surface pour rendre le PDMS plus hydrophile et ont étudié leurs effets durant 6 mois<sup>131</sup>. L'un de ces traitements était un revêtement de la surface du PDMS par du PVP. La conclusion de leur étude est que la technique la plus simple et la plus rapide pour rendre le PDMS hydrophile pendant au moins six mois était de fixer du PVP ou du PEG sur une surface de PDMS activée au plasma d'oxygène.

Or, le PEG est un contaminant déjà connu en protéomique par spectrométrie de masse à cause de sa présence dans les plastiques de laboratoire (cônes de pipettes, capillaires etc.). Ses propriétés physicochimiques lui permettent d'être très facilement ionisé en ESI et par là d'induire des effets de suppression spectrale très délétères pour la détection des peptides et protéines en faible concentration. Ceci en fait un composé peu compatible avec les analyses *bottom-up*. Dans ce projet, la compatibilité du PVP avec la protéomique par spectrométrie de masse et son effet sur l'hydrophobicité du PDMS ont donc été vérifiés.

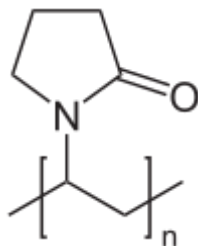


Figure 91 : Structure du polyvinylpyrrolidone (PVP).



### (b) Mise en place du traitement de surface

Des essais préliminaires ont été effectués pour vérifier l'efficacité d'un traitement de surface au PVP sur une puce linéaire simple en PDMS. Cette puce était formée d'une entrée (« in »), un réservoir rectangulaire (appelé aussi canal) et une sortie (« out ») (Figure 93). L'objectif est de vérifier la rétention des peptides extraits de cellules HeLa à travers cette puce.

#### Méthode

Un design de puce (Figure 92) a été dessiné sur le logiciel AutoCAD (*student version, 2020*). Le processus de photolithographie est décrit dans la partie matériel et méthodes. Une fois que le *wafer* est prêt, il est entouré de papier aluminium et 50g de PDMS (Sylgard 184) et son réticulant en ratio 10:1 sont coulés sur le *wafer*. Le *wafer* est placé dans une pompe sous vide pendant 1 heure afin de retirer toutes les bulles d'air et enfin, placé à 70°C pendant 2 heures jusqu'à durcissement du PDMS.

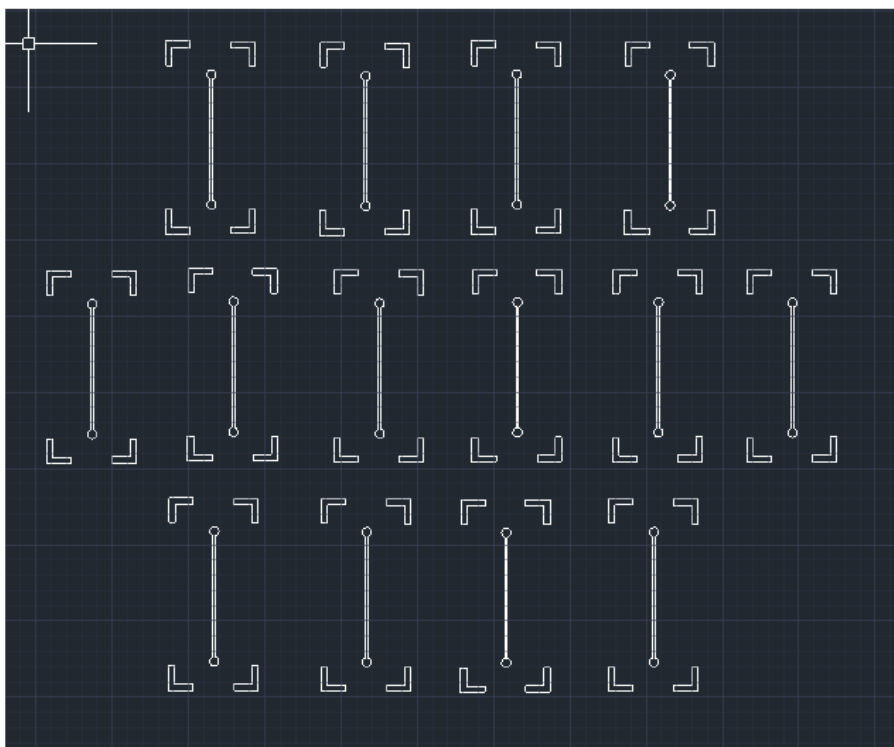


Figure 92 : Dessin sur le logiciel AutoCAD des puces créées pour le test du traitement au PVP sur un *wafer*.

Le PDMS durci est retiré du *wafer*. Les puces sont découpées et à l'aide de poinçons, des trous de la taille requise sont poinçonnés selon les demandes des designs pour former les puces. Les puces sont ensuite traitées au plasma à oxygène en même temps que des lames en verre afin d'activer les surfaces et les coller ensemble de façon permanente. Les puces sont ensuite traitées avec une solution de PVP à 0.2% (v/v) ou à l'eau (condition contrôle) pendant 10 minutes afin de les rendre plus hydrophiles et lavées à l'eau.

Un capillaire de 760µm est inséré dans les trous et connecté à une seringue de 1ml pour laver les puces à grand débit avec de l'eau Milli-Q. Ensuite, une pompe microfluidique Flow EZ® (Fluigent) a été utilisée pour avoir un débit constant à 1µl/min afin d'injecter des échantillons dans les puces ainsi que les solvants d'élution.

## ETUDES ET RESULTATS : Importance du support

Deux concentrations d'un échantillon complexe de peptides de cellules HeLa (50ng/ $\mu$ l et 1.6ng/ $\mu$ l) ont été injectées dans ce réservoir afin d'évaluer la rétention des peptides par le PDMS et l'effet du traitement de surface au PVP sur cette rétention d'une heure. Les peptides sont récupérés directement à la sortie de la puce dans des tubes LC et 1 $\mu$ l est analysé sur un Q Exactive HF. Les médianes des 3 réplicats de préparation sont annotées au-dessus des barres d'intervalles de confiance à 95% (Figure 94).

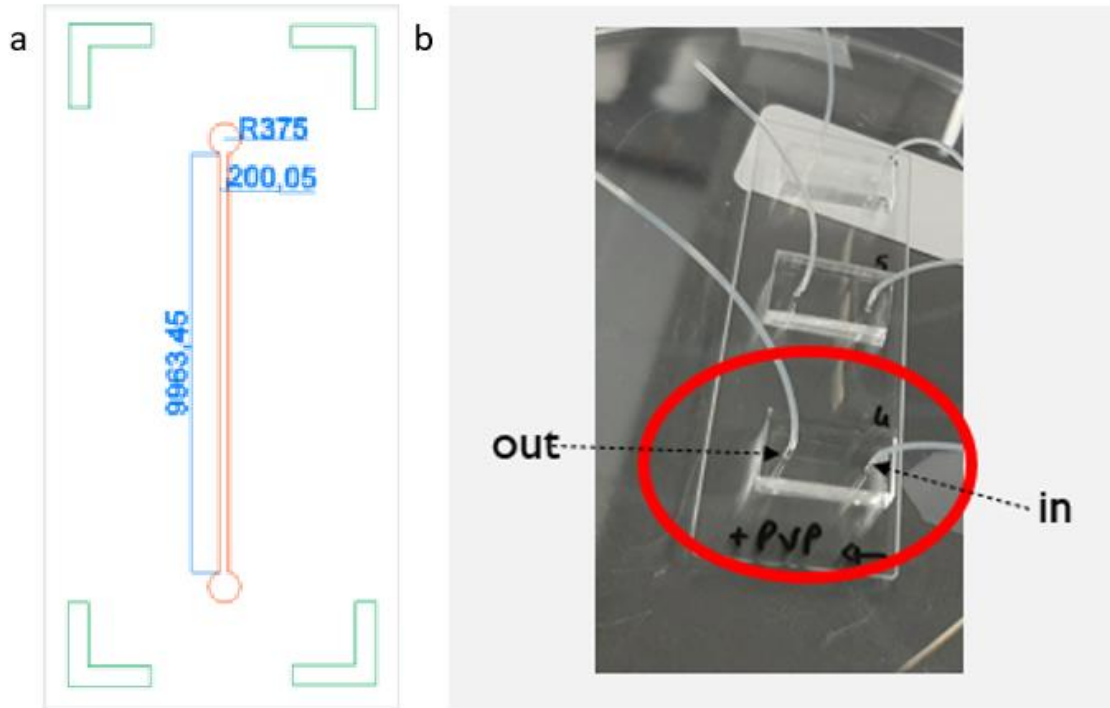


Figure 93 : Design de la puce pour les traitements de surface préliminaires. (a) Design sur le logiciel AutoCAD avec les dimensions des puits créés par photolithographie. (b) Photographie de 3 puces attachées sur une lame de verre. Le cercle rouge indique 1 puce avec son entrée et sa sortie branchées à des capillaires.

### Résultats

Le volume du support a été calculé selon les dimensions du logiciel AutoCAD (Figure 93).

- Le canal : Les dimensions du canal sont 10mm (10000 $\mu$ m) de longueur, 200 $\mu$ m de largeur et 50 $\mu$ m de hauteur. Ces canaux peuvent être modélisés comme des parallélépipèdes rectangles de ces dimensions est posée. Le volume de la surface de contact du canal est donc de :  $10000 \times 200 \times 50 = 100.000.000 \mu\text{m}^3 = 100 \text{ nl} = 0.1 \mu\text{l}$ .
- Les cylindres des puits : Les cylindres formés par les poinçons pour les puits d'entrée et de sortie font 375 $\mu$ m de rayon et 1cm de hauteur, ce qui donne un volume de 4.42  $\mu$ l.
- Le total de la puce est donc  $4.42 + 0.1 + 4.42 = 8.94 \mu\text{l}$  de volume.

Ce volume de 8.85 $\mu$ l permet de comparer l'effet du PVP sur la rétention peptidique et se rapproche de nos plaques micro-puits à 10 $\mu$ l.

### (c) Compatibilité du PVP avec la spectrométrie de masse

Il est important de vérifier la contamination par les polymères utilisés pour le traitement de surface. Cette contamination peut être très importante et peut alors altérer les signaux obtenus en spectrométrie de masse jusqu'à les masquer complètement.

#### *Définitions*

Le chromatogramme en courant total d'ions (TIC, pour *total ion current*) représente la somme des intensités ioniques sur toute la gamme des masses détectées à chaque temps t donné <sup>208</sup>. Le signal maximum obtenu pour l'analyse d'un blanc d'échantillon est de  $1.6 \times 10^8$  et est de  $3 \times 10^9$  pour les échantillons de 50ng de HeLa. Les TICs des échantillons témoins pour chaque condition sont présentés en annexe (Figure 125). Dans les échantillons complexes, le chromatogramme TIC fournit souvent des informations limitées car plusieurs analytes sont élués simultanément, et le TIC n'informe pas sur les espèces individuelles.

Le chromatogramme du pic de base (BPI, pour *base peak intensity*) représente l'intensité du pic le plus intense du spectre à chaque point de l'analyse <sup>208</sup>. Ceci réduit le bruit de fond en se concentrant sur un analyte à chaque point et donne un aspect plus net et donc plus informatif aux chromatogrammes.

#### *Résultats*

Le spectre théorique du polymère PVP <sup>209</sup> et celui d'un contaminant habituel le polyéthylène glycol (PEG) <sup>99</sup> sont présentés en annexe (Figure 124). Les spectres de masse d'un échantillon de PEG et de PVP présentent une distribution de signaux espacés de 44 Da et de 85 Da, respectivement, égal à la masse de l'unité répétitive de la macromolécule. Pour confirmer la présence d'un polymère dans un échantillon, son profil isotopique doit apparaître sur le chromatogramme et dans le spectre. Dans la gamme de masse étudiées ( $m/z$  : 400-1200), les peptides sont rarement monochargés donc les pics monochargés sont attribués à des polymères.

L'analyse d'un blanc d'échantillon ayant traversé la puce traitée ou non au PVP est montrée sur la Figure 95. L'intensité maximale du signal en BPC est à  $1.8 \times 10^8$  et  $8.7 \times 10^7$  pour les deux échantillons. Les deux profils sont inchangés entre les deux échantillons, le traitement au PVP ne semble pas perturber le chromatogramme. De plus, les appels d'ions théoriques ne retrouvent aucun profil isotopique indiquant la présence d'ions PVP dans les échantillons ayant traversé la puce traitée ou non (données non montrées).

L'analyse des échantillons de HeLa (1.6ng et 50ng) ayant traversé la puce traitée ou non au PVP est montrée sur les Figure 96 et Figure 97. Les profils chromatographiques des échantillons étant les mêmes, l'analyse n'apparaît pas perturbée par le traitement au PVP. Les ions théoriques du PVP et le profil isotopique ne sont pas retrouvés dans le BPI ou lors de l'extraction de ces ions (données non montrées).

#### *Bilan*

Les ions théoriques du PVP ne sont pas présents dans les profils chromatographiques après traitement de la puce. Le traitement au PVP ne doit pas être facilement ionisable en condition acide et ne laisse aucune signature polymérique en spectrométrie de masse. Son utilisation pour nos analyses est donc tolérée.

## ETUDES ET RESULTATS : Importance du support

Il faut cependant remarquer la présence des ions du PEG dans tous les échantillons y compris le blanc. Leur profil isotopique est présent avec une différence de masse de 44Da entre les pics (ions 520, 564, 608 etc. dans la Figure 95 par exemple) et le spectre indique des ions monochargés (données non montrées). La contamination par le PEG est courante et attendue en spectrométrie de masse, elle est notamment due aux plastiques de laboratoires et les ions du polymère sont retrouvés aussi bien dans les échantillons témoins que dans les échantillons d'intérêt.

### (d) Application du traitement de surface au PVP sur un mélange complexe

#### *Résultats*

Il faut noter que sur un Q Exactive HF, avec la même configuration de séparation, un digestat de HeLa de 1.6ng permet l'identification de 1279 protéines et 3841 peptides et un digestat de HeLa de 50ng permet l'identification de 2643 protéines et 13316 peptides. Après passage dans les puces, environ 44% des protéines et peptides sont perdus pour 50ng de digestats de cellules HeLa et entre 75 et 85% de protéines et peptides sont perdus pour les échantillons à 1.6ng.

En effet, pour 50ng de digestats de cellules HeLa, 1424 protéines et 6735 peptides ont été identifiés dans les puces non traitées alors que dans les puces traitées au PVP, 1531 protéines et 7420 peptides ont été identifiées. D'après le test statistique effectué, il n'y a pas de différence significative entre les deux conditions de traitement.

Pour 1.6ng de digestats de cellules HeLa, 179 protéines et 375 peptides ont été identifiés dans les puces non traitées alors que 317 protéines et 681 peptides ont été identifiés dans les puces traitées au PVP. D'après le test statistique effectué, il n'y a pas de différence significative entre les conditions traitée ou non traitée avec cette quantité d'échantillon.

#### *Bilan*

Une perte significative est observée dès le passage des digestats cellulaires dans les puces en PDMS. Cette perte est encore plus importante pour des petites quantités d'échantillons. L'adsorption due au passage des échantillons dans les puces en PDMS est délétère à la récupération des peptides et donc à l'identification des protéines. La variabilité d'identifications très importante entre les différentes puces utilisées, due au prototypage de la puce, rend difficile la comparaison des conditions traitées ou non.

Le prototype linéaire utilisé pour cette comparaison n'est pas adapté car le rapport surface de contact/volume est trop faible, le volume du réservoir étant minimal et comportant deux zones de stagnation à l'entrée et à la sortie du réservoir. Le faible renouvellement du contenu de ces deux zones réduit l'efficacité des traitements et lavages. Il faut amplifier le phénomène afin de mieux observer les effets causés par le contact des protéines et peptides avec la surface traitée ou non.

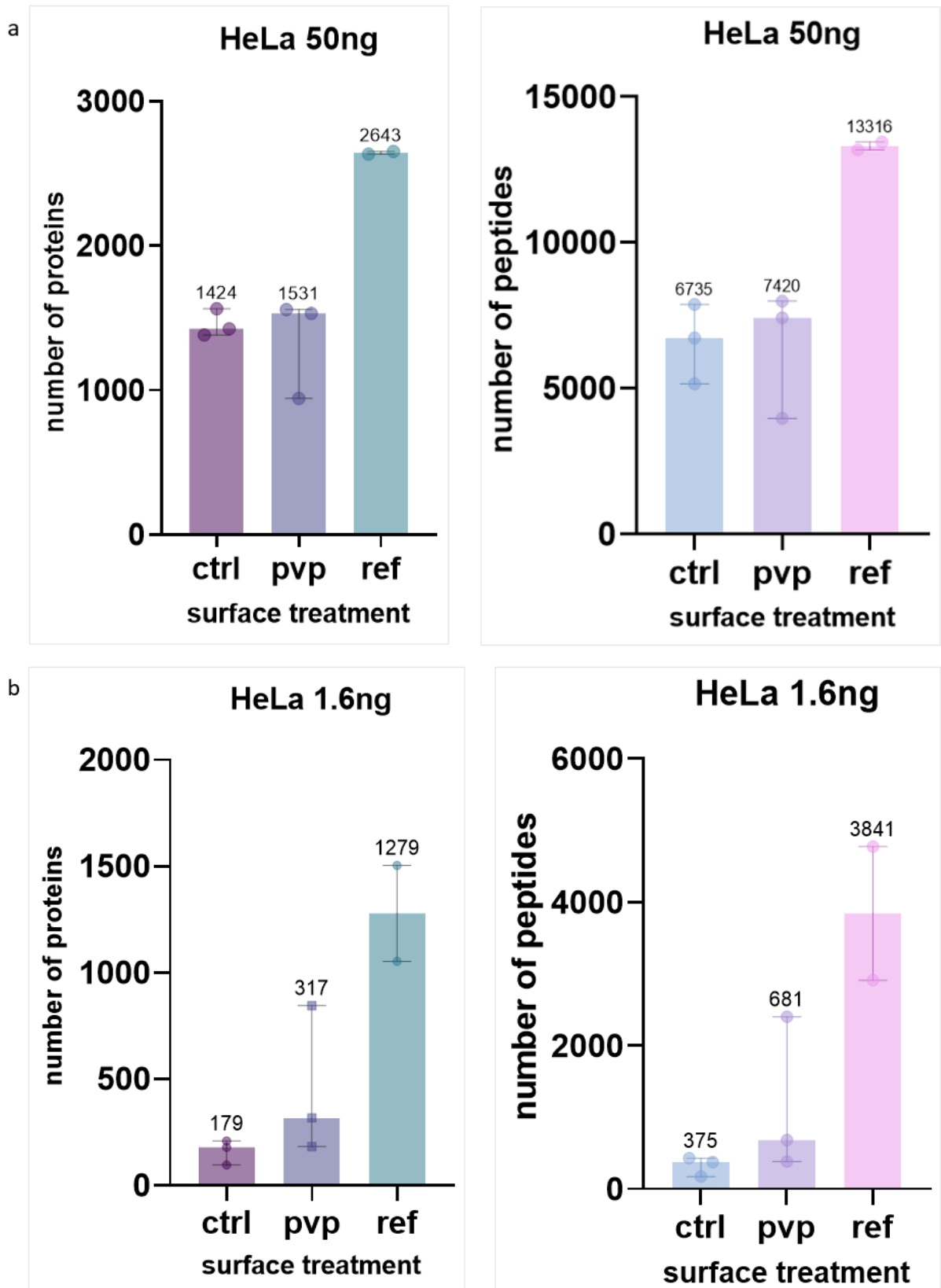


Figure 94 : Nombre de protéines, peptides identifiés à partir de (a) 50ng (b) 1.6ng de digestat de HeLa ayant traversé des puces en PDMS traitées (pvp) ou non (ctrl) au PVP comparée au nombre identifiés à partir de digestat de HeLa n'ayant pas traversé de puce (ref). La médiane des répliquats de préparation est indiquée au dessus de chaque barre d'intervalles de confiance à 95%.

blanc

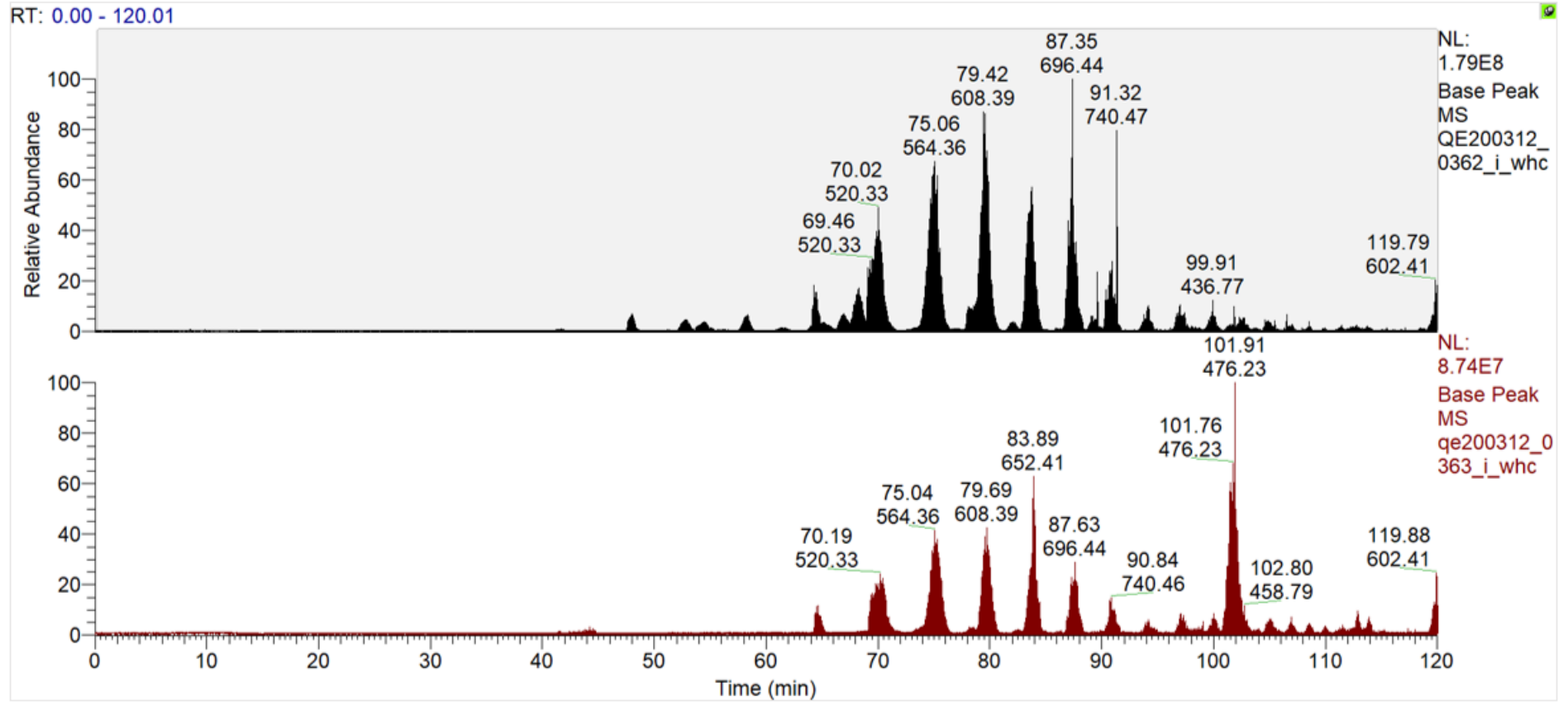


Figure 95 : Chromatogramme en pic de base d'un blanc dans une plaque micro-puits non traitée (en haut) ou dans une plaque micro-puits traitée au PVP (en bas). NL indique les intensités maximales de chaque chromatogramme.

# HeLa 1.6ng

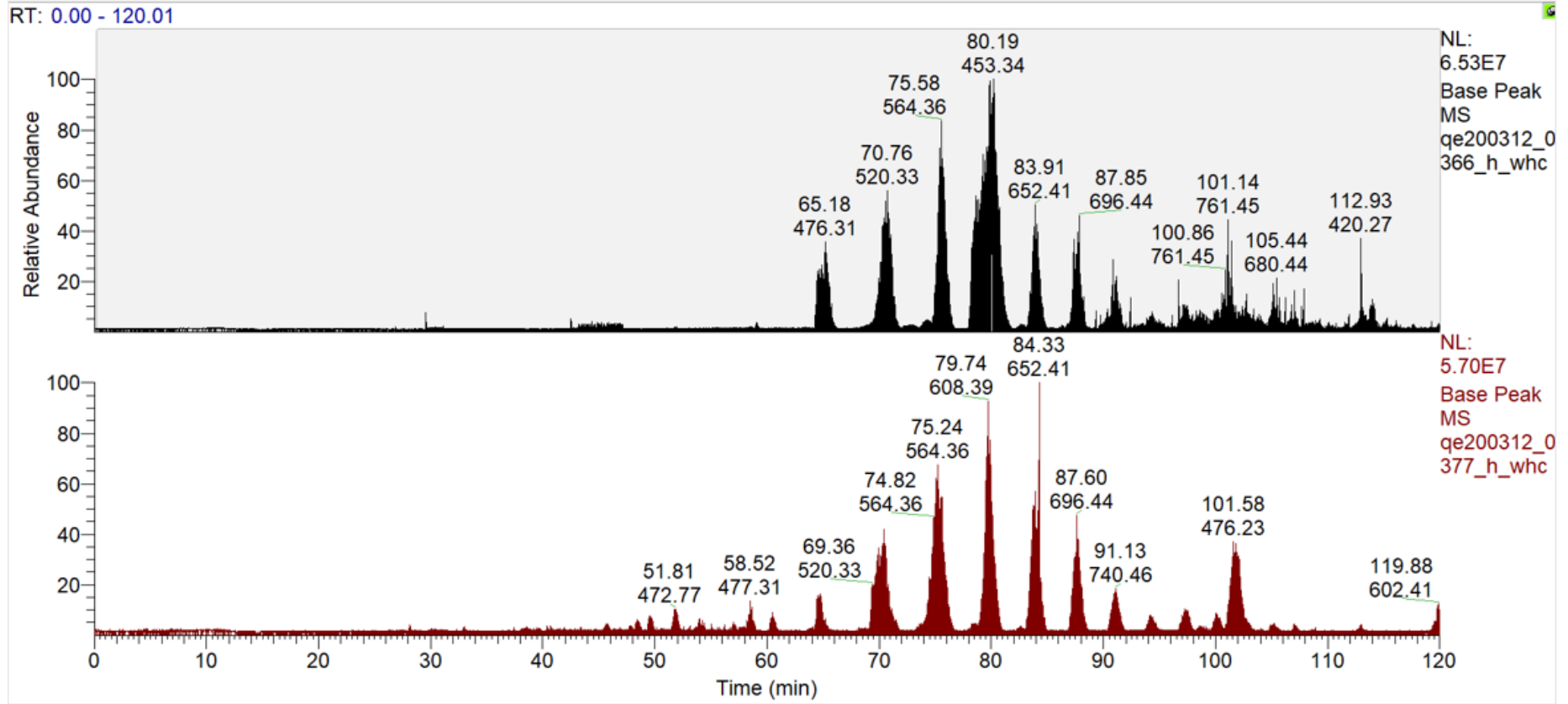


Figure 96 : Chromatogramme en pic de base d'un échantillon de HeLa (1.6ng) dans une plaque micro-puits non traitée (en haut) ou dans une plaque micro-puits traitée au PVP (en bas). NL indique les intensités maximales de chaque chromatogramme.

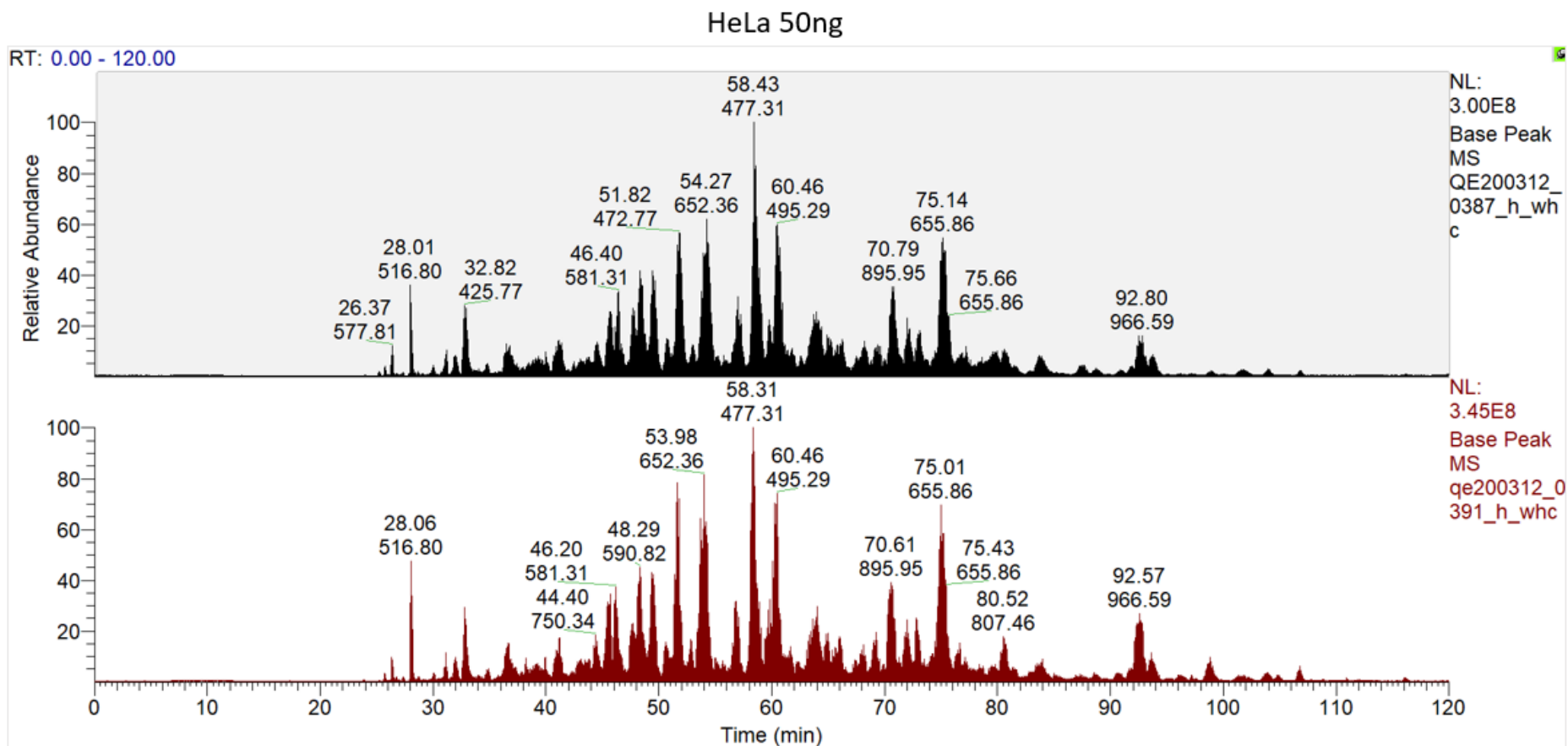


Figure 97 : Chromatogramme en pic de base d'un échantillon de HeLa (50ng) dans une plaque micro-puits non traitée (en haut) ou dans une plaque micro-puits traitée au PVP (en bas). NL indique les intensités maximales de chaque chromatogramme.



### 6.D.2. Le PVP est-il le meilleur traitement de surface ?

Comme discuté précédemment, les caractéristiques hydrophobes de surface du PDMS induisent une perte du matériel peptidique par adsorption. Divers protocoles de traitement hydrophile pour la surface du PDMS ont été sélectionnés dans la littérature afin de limiter cette adsorption, testés dans le cadre du stage de Mathéo La Torre encadré par le Dr. Massamba Ndiaye, et sont présentés dans le tableau (Tableau 9). Afin d'être applicables à notre setup expérimental, les différents traitements doivent pouvoir être réalisés à l'intérieur de canaux. La contrainte est de faire passer les solutions de traitement dans les canaux fermés, ce qui implique une modification des étapes de rinçage en remplaçant les rinçages sous agitation par des renouvellements des solutions dans les canaux. Les échantillons passant dans les dispositifs PDMS sont destinés à une analyse LC-MS. Les traitements doivent donc être stables et ne pas libérer de composés dus aux traitements et susceptibles de perturber l'analyse *en LC-MS*.

#### (a) Optimisation du setup expérimental

Le nombre de Reynolds permet de déterminer le régime d'écoulement en mécanique des fluides. Il permet de définir 3 états d'écoulement : un état laminaire caractérisé par un nombre de Reynolds faible, transitoire et turbulent pour des nombres de Reynolds plus élevés. Tous les dispositifs microfluidiques ont un écoulement laminaire à petit nombre de Reynolds. Il est important de créer des systèmes microfluidiques symétriques avec des zones de turbulence (pas de zones de stagnation) pour récupérer efficacement notre échantillon dans ces systèmes à petit nombre de Reynolds. Il faut également amplifier le contact entre les échantillons et le support afin de mieux observer le phénomène d'adsorption et des différents traitements testés.

Afin de travailler dans un dispositif offrant une surface de contact importante et des conditions fluidiques représentatives, les puces ont été prototypées sous la forme d'un réservoir (incubation et traitement des échantillons) encadré par deux circuits (transfert des échantillons dans des canaux microfluidiques). Les circuits amènent une surface de contact PDMS/liquide plus grande et les réservoirs permettent de faire passer dans la puce un volume suffisant pour être analysé en LC-MS et le réservoir en forme de prisme droit avec une base de trapèze isocèle permet de minimiser les zones de stagnation (Figure 98).

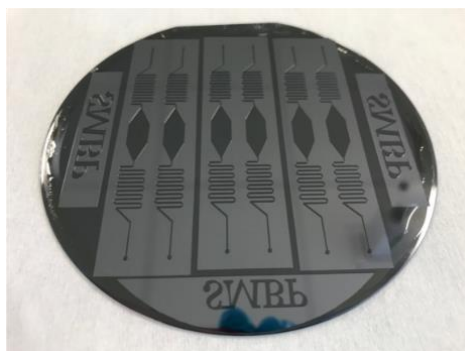


Figure 98 : Photographie du *wafer* du prototype 2.

Les caractéristiques géométriques des puces sont regroupées dans le tableau ci-dessous (Tableau 8).

## ETUDES ET RESULTATS : Importance du support

Dessin canaux	Profondeur	Surface	Volume total
1	70µm	192mm <sup>2</sup>	11µl
2	70µm	163mm <sup>2</sup>	9µl

Tableau 8 : Dimensions caractéristiques des dispositifs microfluidiques avec surface totale de contact PDMS-Solution et volume d'un canal arrondi à l'inférieur.

### Méthode

Chaque dispositif microfluidique a été rempli avec 9µl (volume total des canaux) de solution de HeLa à 10ng/µl. L'ensemble est resté en incubation pendant 1 heure puis la solution a été prélevée à la sortie du canal pour analyse sur un Tims TOF Flex. Les analyses ont été effectuées en triplicat.

### Résultats

Le nombre de peptides identifiés à partir d'une solution de digestat de cellules HeLa à 10 ng/µl et à partir de la même solution ayant été en contact pendant 1 heure avec la puce en PDMS est montré sur la Figure 99. Les résultats indiquent une baisse de 53% du nombre de peptides, passant de 8838 dans la solution de référence à 4039 peptides en sortie de PDMS. Ces résultats sont en concordance avec ce qui a été observé dans notre essai préliminaire.

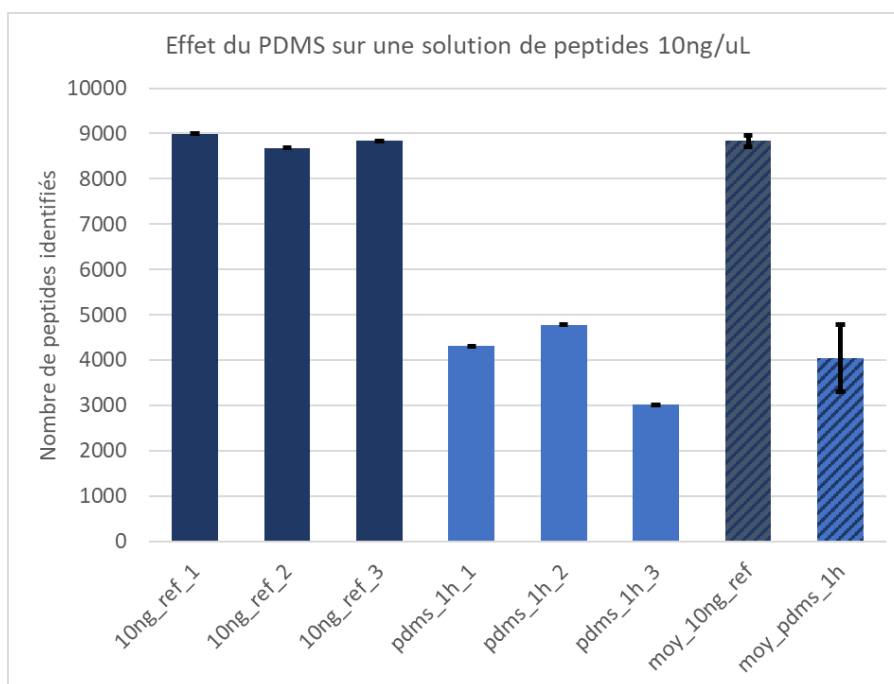


Figure 99: Étude de l'effet de la présence d'une surface de PDMS en contact avec une solution de digestat tryptique d'extrait protéique total de cellule HeLa à 10ng/µl. Les analyses ont été réalisées en triplicat pour la solution test (bleu sombre, 10ng\_ref\_1, 2 et 3) et pour la solution qui est passée dans le dispositif en PDMS (bleu clair, pdms\_1h-1, 2 et 3). La moyenne des résultats de chaque série est présentée en hachuré

## ETUDES ET RESULTATS : Importance du support

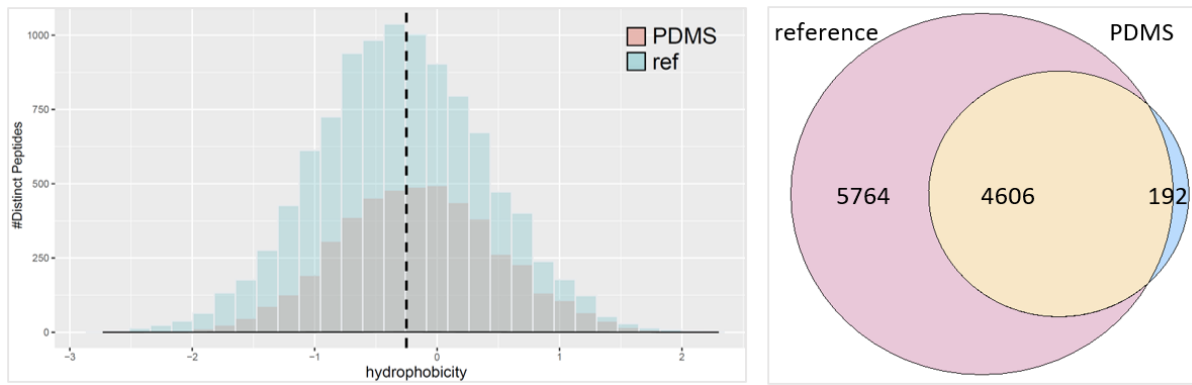


Figure 100: Analyse différentielle des identifications de peptides après mise en contact avec une surface PDMS. (À gauche) : Distribution des peptides selon leur indice d'hydrophobicité (échelle Kyte-Doolittle). Plus un peptide est hydrophobe, plus cet indice est grand. En vert, les peptides identifiés à partir de la solution de référence (10ng/μl), en orange, les peptides identifiés après avoir séjourné dans le dispositif microfluidique en PDMS pendant 1h. (à droite) : Étude des peptides identifiés en commun en absence (Reference) et en présence d'une surface PDMS (PDMS).

La Figure 100 indique le nombre de peptides identifiés en fonction du caractère hydrophobe du peptide. Cette figure confirme la réduction de la moitié du nombre de peptide reconnus après passage sur le PDMS. Il apparaît cependant que les peptides hydrophiles sont ceux qui sont les moins bien détectés, ce qui ne peut pas s'expliquer directement par le caractère hydrophobe de la surface du PDMS. Ce résultat est corroboré par la distribution des temps de rétention obtenus pour ces analyses comme illustré dans la Figure 101. De plus, l'analyse des peptides identifiés en commun est illustrée par le diagramme de Venn (Figure 100). Seul 1.8% des peptides ne sont vus de façon reproductible qu'après passage sur PDMS, soit 192 séquences uniques sur 10562 au total.

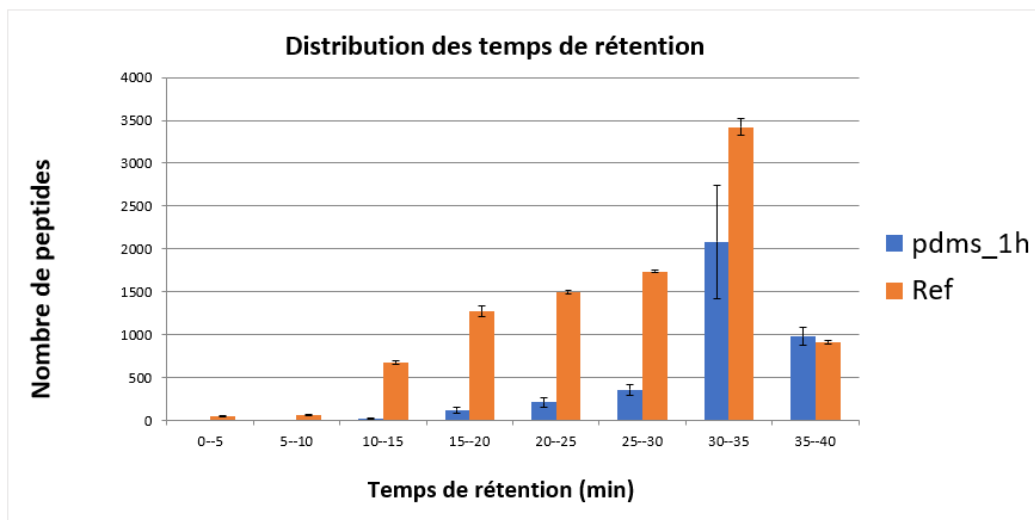


Figure 101 : Distribution des temps de rétention des peptides identifiés dans la solution de référence (orange) et dans la solution au contact du PDMS (bleu).

D'autres propriétés pourraient ainsi intervenir et jouer un rôle important dans l'adsorption des peptides, comme la polarité, la formation de liaisons hydrogènes ou des effets électrostatiques avec les tri-méthyl-silanes et les silanols en surface. Ces effets mériteraient

d'être investigués plus en détail par traitement statistique, en tenant compte du pH de la solution (pH=5).

De plus il est possible d'évaluer la variabilité induite par l'utilisation de différents dispositifs en alignant les analyses deux par deux. L'analyse par diagrammes de dispersion (*multi-scatter plots*) met en évidence plusieurs phénomènes (Figure 102). Le premier est la très grande reproductibilité de l'analyse, ce qui correspond à l'alignement des triplicatas obtenus sur l'échantillon de 10ng/μl de référence (ref\_1 à 3) avec un coefficient de corrélation très élevé ( $\geq 0.985$ ). Le second est la moins bonne reproductibilité des échantillons passés sur PDMS, et notamment en ce qui concerne l'échantillon PDMS\_3. Cette variabilité peut être imputée à des effet induits par le traitement similaire mais distinct des aliquots d'un même échantillon dans trois dispositifs. Le troisième est issu de la distribution des points par rapport à la première bissectrice, qui représente le cas idéal d'une reproductibilité parfaite (qu'on retrouverait sur la diagonale vide où chaque échantillon est comparé à lui-même). La distribution est symétrique quand deux échantillons de même nature sont comparés. Elle est asymétrique quand on compare un échantillon PDMS à un échantillon référence, avec majoritairement des peptides identifiés en commun avec une plus grande intensité en PDMS. En effet, en regardant l'agrandissement, il est clair que l'ensemble des points est symétrique selon un axe quasi parallèle et au-dessus de la première bissectrice.

Il y a donc moitié moins de peptides identifiés mais ils sont détectés avec une intensité supérieure, ce qui à nouveau ne semble pas refléter directement la variation d'abondance absolue. Une hypothèse est que la plus faible complexité de l'échantillon au bénéfice d'espèces a priori plus hydrophobes fait chuter les effets de suppression et de saturation et permet d'augmenter le rapport signal/bruit de l'analyse, rappelant ici que l'analyse par spectrométrie de masse n'est pas une analyse quantitative *per se*.

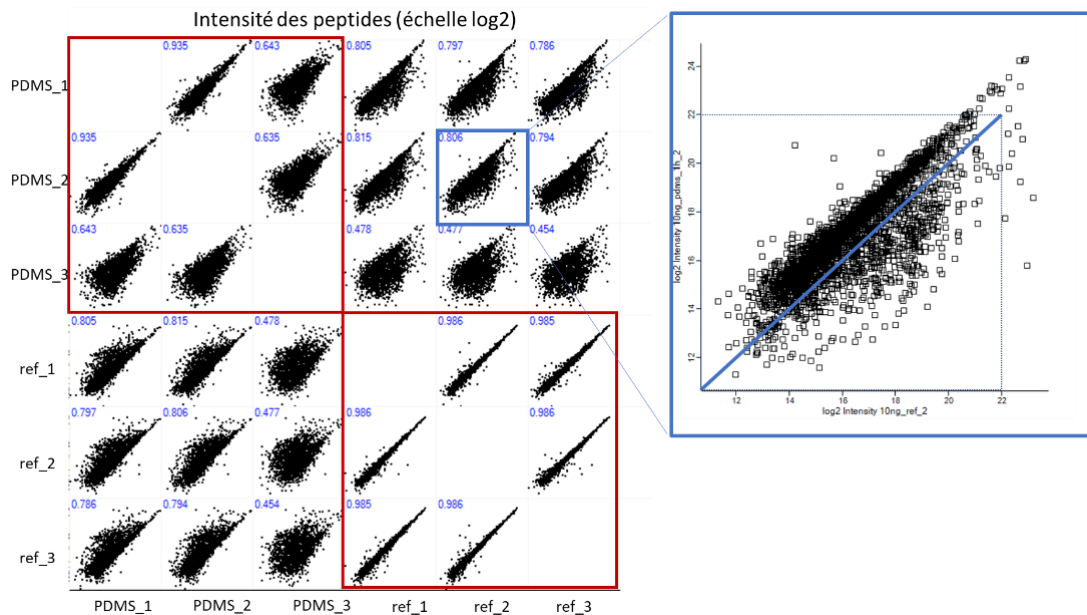


Figure 102: Comparaison des analyses deux à deux. Chaque point correspond à un peptide représenté en échelle d'intensité pour les deux conditions comparées. Pour chaque couple, le coefficient de corrélation de Pearson est indiqué. Le cadre en haut à gauche correspond aux analyses sur PDMS, le cadre en bas à droite correspond aux analyses de l'échantillon contrôle. À droite, l'agrandissement d'une des comparaisons montre la position du nuage de points par rapport à la première bissectrice, cas idéal de reproductibilité et d'une corrélation parfaite.

## ETUDES ET RESULTATS : Importance du support

Ce nouveau prototype confirme la perte de la moitié des peptides après adsorption par contact des peptides sur la surface du PDMS. Il sera utilisé pour vérifier l'apport du traitement au PVP mais aussi d'autres traitements de surface trouvés dans la littérature qui seront comparés à l'échantillon de référence n'ayant pas été en contact avec le PDMS.

### (b) Évaluation des différents traitements de surface

Une recherche bibliographique sur les divers traitements de surface du PDMS compatibles avec notre setup expérimental a amené à sélectionner 8 possibilités, reprises dans le tableau ci-dessous (Tableau 9).

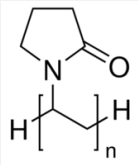
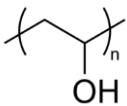
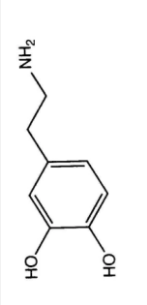
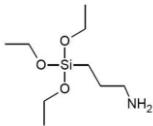
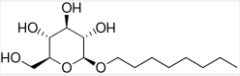
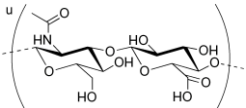
Structure	Traitement	Activation	Réaction	Lavage
	<b>PVP</b> , <sup>131</sup>	Plasma d'oxygène	PVP 0.2%, 10min	Eau mQ, 1mL
	<b>PVA</b> , <sup>137</sup>	H <sub>2</sub> O :H <sub>2</sub> O <sub>2</sub> :HCl 5 :1 :1, 35min	PVA 1%, 10min	Eau mQ, 1mL
	<b>PD1</b> , <sup>210</sup>	H <sub>2</sub> O :H <sub>2</sub> O <sub>2</sub> :HCl 5 :1 :1, 45min	Dopamine 5mg/ml, Tris-HCl 50mM, pH=8.8, 72h	Eau mQ, 3mL, 72h
	<b>PD2</b>	NaIO <sub>4</sub> 20mM dans solution de réaction	Dopamine 2mg/ml, AcONa 50mM, pH=5, 24h	Eau mQ, 3mL, 72h
	<b>PD3</b>	O <sub>2</sub> dissous	Dopamine 5mg/ml, Tris-HCl 50mM, pH=8.8, 72h	Eau mQ, 2mL, 24h + 48h à 37°C
	<b>APTES</b> , <sup>211</sup>	H <sub>2</sub> O :H <sub>2</sub> O <sub>2</sub> :HCl 5 :1 :1, 20min	APTES 5% dans chloroforme, 24h	EthOH, 2mL + eau mQ, 2mL
	<b>Bêta-octyl</b>	H <sub>2</sub> O :H <sub>2</sub> O <sub>2</sub> :HCl 5 :1 :1, 20min	Bêta-octyl 1%, Tris-HCl 50mM, pH=8.8, 24h	Eau mQ, 1mL
	<b>HA</b> , <sup>212</sup>	H <sub>2</sub> O :H <sub>2</sub> O <sub>2</sub> :HCl 5 :1 :1, 20min	HA 1%, 20min HA 0.1% 24h	Eau mQ, 1mL

Tableau 9 : Liste des traitements de surface PDMS évalués. PVP : polyvinylpyrrolidone, PVA : Poly-acétate de vinyle, PD : Polydopamine, APTES : (3-Aminopropyl)triethoxysilane, Bêta-octyl : bêta-octyl-glucoside, HA : Hyaluronic acid.

## ETUDES ET RESULTATS : Importance du support

### Méthode

Les protocoles détaillés des traitements de surface sont repris dans la partie matériels et méthodes. Pour chaque traitement, la procédure pour évaluer son efficacité est identique. Une solution de peptides protéolytiques issus du digestat de référence de protéines de cellules HeLa solubilisées selon les recommandations du fournisseur, a été reprise à 10ng/μl. Cette solution a été incubée dans les canaux en PDMS (1h) à température ambiante, avant d'être récupérée sous pression douce (à la seringue) dans un tube en verre (volume récupéré ~ 9μl, volume injecté = 2μl), pour analyse sur le Tims TOF Flex. Les acquisitions n'ont pas cherché l'exhaustivité mais la reproductibilité des conditions expérimentales en minimisant les délais entre les diverses acquisitions. Les résultats peuvent être analysés en termes de comparaison du nombre de peptides identifiés et de la distribution de l'index d'hydrophobicité/distribution temps de rétention.

### Résultats

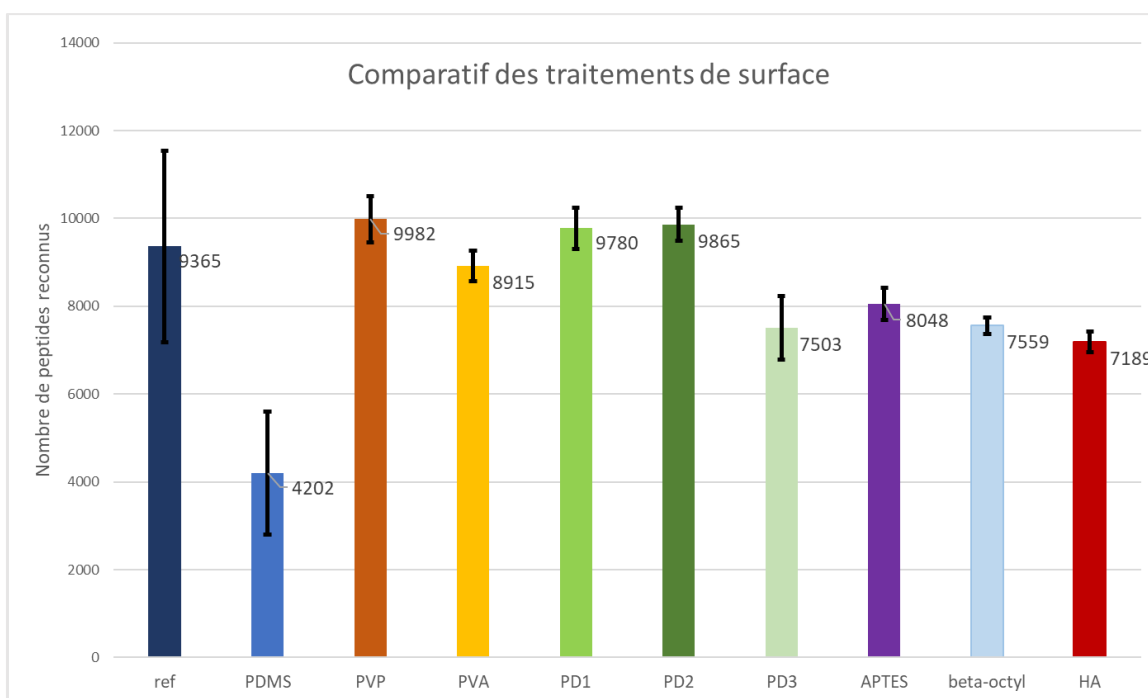


Figure 103 : Évolution du nombre moyen de peptides détectés après passage dans un dispositif en PDMS selon le traitement de surface. L'échantillon de référence à 10ng/μl (ref) a été injecté dans des dispositifs en PDMS non traités (PDMS) ou traités (PVP, PVA, PD1, PD2, PD3, APTES, bêta-octyl, HA). Les chiffres indiquent le nombre moyen de peptides identifiés et passant les filtres de validation sur l'ensemble des analyses réalisées, indépendamment des aliquots. Ceci explique la plus grande variabilité des résultats ref et PDMS qui sont moyennés sur l'ensemble des campagnes d'analyse pour tous les traitements.

En considérant le nombre de total de peptides identifiés selon chacune des conditions (un minimum de 3 traitements identiques distincts a été fait par condition), tous les traitements testés améliorent la récupération peptidique (Figure 103). Les traitements au PVP, PVA et à la PD (en présence de H<sub>2</sub>O<sub>2</sub> ou de NaIO<sub>4</sub>) apparaissent les plus performants. En effet, le nombre total de peptides identifiés par ces traitements n'est pas significativement différent de celui obtenu en moyenne à partir de la solution de référence de digestat de 10ng d'extrait protéique total de HeLa.

## ETUDES ET RESULTATS : Importance du support

En revanche en considérant le nombre de peptides identifiés de façon reproductible sur plus de la moitié des dispositifs traités de manière identique, les traitements PD3, APTS HA et bêta-octyl sont bien moins performants. Ce traitement de données est détaillé dans la suite sur les histogrammes qui permettent aussi de mettre en évidence la distribution de la population de peptides identifiés de façon reproductible en fonction de leur indice selon Kyte-Doolittle, qui indique une estimation du caractère hydrophobes des espèces. Il est aussi intéressant de regarder le recouvrement des populations de peptides identifiés de façon reproductibles, afin de voir si la modification des surfaces des contenants exerce une influence sur l'identité des peptides identifiés de façon spécifique. Ceci est illustré dans la suite par les diagrammes de Venn pour chaque condition. Ces diagrammes mettent en évidence que sans exception, certains peptides ne sont identifiés que dans les échantillons ayant été en contact avec le PDMS traités, alors qu'ils n'étaient pas identifiés systématiquement dans les échantillons de référence.

Polydopamine tamponnée (pH = 8.8) sans oxydation

Ce traitement est noté PD3. Seuls 45% des peptides identifiés dans l'échantillon de référence sont détectés après traitement PD3 (Figure 104). La distribution des peptides sur l'échelle de Kyte et Doolittle montre une répartition légèrement décalée des peptides en comparaison avec le PDMS, avec une augmentation relative des peptides les moins hydrophobes, ce qui recentre la distribution en PD3 vers la distribution de référence. Il faut cependant noter que seulement 44% des peptides sont identifiés en commun, ce qui donne un taux d'identification proche du cas du PDMS non traité. Ce traitement n'a donc pas d'intérêt pour nos applications en protéomique *bottom-up*.

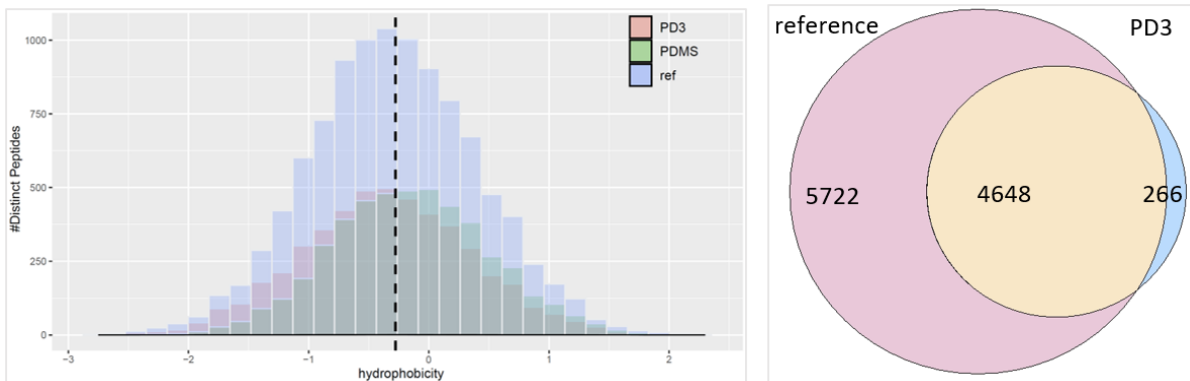


Figure 104 : Effet du traitement de la surface du PDMS à la polydopamine sans traitement oxydant, en présence d'O<sub>2</sub> (PD3) sur la population des peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (PD3) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence.

Polydopamine tamponnée (pH = 8.8) avec oxydation préliminaire

Ce traitement est noté PD1. Il est réalisé au même pH que précédemment, mais une première étape d'oxydation préliminaire au peroxyde, toute conditions étant égales par ailleurs avec les conditions PD3. Dans ces conditions, le nombre de peptides identifiés augmente (Figure 105). Environ 68% des 11543 peptides identifiés sont communs aux deux conditions. Il faut noter que 1173 peptides (10%) sont identifiés alors qu'ils ne l'étaient pas à partir de l'échantillon de référence. De plus 22% des peptides identifiés en référence, ne sont pas validés sur l'échantillon PD1. Ce traitement est donc assez efficace pour limiter les pertes

## ETUDES ET RESULTATS : Importance du support

de peptides observées par adsorption sur la surface PDMS. Il ne semble pas induire de biais sur les caractéristiques d'hydrophobie des peptides, et corrige le décalage de la distribution vers les peptides moins hydrophobes.

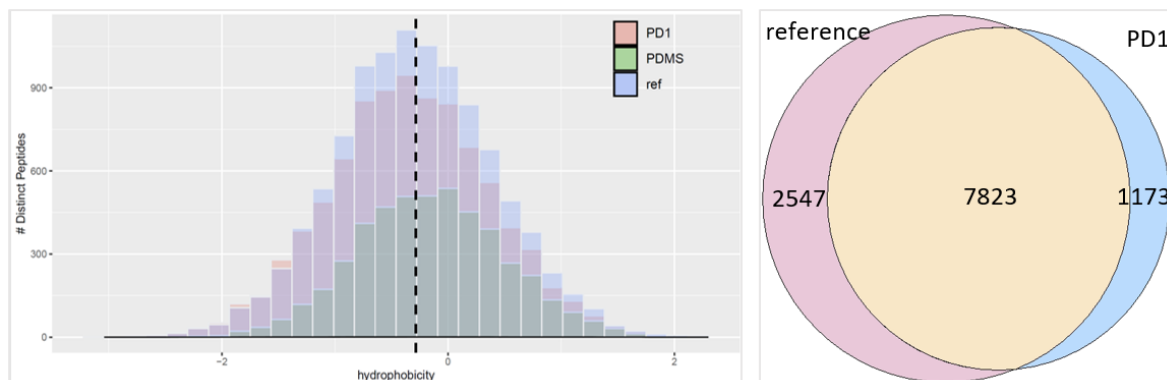


Figure 105 : Effet du traitement de la surface du PDMS à la polydopamine avec traitement oxydant au peroxyde et HCl concentré (PD1) sur les peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (PD1) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence.

Polydopamine tamponnée (pH = 5) avec oxydation

Ce traitement est noté PD2. Ce traitement est réalisé en condition acide après activation au sodium periodate,  $\text{NaIO}_4$ , un agent oxydant. Les résultats sont meilleurs que ceux obtenus avec le traitement PD1. Nous avons une couverture des peptides relativement équivalente, avec la récupération d'espèce moins hydrophobes, bien que la qualité des analyses semble légèrement inférieure avec ce lot d'échantillon (Figure 106). Au total 13067 peptides uniques sont identifiés, 62% sont communs, 18% ne sont validés que dans l'échantillon de référence, et 20,5% ne sont identifiés que dans l'échantillon en contact de la surface PDMS\_PD2, ce qui est surprenant. Cela souligne que les conditions d'analyse de l'échantillon de référence ne sont pas à même de couvrir tout l'échantillon, ce qui pourrait être expliqué notamment par un défaut important de capacité de pic de l'analyse nano LC-MS/MS, et par un effet de suppression de signal plus important dans le cas de l'échantillon de référence de composition plus complexe.

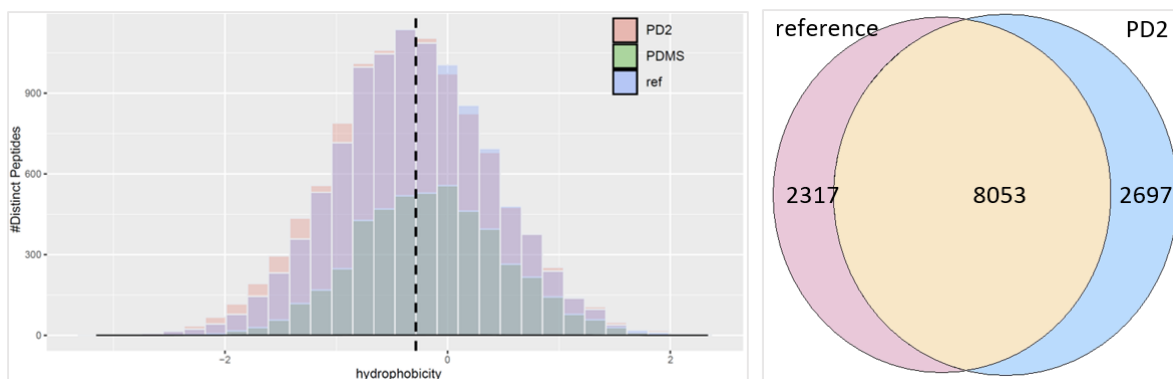


Figure 106 : Effet du traitement de la surface du PDMS à la polydopamine avec traitement oxydant au périodate de sodium sur la population des peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (PD2) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence.



## ETUDES ET RESULTATS : Importance du support

### Traitement par le PVA

Le traitement par le PVA a semblé relativement efficace. Sur les 12584 peptides identifiés, 64% sont identifiés en commun, 18% ne sont vus que dans les échantillons de référence, et 17.5% ne sont identifiés que dans l'échantillon en contact avec la surface PDMS-PVA. Les résultats sont donc comparables à ceux obtenus avec le traitement PD2. Au vu des résultats obtenus avec d'autres traitements qui ne nécessitent pas de travailler en condition acide concentrée, sa mise en œuvre n'a pas été poursuivie.

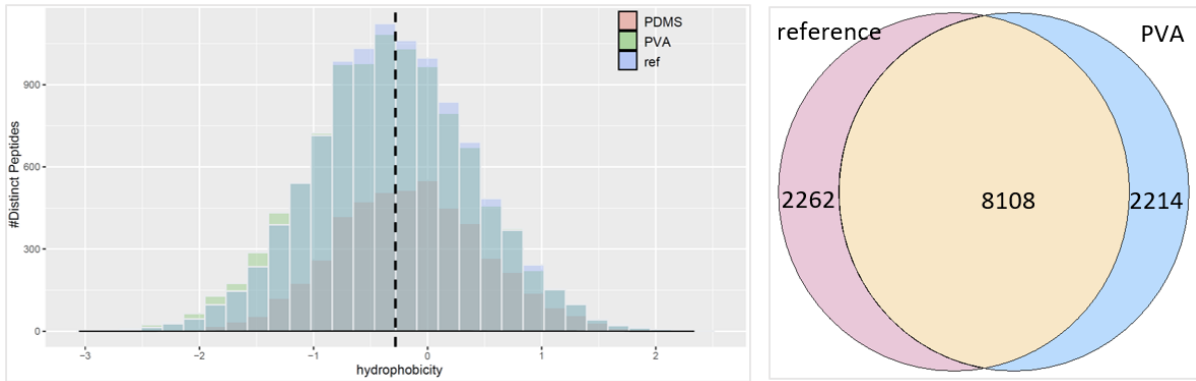


Figure 107 : Effet du traitement de la surface du PDMS au polyacétate de vinyle après activation avec traitement oxydant au peroxyde et HCl concentré (PVA) sur la population des peptides identifiés sans passage (ref) ou avec passage dans le dispositif microfluidique traité (PVA) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence

### Traitement par le PVP

Le traitement au PVP est efficace. Il corrige parfaitement la distribution des peptides qui se superpose avec la population originale. Le nombre d'identification est sensiblement équivalent (voir les estimations d'incertitude dans la Figure 108).

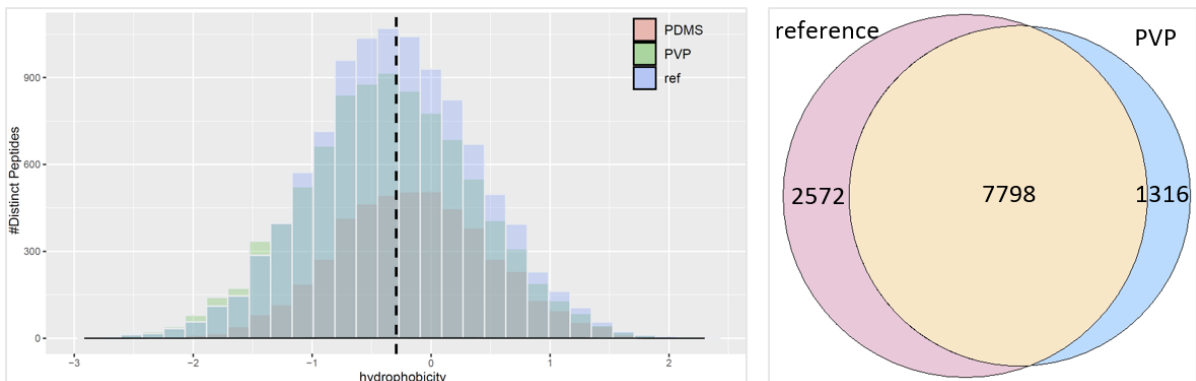


Figure 108 : Effet du traitement de la surface du PDMS au polyvinylpyrrolidone PVP après activation au plasma O<sub>2</sub> (PVP) sur la population des peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (PVP) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence

## ETUDES ET RESULTATS : Importance du support

### Traitement par l'APTES

Le traitement semble moins efficace en moyenne, ici il n'a pas permis d'atteindre une identification significativement comparable à l'échantillon de référence même si le nombre de peptides identifiés est plus élevé qu'en PDMS (Figure 109). Il semble être moins favorable aux espèces plus hydrophobes. De plus les conditions opératoires pour sa mise en œuvre sont moins douces, puisque l'activation de fait en milieu acide HCl concentré suivi par une incubation de 24h dans le chloroforme. Il n'a donc pas été retenu pour la suite de l'étude.

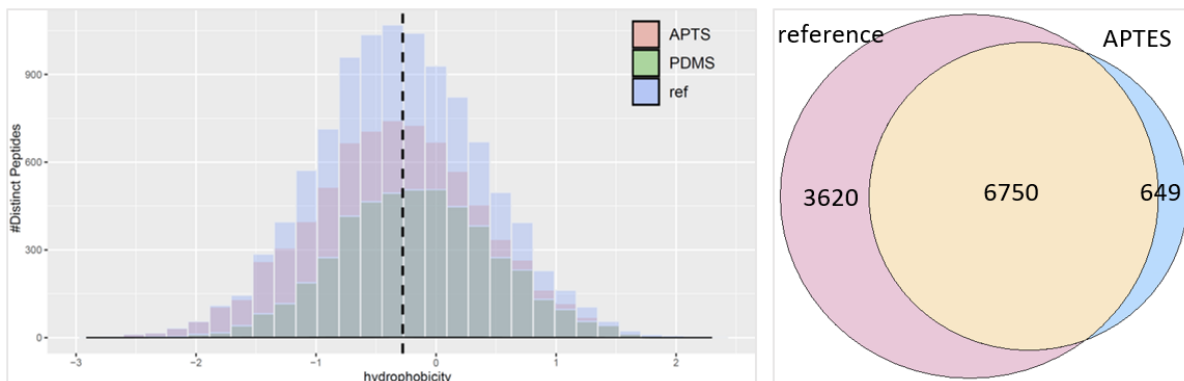


Figure 109 : Effet du traitement de la surface du PDMS au (3-aminopropyl) triéthoxysilane (APTES) après activation au plasma O<sub>2</sub> sur la population des peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (APTES) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence

### Traitement par le $\beta$ - octyl

Le traitement est efficace, mais les résultats de cette campagne d'analyse ne répondent pas aux mêmes critères de contrôle qualité avec des valeurs d'identification en deçà des valeurs moyennes, même pour l'analyse de l'échantillon de référence. La qualité d'analyse et de couverture d'échantillon semble donc comparable entre les deux conditions comme indiqué dans le diagramme de suivi du nombre de peptides identifiés et validé pour chaque analyse (Figure 110).

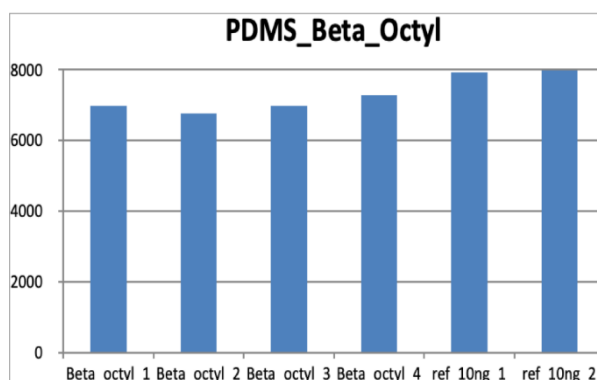


Figure 110 : Évolution du nombre de peptides identifiés pour chaque enregistrement avec ou sans mise en contact avec une surface PDMS- bêta-octylglucoside.

En revanche, une analyse globale comparative toutes campagnes d'analyse confondues montre la faiblesse des performances analytiques particulière pour cette série (Figure 111). L'histogramme permet de montrer la meilleure couverture des espèces peu hydrophobes

## ETUDES ET RESULTATS : Importance du support

cependant. Au vu des conditions de traitement plus exigeantes, ce traitement n'a pas été considéré plus avant car il n'apporte pas de meilleurs résultats que PD2, PVP et PVA.

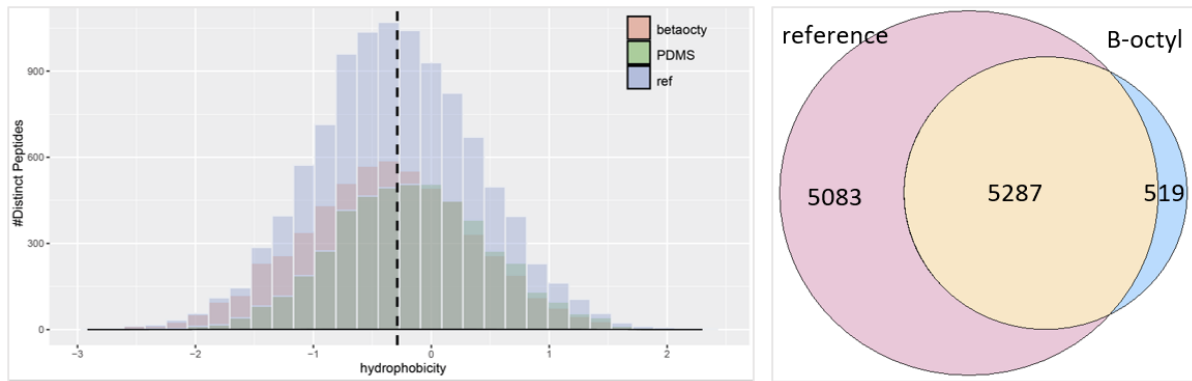


Figure 111: Effet du traitement de la surface du PDMS au bêta-octylglucoside (betaoctyl) après activation  $H_2O_2$ , HCl concentré sur la population des peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (betaoctyl) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence

Traitements par l'acide hyaluronique (HA)

Le traitement HA produit un effet limité et donne des résultats comparables à ceux obtenu avec le bêta-octylglucoside. Pour les mêmes raisons que pour ce dernier, le traitement nécessaire moins favorable et plus exigeant en termes de mise en œuvre nous a conduit à ne pas aller plus avant dans son étude.

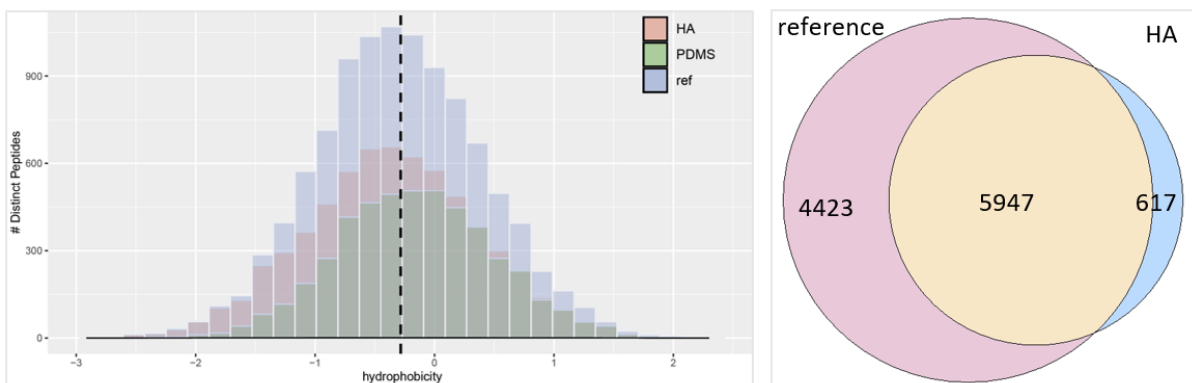


Figure 112 : Effet du traitement de la surface du PDMS à l'acide hyaluronique (HA) après activation  $H_2O_2$ , HCl concentré sur la population des peptides identifiés dans l'échantillon référence sans passage (ref) ou avec passage dans le dispositif microfluidique traité (HA) ou non traité (PDMS). À gauche, répartition des peptides selon leur indices d'hydrophobie ; à droite recouvrement des peptides identifiés dans un même lot après traitement et dans l'échantillon de référence.

### Bilan

Les traitements hydrophiles exercent une influence certaine sur le nombre de peptides reconnus et corrigent vraiment l'adsorption du PDMS. Certains semblent même annuler cette adsorption. Il faut ici souligner que les résultats montrent que l'analyse est loin de couvrir efficacement l'échantillon puisque même lorsqu'un nombre de peptides équivalent est retrouvé, il n'y a pas plus de 70% des séquences identifiées de façon reproductible en

## ETUDES ET RESULTATS : Importance du support

commun. Les rapides tests réalisés n'ont pas permis de détecter ou d'identifier un paramètre physicochimique qui serait discriminant dans le phénomène d'adsorption.

Pour comparer les rendements de chaque traitement, on peut dire que parmi les 6 protocoles issus de la littérature, les traitements PVP, PVA, PD1 et PD2 sont les plus prometteurs. Parmi les deux traitements conçus au laboratoire, c'est celui au  $\beta$ -octyl qui produit le meilleur rendement, mais leur efficacité est inférieure. Outre l'efficacité du traitement, il y a deux paramètres importants. Le premier est la mise en œuvre. Certains traitements sont particulièrement longs et/ou difficiles à mettre en œuvre. Parmi les traitements à très bon rendement, celui au PVP est de loin le plus rapide et facile. Le second paramètre est la stabilité et la neutralité du traitement. C'est l'ensemble de ces paramètres qui doit déterminer le choix du traitement optimal. Et il n'est peut-être pas le même en fonction des applications.

### Cas particulier du PVP

Il n'y a pas de biais apparent significatif sur les distributions des peptides en fonction de leur paramètres physicochimiques. Parmi les 11686 séquences comparées, 67% (7798 peptides) sont détectées en commun (Figure 108).

Une étude complémentaire de la distribution des peptides selon les paramètres définis par Cruciani permet de considérer les paramètres : PP1 pour la polarité, PP2 pour l'hydrophobicité (une variante de l'index de Kyte Doolittle), et PP3 pour la capacité à former des liaisons hydrogènes<sup>213</sup>. Les peptides spécifiquement récupérés après traitement PVP sur PDMS pourraient peut-être être légèrement plus polaires et moins hydrophobes que la statistique appliquée à l'échantillon de référence, comme illustré sur la Figure 113 : Évolution de la distribution des peptides selon les paramètres physicochimiques définis par Cruciani<sup>213,214</sup>. L'axe des x représente des échelles PP1 : polarité, PP2 : Hydrophobicité et taille, PP3 : capacité à former des liaisons-H.. Le paramètre PP3 ne donne pas de tendance significative.

La corrélation entre les différents canaux traités est beaucoup plus forte que dans le cas du PDMS non traité (Figure 114). Les coefficients de corrélation sont tous supérieurs à 0.985, ce qui est à comparer avec la Figure 102 obtenue dans le cas du PDMS non traité au sein d'une même campagne analytique. Les traitements semblent donc améliorer la reproductibilité des expériences de mise en contact des peptides avec les surfaces des canaux microfluidiques. On peut donc émettre l'hypothèse que c'est le phénomène d'adsorption qui faisait chuter la reproductibilité. Le traitement a donc un effet complémentaire au contrôle de l'adsorption en normalisant les états de surface pour améliorer la reproductibilité des analyses.

Il existe bien un phénomène d'adsorption des peptides dans les puces en PDMS. Cependant, des traitements hydrophiles de la surface du PDMS sont possibles et permettent de réduire fortement l'adsorption. L'utilisation de ces traitements permet aussi une meilleure reproductibilité des expériences menées dans les puces.

Les traitements permettent tous d'obtenir un bon (voir très bon) rendement en nombre de peptides reconnus. Ils sont tous compatibles avec la méthode d'analyse LC-MS des peptides (pas de contamination). En comparant le rendement ainsi que la durée et la facilité des traitements, certains peuvent être écartés comme l'APTES (très difficile à mettre en œuvre) ou la PD3 (très long et à faible rendement). Le traitement au PVP reste le plus rapide

## ETUDES ET RESULTATS : Importance du support

et le plus propre (grâce à l'activation par plasma d'oxygène) des traitements testés et permet un rendement de récupération efficace des peptides.

Mais il reste des expériences à mener pour trancher définitivement. En effet, la longévité du traitement et son efficacité à long terme sont des facteurs importants pour le choix.

Pour la suite du projet, la vérification s'est concentrée sur le traitement PVP fait extemporanément de nos expériences et à son apport sur les différentes étapes du protocole de préparation des échantillons.

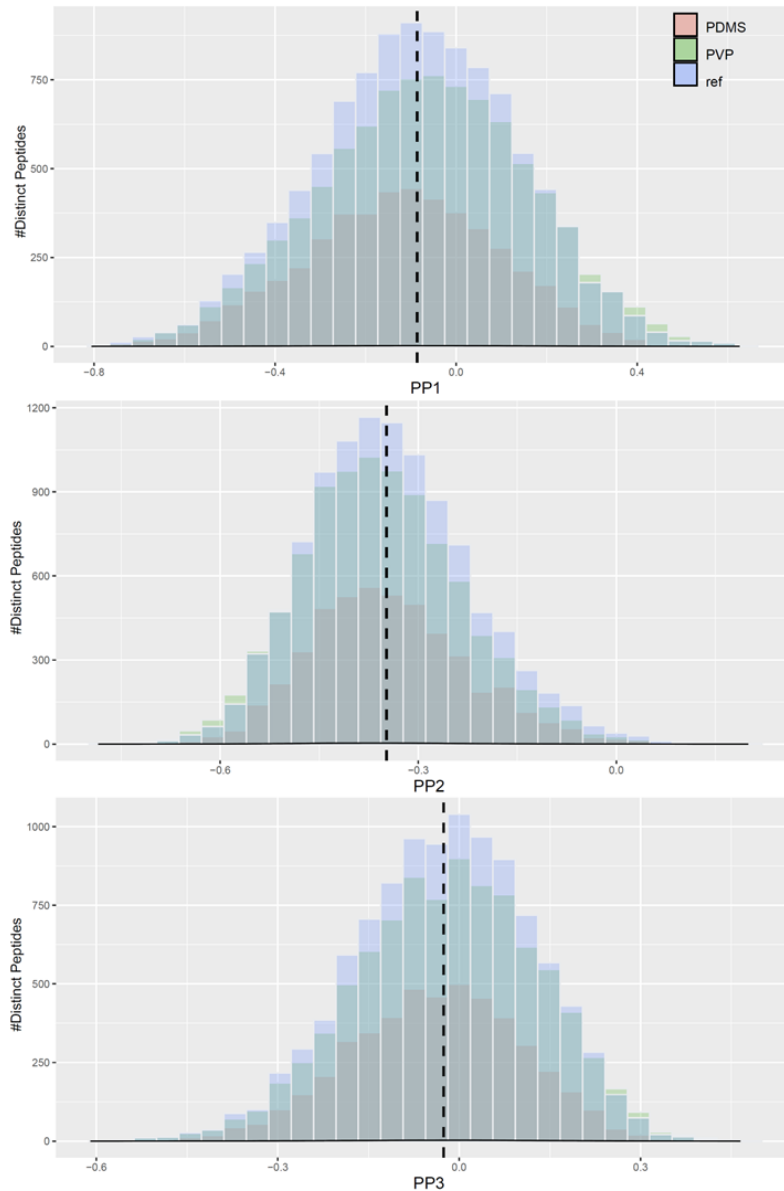


Figure 113 : Évolution de la distribution des peptides selon les paramètres physicochimiques définis par Cruciani <sup>213,214</sup>. L'axe des x représente des échelles PP1 : polarité, PP2 : Hydrophobicité et taille, PP3 : capacité à former des liaisons-H.

## ETUDES ET RESULTATS : Importance du support

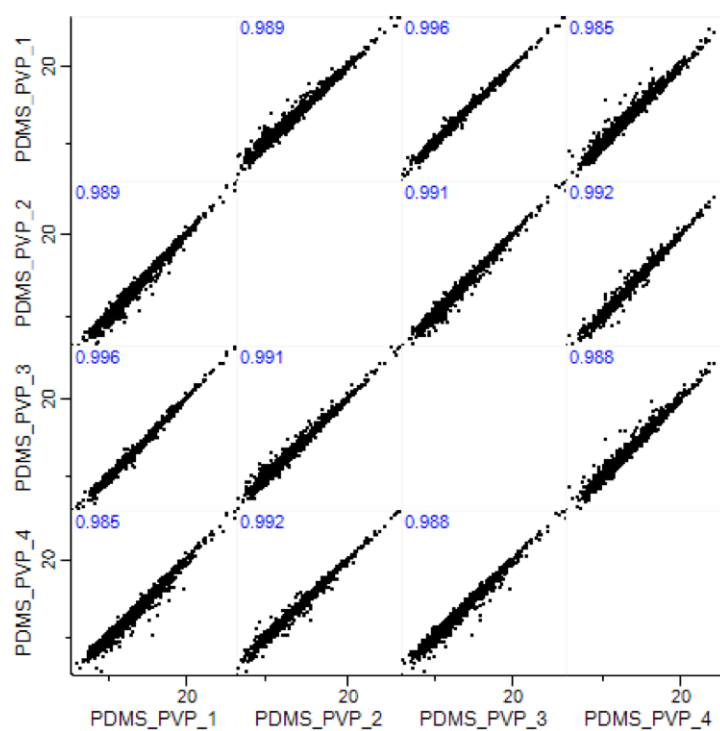


Figure 114 : Diagramme de dispersion des analyses dans les canaux PDMS traités au PVP. Les intensités obtenues pour chaque peptide sont comparées enregistrement par enregistrement. Le coefficient de corrélation de Pearson est indiqué en haut à gauche de chaque comparaison.

### 6.D.3. Le PVP est-il compatible avec la protéolyse des échantillons ?

La compatibilité du traitement de surface au PVP avec une protéolyse d'extraits protéiques de levure a été vérifiée.

#### Méthode

Sur une plaque micro-puits en PDMS (PDMS 10 $\mu$ l) traitée ou non au PVP, des aliquots d'1 $\mu$ l de 50ng de protéines extraites de cellules de levure *Saccharomyces cerevisiae* (Promega) ont été digérées par ajout d'1 $\mu$ l de trypsine-Lys C (ratio 1:20 (E/S)). L'équivalent de 10ng de ces échantillons a été analysé sur un Q Exactive HF. La médiane des réplicats est annotée au-dessus des barres d'intervalles de confiance à 95% de chaque condition sur la Figure 115. Neuf réplicats réalisés sur des puits non traités sont comparés à 3 réplicats effectués sur des puits traités par le PVP.

#### Résultats

Pour les 10ng de protéines analysées après leur digestion dans la plaque micro-puits (PDMS 10 $\mu$ l), 379 protéines, 819 peptides et 920 PSM ont été identifiés dans les puits de la plaque micro-puits non traités au PVP. Dans les puits traités au PVP, 665 protéines, 2093 peptides et 2507 PSM ont été identifiés pour les 10ng de protéines analysées. Pour ces échantillons de levure, une différence significative a été observée entre les deux conditions après le test statistique effectué.

Le traitement de surface au PVP double la récupération des peptides et PSM et les identifications des protéines pour les quantités d'échantillon analysées sur la plaque micro-puits en PDMS (ici 10ng de protéines).

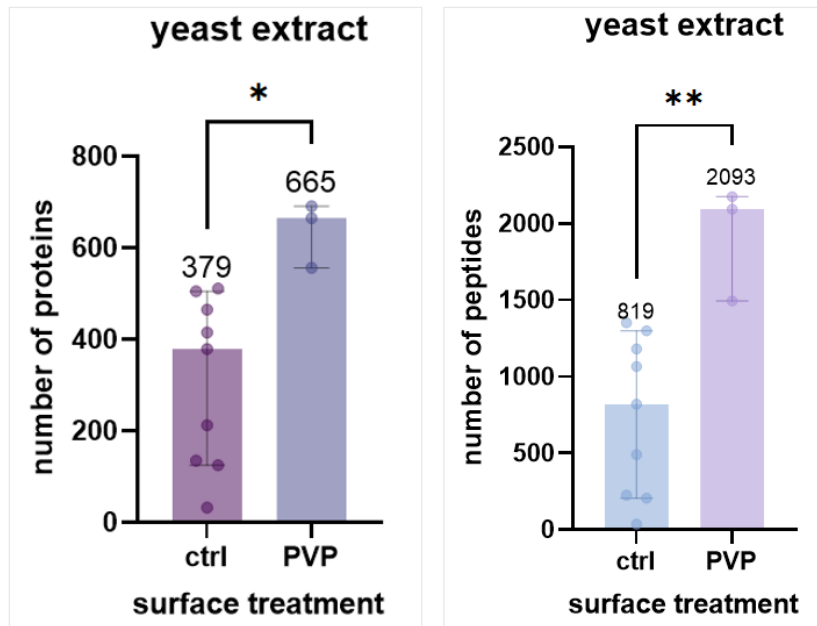


Figure 115 : Nombre de protéines et peptides identifiés à partir de protéines extraites de levure digérées sur la plaque micro-puits en PDMS. Les barres d'intervalles de confiance à 95% sont annotées des médianes des réplicats. Les \* indiquent une p- value <0.05 obtenue par un t test non apparié (avec \*<0.05, \*\*<0.005, \*\*\*<0.0005).

#### 6.D.4. Le PVP est-il compatible avec de plus petites quantités de matériel ?

L'effet du traitement de surface au PVP a été testé avec une quantité d'échantillon de départ moindre sur les différents supports (une plaque standard (PP 120µl) et une plaque micro-puits (PDMS 10µl)). De plus, les différentes étapes du protocole optimisées dans la première partie du manuscrit (ajout de DMSO, ratio enzyme/substrat optimisé) sont introduites afin d'améliorer le rendement d'identifications pour de très petites quantités d'échantillons.

##### Méthode

Des aliquots d'1µl de 10ng extraits protéiques de cellules humaines et de levure *Saccharomyces cerevisiae* ont été déposés avec 1µl de DMSO (5% final) sur une plaque standard (PP 120µl) et sur une plaque micro-puits (PDMS 10µl). Ces échantillons ont été digérés par l'ajout de 1µl de trypsine-Lys C (ratio 2:1 (E/S)). L'équivalent de 2ng a été analysé sur le Tims TOF Pro. La médiane des 12 (et 24) réplicats (pour la condition PVP) effectués est annotée au-dessus des barres d'intervalles de confiance à 95% de chaque condition sur la Figure 116.

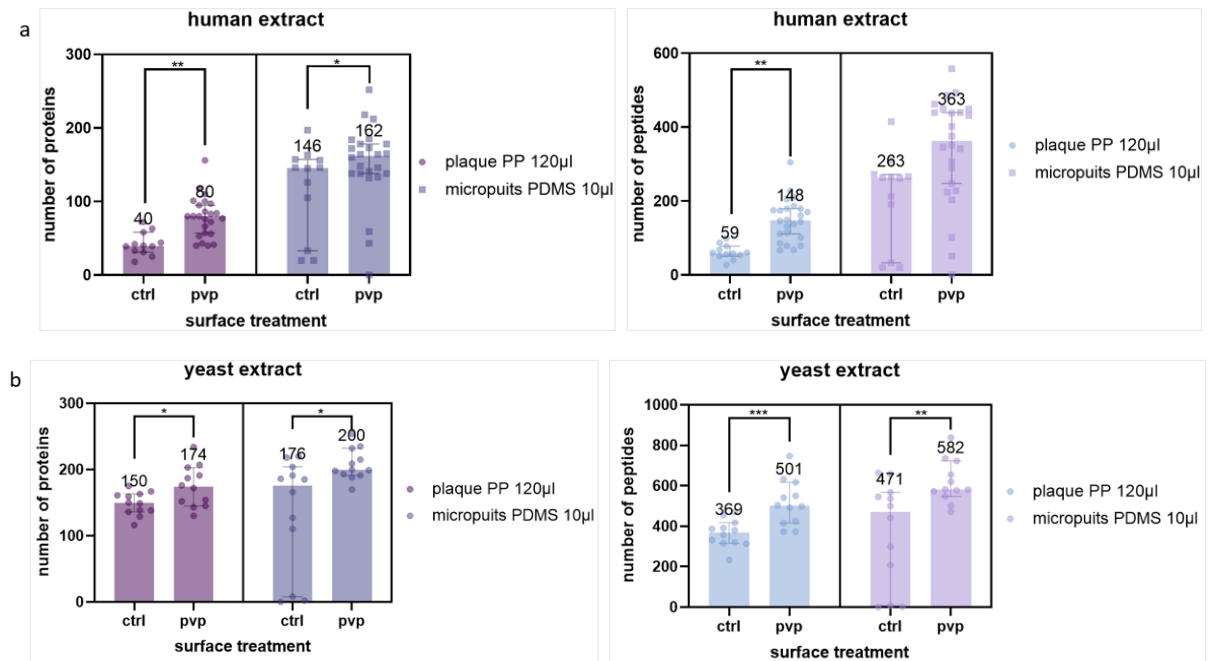


Figure 116 : Nombre de protéines et peptides identifiés après digestion de protéines extraites de cellules humaines ou de levure déposées sur une plaque standard 384 puits ou une plaque micro-puits traitées ou non au PVP. Les médianes des réplicats sont indiquées au-dessus des barres d'intervalle de confiance à 95%. Les \* indiquent une p-value significative après t-test non apparié (avec \*<0.05, \*\*<0.005, \*\*\*<0.0005).

##### Résultats

L'analyse de l'effet du traitement PVP dans chaque support indique un nombre significativement supérieur de protéines, peptides et PSM dans la condition traitée par rapport à la condition contrôle pour les extraits de protéines humaines et de levure d'après le test statistique effectué à l'exception du nombre de peptides dans la plaque micro-puits qui est supérieur dans la condition traitée mais pas significativement différent.



## ETUDES ET RESULTATS : Importance du support

Le traitement de surface apporte une amélioration d'identification pour des quantités d'échantillons équivalentes à 2 ng injectés quel que soit le support utilisé. Le traitement au PVP permet également de réduire les intervalles de confiance dans la plaque micro-puits par rapport à la plaque micro-puits non traitée.

### *Bilan*

L'utilisation combinée d'une plaque micro-puits (PDMS 10  $\mu$ l) et du traitement au PVP permet une meilleure récupération des peptides et donc une meilleure identification protéique pour 2ng de protéines analysées.

## Chapitre 7 La lyse des cellules

<b>7.A Cas de la lyse cellulaire par traitement thermique .....</b>	<b>136</b>
7.A.1. Exemple des cellules Jurkat .....	136
7.A.2. Exemple des cellules THP-1 .....	138
<b>7.B Cas de la lyse cellulaire par sonication.....</b>	<b>140</b>
<b>7.C La méthode de fabrication influence l'état de surface du support .....</b>	<b>143</b>

### 7.A Cas de la lyse cellulaire par traitement thermique

L'efficacité du traitement PVP pour les étapes de protéolyse des protéines et de transfert des peptides protéolytiques a été validée. Afin de compléter l'étude, nous avons souhaité remonter aux étapes précédentes du protocole, notamment à l'étape menant de la cellule intacte à l'extrait protéique. Pour le traitement des cellules, une étape de lyse est nécessaire pour libérer les protéines et les traiter. Une étude avec des cellules Jurkat *wt* et des cellules THP-1 a été effectuée.

#### 7.A.1. Exemple des cellules Jurkat

##### Méthode

Des cellules Jurkat au nombre de 200, 50 et 10, ont été déposées dans une plaque standard (PP 120µl) ou une plaque micro-puits en PDMS (PDMS 10µl) par un trieur de cellules. Ces plaques ont été traitées au PVP en amont de l'expérience. Les cellules ont été lysées par lyse thermique (cycle de froid/température ambiante) en présence d'1µl de DMSO. Les protéines ont été digérées par ajout d'1µl de trypsine-Lys C (ratio 2:1 (E/S)). Puis, l'équivalent de 40, 10 et 2 cellules respectivement ont été analysés sur un Tims TOF Pro. La médiane des 12 réplicats de préparation est annotée au-dessus des barres d'intervalles de confiance à 95% de chaque condition sur la Figure 117 et Tableau 10.

##### Résultats

	40 cellules (2ng)			10 cellules (0.5ng)			2 cellules (0.1ng)		
	Protéines	Peptides	PSM	Protéines	Peptides	PSM	Protéines	Peptides	PSM
PP 120µL	324	832	901	99	187	200	13	17	17
PDMS 10µL	241	532	567	79	144	149	21	23	23

Tableau 10 : Comparaison de la lyse thermique de cellule Jurkat sur plaque PP 120µL commerciale et plaque PDMS 10µL *homemade* (12 réplicats/condition).

Une meilleure identification sur la plaque standard que sur la plaque micro-puits est observée pour 10 et 40 cellules Jurkat analysées lorsque la lyse est effectuée par traitement thermique (Tableau 10).

La tendance semble s'inverser pour 2 cellules Jurkat analysées pour lesquelles l'identification protéique devient meilleure dans la plaque micro-puits, mais sans validation de la significativité par test statistique. Il faut noter le manque de reproductibilité avec une très forte variation des valeurs (Figure 117) selon le lot de cellules considéré parmi les 12 réplicats analysés pour chaque condition. Cette très grande hétérogénéité des résultats observée pour ces réplicats biologiques ne permet pas de tirer de conclusion simple sur l'efficacité du traitement thermique.

ETUDES ET RESULTATS : La lyse des cellules

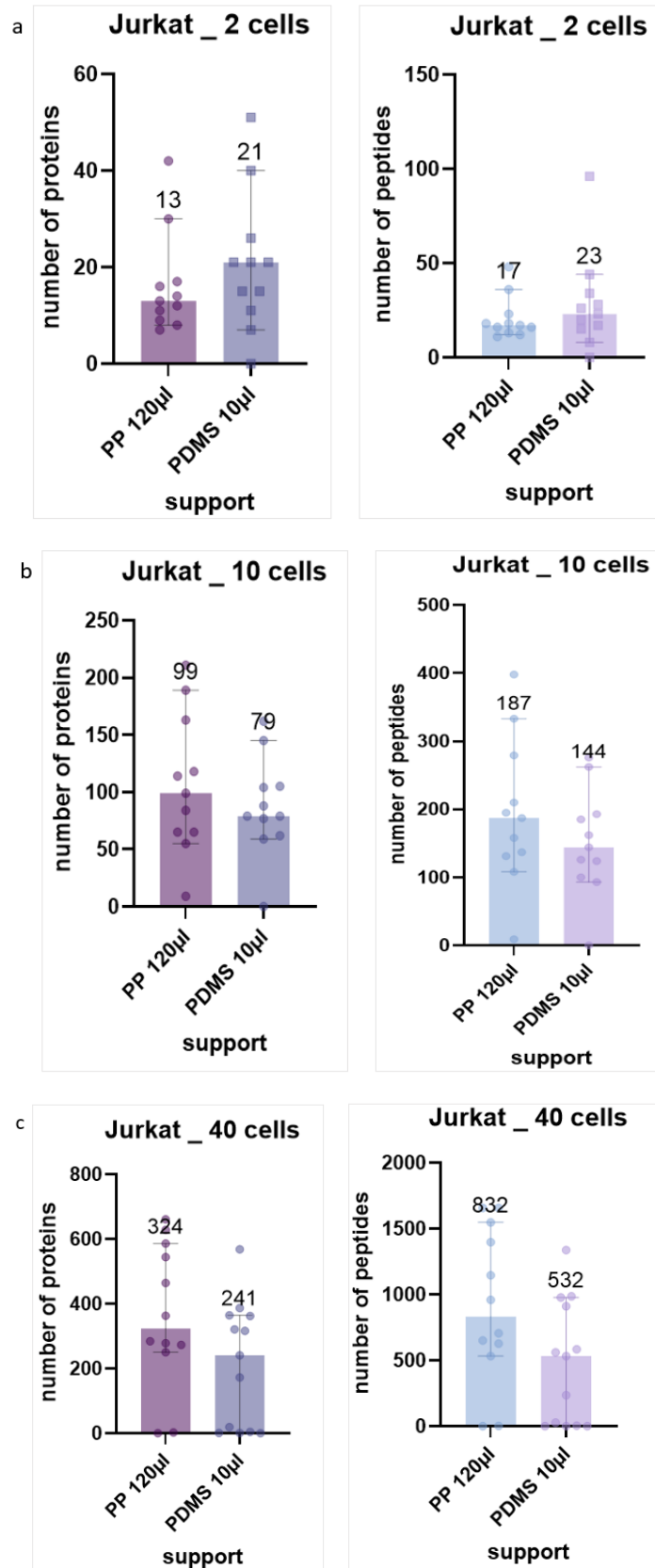


Figure 117 : Nombre de protéines et peptides après analyse de 2, 10 et 40 cellules traitées sur une plaque standard ou une plaque micro-puits en PDMS. Les médianes des réplicats sont indiquées au-dessus de chaque barre d'intervalles de confiance à 95%. Les \* indiquent une p-valeur significative après t-test non apparié (avec \*<0.05, \*\*<0.005, \*\*\*<0.0005).

### 7.A.2. Exemple des cellules THP-1

#### Méthode

La même expérience a été réalisée avec des cellules THP-1. La médiane de 18 réplicats est représentée dans chaque condition (Figure 118 et Tableau 11).

#### Résultats

	40 cellules (4ng)			10 cellules (1ng)			2 cellules (0.2ng)		
	Protéines	Peptides	PSM	Protéines	Peptides	PSM	Protéines	Peptides	PSM
PP 120µL	686	2090	2311	205	414	448	59	82	85
PDMS 10µL	509	1230	1330	121	185	197	54	63	64

Tableau 11 : Comparaison de la lyse thermique de cellule THP1 sur plaque PP 120µL commerciale et plaque PDMS 10µL *homemade* (18 réplicats/condition)

Pour 40 cellules analysées, la différence en nombre d'identifications est significative selon le test statistique effectué avec une identification supérieure dans la plaque standard. Cette même différence en faveur de la plaque standard est observée pour 10 cellules THP-1 analysées. Enfin, pour 2 cellules THP-1 analysées, aucune différence significative selon le test statistique effectué n'est observée.

#### Bilan

Il est à noter qu'un nombre plus important de protéines est obtenu pour les cellules THP-1 (monocytes) ce qui est attendu car leur diamètre est le double de celui des cellules Jurkat (lymphocytes T). Les THP-1 contiennent donc plus de matériel protéique que les cellules Jurkat. Les tendances observées pour les THP1 et les Jurkat sont donc très cohérentes.

Deux étapes peuvent influencer la qualité des résultats : l'étape de distribution des cellules et l'étape de lyse thermique. L'étape de distribution a été vérifiée sur la plaque micro-puits en PDMS 10µl et le trieur de cellules était calibré avant chaque distribution pour chaque support. Ceci suggère donc que la lyse thermique est donc moins efficace dans nos dispositifs en PDMS que dans les dispositifs en PP.

Le PDMS est un polymère avec une grande stabilité thermique, à forte capacité d'isolant<sup>215</sup>. Des études récentes suggèrent même l'utilisation du PDMS comme matrice antigél<sup>215,216,217</sup>.

La lyse thermique (cycle de congélation puis décongélation à température ambiante) n'est peut-être pas la méthode de lyse la plus adaptée en combinaison avec un support en PDMS, surtout si la taille des puits est minimisée, ce qui implique que l'épaisseur du PDMS (plus de 10mm) est très importante, surtout au regard des plaques commerciales en PP qui sont constituées de couches relativement fines de polypropylène (environ 1mm). L'étape de lyse cellulaire, comme discuté dans l'introduction, peut être thermique ou mécanique. Dans les expériences précédentes, une lyse thermique a été utilisée. Nous avons donc souhaité évaluer un autre mode de lyse, comme la lyse par sonication, qui a été réalisée sur des cellules THP-1.

ETUDES ET RESULTATS : La lyse des cellules

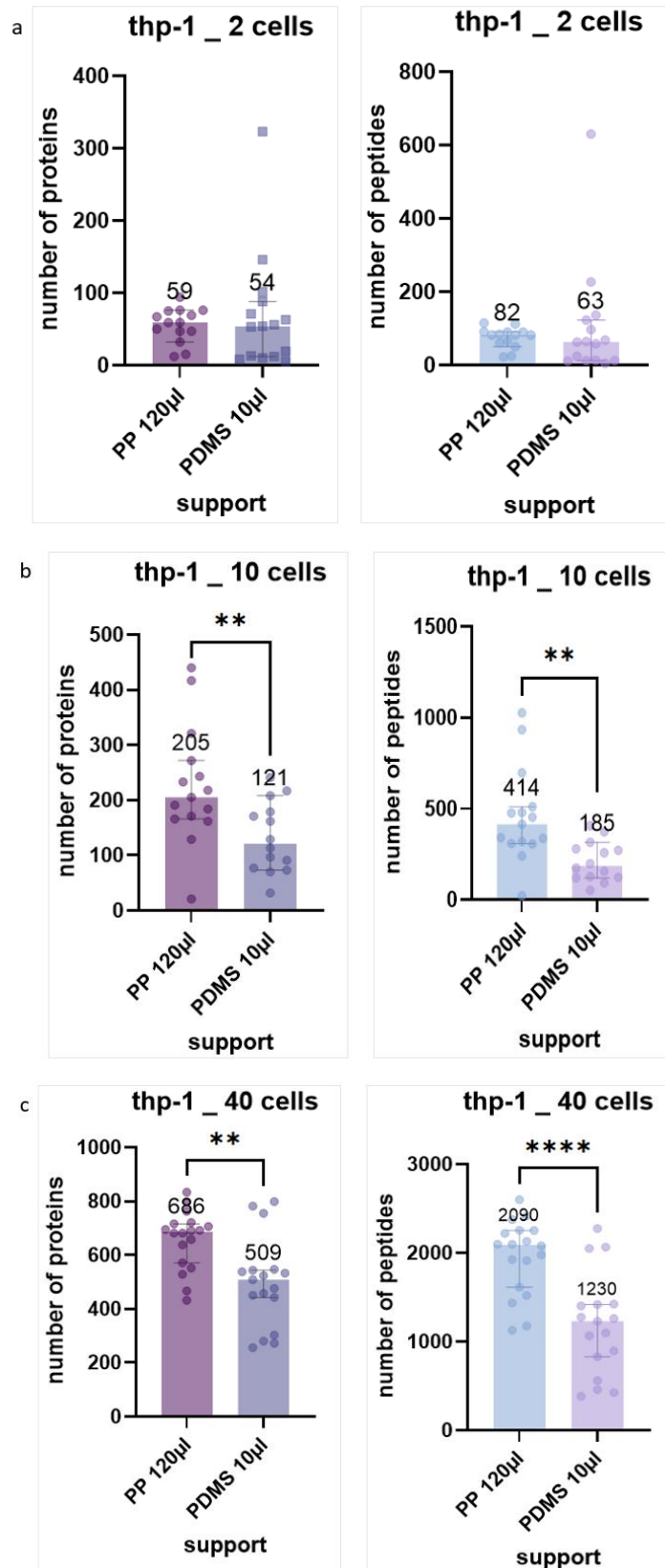


Figure 118 : Nombre de protéines et peptides après analyse de 2, 10 et 40 cellules traitées sur une plaque standard ou une plaque micro-puits en PDMS. Les médianes des réplicats sont indiquées au-dessus de chaque barre d'intervalles de confiance à 95%. Les \* indiquent une p-valeur significative après t-test non apparié (avec \*<0.05, \*\*<0.005, \*\*\*<0.0005).

### 7.B Cas de la lyse cellulaire par sonication

Afin de prendre en compte l'effet de masse du matériel, la nature du matériau et le volume des micro-puits, un nouveau design de plaque multi-puits, à l'image des plaques PDMS 10 $\mu$ l mais en polypropylène (PP). La fabrication de plaques micro-puits en PP a été effectuée par impression 3D à partir de filaments de PP (voir « Miniaturisation des supports »). De plus, nous avons utilisé des plaques commerciales ayant des puits plus restreints de 58 $\mu$ l par puits.

#### Méthode

Trois supports ont été comparés dans cette expérience :

- Une plaque commerciale standard (ici 58 $\mu$ l/puits au lieu de 120 $\mu$ l/puits utilisées dans les autres expériences) = PP 58 $\mu$ l
- Une plaque micro-puits en PDMS (10 $\mu$ l/puits) = PDMS 10 $\mu$ l
- Une plaque micro-puits en PP (10 $\mu$ l/puits) = PP 10 $\mu$ l

Les supports ont été traités au PVP en amont de l'expérience. Des cellules THP-1 au nombre de 200, 50 et 10 cellules ont été déposées au fond des puits des 3 supports par un trieur de cellules en présence d'1 $\mu$ l de DMSO. Une lyse par sonication (10 min dans un bain à ultrasons) a été effectuée. Les protéines ont été digérées par ajout d'1 $\mu$ l de trypsine-Lys C (ratio 2:1 (E/S)). Puis, l'équivalent de 40, 10 et 2 cellules respectivement ont été analysés sur un Tims TOF Pro. La médiane des 6 réplicats est annotée au-dessus des barres d'intervalles de confiance à 95% de chaque condition sur la Figure 119 et le Tableau 12.

#### Résultats

	40 cellules (4ng)			10 cellules (1ng)			2 cellules (0.2ng)		
	Protéines	Peptides	PSM	Protéines	Peptides	PSM	Protéines	Peptides	PSM
PP 58 $\mu$ L	645	1725	1823	180	373	410	10	20	22
PDMS 10 $\mu$ L	764*	2048	2194	90*	152	158	5	6	6
PP 10 $\mu$ L	645	1642	1757	180	354	389	22	38	41

Tableau 12 : Comparaison de la lyse thermique de cellule THP1 sur plaque PP 120 $\mu$ L commerciale et plaque PDMS 10 $\mu$ L *homemade* (18 réplicats/condition). (\*) différence significative.

#### Bilan

Ici encore, la variabilité inter-réplicats observée est très importante dans les plaques micro-puits. La variabilité de cellule à cellule est d'autant plus influente que le nombre de cellules analysées simultanément est faible. Le nombre de réplicats pour les conditions à 2 cellules doit être augmenté pour pouvoir faire une étude statistique.

Les deux supports en PP (plaque standard 58 $\mu$ l/puits et plaque micro-puits 10 $\mu$ l/puits) ne présentent aucune différence significative de performances pour les différentes quantités de cellules analysées. La nature du support PP vs. PDMS a une très forte influence même après traitement PVP et le support PP semble particulièrement intéressant pour l'étape de lyse cellulaire pour les faibles quantités.

## ETUDES ET RESULTATS : La lyse des cellules

Dans le cas du support PDMS 10 $\mu$ L, le nombre d'identifications a été amélioré pour 40 cellules analysées (754 protéines identifiées avec la lyse par sonication au lieu de 509 protéines avec la lyse thermique). En revanche, pour 10 et 2 cellules analysées, le nombre d'identifications diminue : 90 protéines pour 10 cellules et 5 protéines pour 2 cellules, au lieu de 121 et 54 avec la lyse thermique respectivement (Figure 119 et Figure 118).

Dans le cas des supports en PP, les identifications obtenues pour 40 et 10 cellules sont similaires à celles observées avec le PP 120 $\mu$ l avec une lyse thermique (645 vs. 685 et 180 vs. 220 respectivement). En revanche ce nombre chute pour 2 cellules (10/25 selon la taille des puits 58  $\mu$ L/10  $\mu$ L respectivement vs. 59 dans le cas de la lyse thermique). Le nombre de peptides diminue aussi avec la plaque PP 58 $\mu$ L vs la plaque PP 10 $\mu$ L. La nature de la lyse affecte les identifications différemment pour 40 cellules analysées ou pour des plus petites quantités de cellules. L'état de surface de la plaque est très différent entre les plaques commerciales PP 120 $\mu$ L et PP 58 $\mu$ L produites par moulage et la plaque PP 10 $\mu$ L produite par impression 3D et qui présente donc des inégalités de surface. Ceci peut expliquer la grande similarité des performances entre la plaque PP 58 $\mu$ L et la plaque PP 10 $\mu$ L proposant une plus petite surface mais avec plus d'aspérités.

L'intérêt de la sonication est plus important pour le PDMS très isolant que pour le PP, ce qui se constate sur les essais à 40 cellules. Il s'agit donc d'une bonne alternative pour le PDMS mais peut-être moins pour le PP. De plus le traitement PVP est peut-être moins résistant aux ultrasons qu'aux variations thermiques.

Pour conclure sur cette étape de lyse cellulaire, le traitement thermique semble moins favorable pour les supports en PDMS plus isolants, les dimensions des puits ne peuvent être comparées qu'à état de surface comparables, la variabilité entre cellules demande un plus grand nombre de réplicats pour les faibles nombres de cellules et enfin il serait intéressant de vérifier la stabilité du traitement PVP pendant une sonication.

Afin de regarder l'influence de l'état de surface des plaques nous avons donc décidé de nous affranchir de l'étape de lyse cellulaire pour réaliser une étude comparative des résultats juste après cette étape sur le même modèle cellulaire.



ETUDES ET RESULTATS : La lyse des cellules

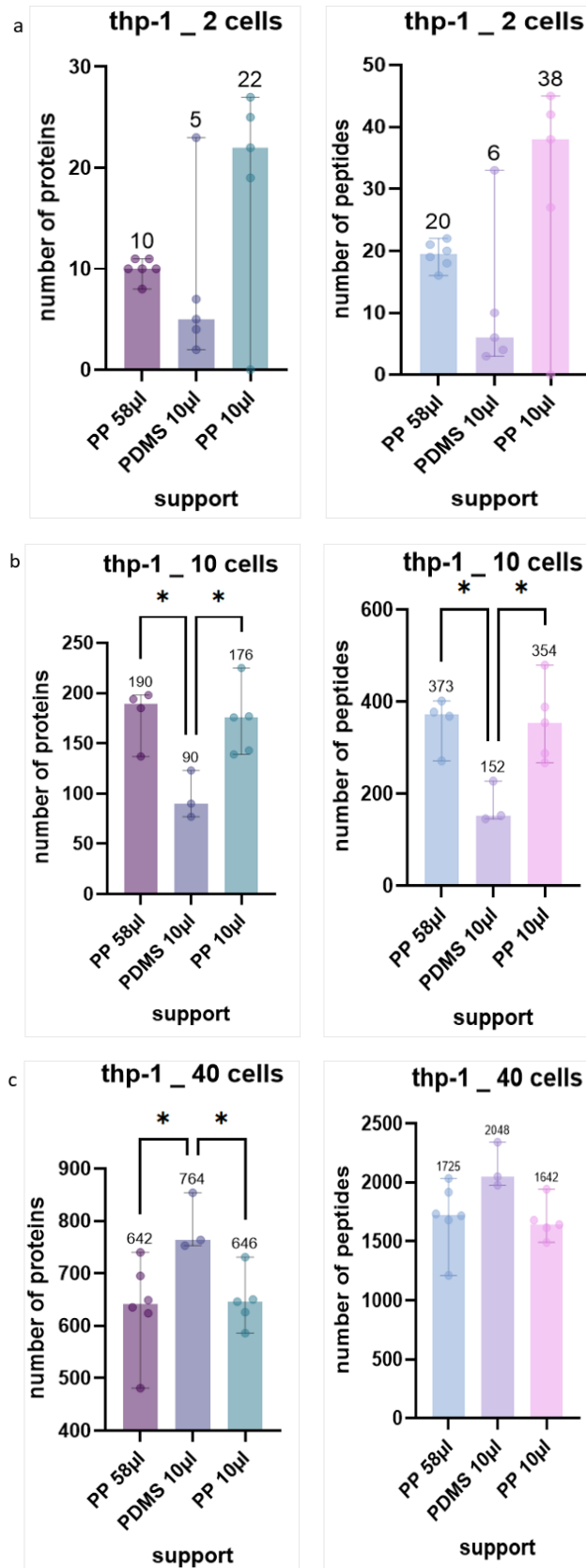


Figure 119 : Nombre de protéines et peptides identifiés pour 2, 10 et 40 cellules analysées après une lyse mécanique par sonication. Le nombre d'identification médian est annoté au-dessus de chaque barre d'intervalles de confiance à 95%. Les \* indiquent une p-value significative après un Anova à un facteur (avec \* $<0.05$ , \*\* $<0.005$ , \*\*\* $<0.0005$ ).

### 7.C La méthode de fabrication influence l'état de surface du support

Pour mieux comprendre les résultats précédents et approfondir l'étude sur le rôle de l'état de surface des supports, nous nous sommes concentrés sur la comparaison des deux supports en polypropylène avec des échantillons prétraités en *bulk*. Une plaque commerciale en PP avec les plus petits volumes de puits (58µl) a été comparée à une plaque micro-puits en PP imprimée en laboratoire.

Des digestats de cellules THP-1 traitées par un protocole de protéomique en bulk ont été déposés sur deux supports en PP traités au PVP. Ces échantillons ont été soumis à des conditions mimant la digestion afin de vérifier les effets du chauffage sur l'identification des protéines de digestats sur deux supports en PP.

#### *Méthode*

Deux millions de cellules THP-1 ont été traitées en bulk en suivant le protocole du kit de préparation de protéines iST (Preomics). La concentration du digestat obtenu a été mesurée par un kit de mesure de la concentration peptidique par fluorescence (Pierce). Puis, ce digestat a été dilué dans AF 0.2% pour correspondre à 10.000, 1000 et 100 cellules. Un aliquot d'1µl de ces dilutions a été déposé sur 2 supports, une plaque commerciale standard 384-puits avec 58µl/puits (PP 58µl) et une plaque micro-puits en PP 10µl (PP 10µl). Ces supports ont été traités au PVP en amont de l'expérience. Les échantillons ont été chauffés à 37°C pendant 3 heures afin de mimer les effets de la digestion. L'équivalent de 2000, 200 et 20 cellules respectivement est analysé sur l'Eclipse Tribid. Le nombre médian du nombre de protéines identifiées pour les 3 réplicats de préparation est annoté au-dessus des barres d'intervalles de confiance à 95% de chaque condition sur la figure (Figure 120).

#### *Résultats*

Pour l'équivalent de 2000 et 200 cellules analysé, une différence significative est observée entre la plaque commerciale (58µl/puits) et la plaque micro-puits PP 10µl en nombre de protéines identifiées. Pour 20 cellules, aucune différence significative n'est observée entre les 2 supports.

La plaque commerciale permet une meilleure identification pour de grandes quantités de matériel analysé malgré la miniaturisation des puits en PP imprimés en 3D. Ici, malgré une surface de contact plus faible, les supports en PP imprimés ne donnent pas les résultats attendus. Une hétérogénéité des surfaces est mise en évidence.

Les résultats indiquant une meilleure performance des plaques commerciales par rapport aux plaques micro-puits sont également retrouvés sur des échantillons traités en *bulk*, ce qui suggère qu'en plus de l'étape de lyse cellulaire, le support et sa fabrication jouent aussi un rôle dans ces résultats inattendus.

La différence entre les deux supports réside d'une part dans le volume des puits mais il a été démontré auparavant que la miniaturisation augmentait l'identification protéique et d'autre part dans la technique de fabrication. En effet, les plaques commerciales en PP sont fabriquées par moulage de polypropylène pur et leur fabrication industrielle est contrôlée pour être reproductible et obtenir une surface lisse. Or, les plaques imprimées en laboratoire avec une résolution moyenne limitée par le temps de fabrication et la taille de la buse utilisée (100µm) possèdent une surface rugueuse (Figure 121).

Une hétérogénéité des surfaces mise en évidence dans cette expérience démontre que le gain de la miniaturisation du volume est perdu à cause de l'état de surface non optimal des plaques en PP imprimées.

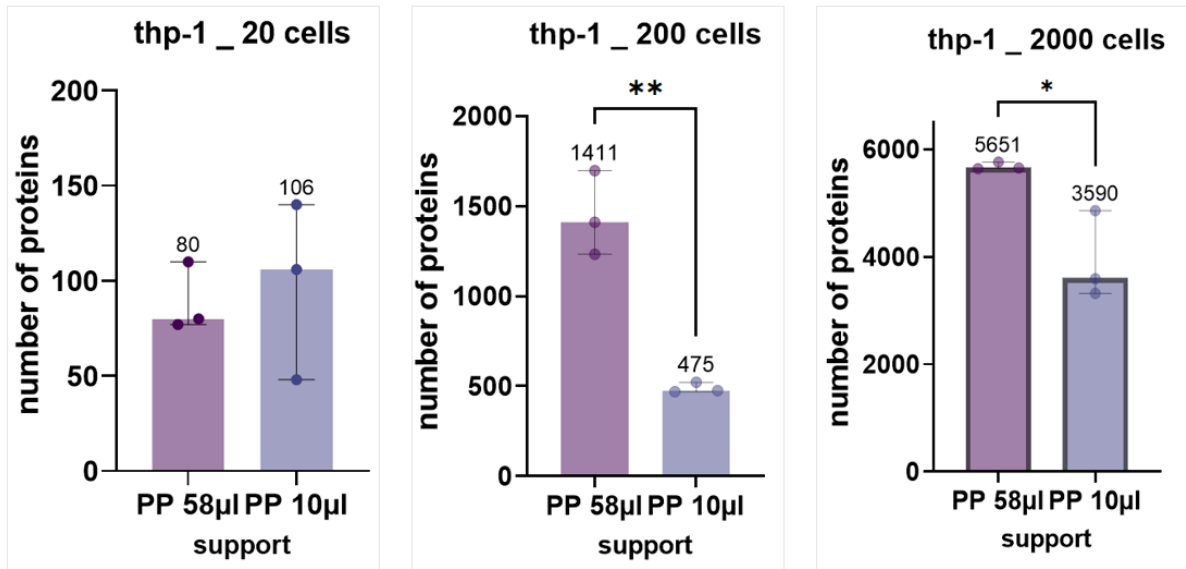


Figure 120 : Nombre de protéines identifiées après dépôt de différentes concentrations d'un digestat de cellules Thp-1 sur une plaque standard 384 puits avec 58µl/puits (PP 58µl) et une plaque micro-puits PP 10µl (PP10µl). Les médianes des répliquats de préparation sont indiquées au-dessus de chaque barre d'intervalle de confiance à 95%. Les \* indiquent une p-value significative après t-test (avec \*<0.05, \*\*<0.0005, \*\*\*< 0.0005).



Figure 121 : Photographie de l'observation sous microscope optique d'un puits de la plaque micro-puits en PP imprimée en 3D sur lequel on peut observer une surface non lisse avec des cercles concentriques montrant les effets de l'impression 3D donnant la surface rugueuse.

## IV. CONCLUSION GENERALE

<b>Chapitre 8 Bilan des résultats et discussion .....</b>	<b>146</b>
<b>8.A Des protéines aux peptides : une série d'étapes remaniées .....</b>	<b>147</b>
<b>8.B Conditions de séparation chromatographique.....</b>	<b>148</b>
<b>8.C Les surfaces de contact : un vrai piège à peptides ? .....</b>	<b>150</b>
8.C.1. Un peu de géométrie.....	150
8.C.2. L'impact de la nature du matériau de surface .....	152
<b>8.D La manipulation des cellules : une étape très en amont à fort impact .....</b>	<b>153</b>
8.D.1. Sélection et comptage .....	153
8.D.2. Lyse cellulaire .....	153
<b>8.E Tentatives d'automatisation .....</b>	<b>154</b>
<b>Chapitre 9 Conclusion et perspectives.....</b>	<b>156</b>

## Chapitre 8 Bilan des résultats et discussion

Ce projet s'inscrit dans les études de recherche de sensibilité pour la protéomique. Au cours de son évolution constante depuis les années 90, la protéomique *bottom-up* a vu se développer de nombreuses solutions mettant en œuvre des protocoles multiples pour répondre à ce besoin d'analyse ultra-sensible résultant de notre incapacité à amplifier le signal comme en transcriptomique ou en génomique. De nombreuses étapes peuvent présenter des écueils rédhibitoires pour détecter les espèces minoritaires, et ils se cachent souvent dans les détails. Ces aspects sont rarement considérés en eux-mêmes comme un sujet d'étude, pourtant ils peuvent conditionner la réussite d'une analyse. Depuis les années 2000, un effort considérable a été fait d'un point de vue technologique en instrumentation en spectrométrie de masse, avec une forte implication des constructeurs et notamment en mode d'analyse de type nano LC-MS/MS. En revanche, s'il est indispensable de travailler avec des analyseurs et des modes de détection performants, il est tout aussi, voir même plus important de savoir manipuler et traiter les échantillons avec des rendements optimaux. Sans pour autant être exhaustive, voici une liste des points d'attention que nous avons identifiés dans ce dernier aspect : la manipulation sans dégradation ou perte des objets d'étude, par exemple les cellules en petits nombre, l'extraction spécifique et non biaisée des protéines, les chutes de rendement au cours de la protéolyse, les pertes de polypeptide par adsorption selon les conditions d'échantillonnage, la composition des tampons et solvants ou la nature des contenants. La littérature regorge d'études et d'applications prometteuses, mais lorsqu'il faut comparer les divers protocoles, certains aspects sont rarement discutés comme le choix ou la géométrie des matériaux dont les surfaces sont en contact avec les molécules d'intérêt, ou encore l'utilisation de certains réactifs et leur concentration en fonction de la taille des échantillons, ou encore les conditions de chromatographie.

Certains protocoles classiques sont unanimement repris par la communauté scientifique pour le traitement et l'analyse de plusieurs millions de cellules, c'est-à-dire de quelques centaines de  $\mu\text{g}$  de matériel. Les règles changent lorsque la quantité de matériel disponible chute de plusieurs ordres de grandeur, et qu'il s'agit de travailler sur un faible nombre de cellules, sachant que quelques milliers de cellules correspondent alors à quelques centaines de ng de matériel de départ (voir Tableau 1). Si la sensibilité de détection peut être évaluée par dilution, le traitement et la manipulation des échantillons sont tout autres, les volumes changent et avec eux, la cinétique réactionnelle, l'importance relative des phénomènes d'adsorption non spécifique, la compatibilité des solvants, les aspects fluidiques et l'impact des volumes morts, par exemple. Dans le cadre de ce travail, nous avons cherché à reprendre certaines solutions indépendantes présentées comme des protocoles de routine afin de les étudier mais aussi de voir si la combinaison de ces protocoles qui ne sont pas souvent remis en question permettait d'améliorer les performances de l'analyse, indépendamment de l'aspect instrumental en spectrométrie de masse. Dans ce cadre, quelques aspects ont été plus particulièrement considérés : les conditions de traitement de l'échantillon, l'impact de la configuration nano LC et l'adsorption des peptides sur les surfaces. D'autres aspects ont été abordés concernant la manipulation des cellules et permettent d'identifier un certain nombre de points d'attention dans la perspective d'études futures. Enfin un enjeu particulier a été constamment pris en compte : si chaque étape peut être optimisée séparément, les conditions optimales ne sont pas nécessairement celles qui produisent les meilleurs

résultats/rendements indépendamment mais celles qui produisent des résultats satisfaisants et restent compatibles en aval comme en amont, avec la chaîne d'étapes nécessaires à l'analyse protéomique complète.

### 8.A Des protéines aux peptides : une série d'étapes remaniées

Nos résultats concernant la protéolyse des protéines montrent que pour un échantillon de 10ng de protéines extraites de cellules humaines ou de levure (soit l'équivalent de quelques centaines de cellules), un protocole simplifié avec une étape de digestion seule fournit de meilleurs résultats en termes de nombre de protéines identifiées et de peptides séquencés analysés qu'un protocole incluant des étapes de réduction et d'alkylation. Bien sûr, l'absence de ces étapes de réduction/alkylation proscrit théoriquement l'accès à une couverture de séquence optimale de nombreuses protéines, et limite l'efficacité de dénaturation. Nous avons émis l'hypothèse que pour de petites quantités d'échantillons, le gain théorique apporté par les étapes de réduction et d'alkylation était entravé par la perte due à l'adsorption des peptides et protéines à la surface à cause de la durée supérieure de prolongement du contact entre l'échantillon et le support lors de l'ajout des étapes de réduction et d'alkylation. La même tendance a été démontrée sur des échantillons obtenus à partir de 100 et 1000 cellules Jurkat après une lyse thermique. En comparant les résultats obtenus avec et sans les étapes particulières de dénaturation, réduction et alkylation, un nombre supérieur d'identifications a été constaté avec les protocoles n'impliquant pas de réduction et d'alkylation. Par ailleurs, le protocole impliquant une dénaturation a démontré une meilleure identification pour 100 cellules Jurkat analysées mais ce protocole ayant été testé en tube, cette étape est évitée lors de l'analyse des échantillons en plaque plus prônes à l'évaporation. Lorsque les volumes d'échantillons sont très faibles, cette étape nécessiterait un contrôle précis de l'environnement en contact avec la solution pour limiter l'évaporation, notamment en travaillant à la pression de vapeur saturante de l'eau à 37°C, pour éviter tout phénomène d'évaporation/condensation.

L'étape de protéolyse enzymatique a aussi été étudiée. D'un point de vue cinétique, des temps d'incubation compris entre 15 minutes et 3 heures ont été comparés et le nombre d'identification protéiques et peptidiques a augmenté avec le temps jusqu'à 3 heures. Lorsque les échantillons analysés sont plus concentrés, les phénomènes d'adsorption sont moins importants d'un point de vue relatif, c'est-à-dire que la perte des molécules par adsorption reste négligeable par rapport à la quantité totale de molécules en présence. Dans le cas d'échantillon très faiblement abondants, nous avons limité cette comparaison à 3 heures afin de limiter l'adsorption due à un temps de contact trop long entre l'échantillon et le support.

Selon la littérature, une augmentation considérable de la concentration en trypsine peut compenser partiellement les effets de la réduction des concentrations de substrat<sup>53</sup>. Cependant, il y a une limite au-delà de laquelle, l'autolyse et l'activité chymo-tryptique résultant de l'excès de trypsine interfèrent avec l'analyse. Pour augmenter la concentration, il est possible d'augmenter la quantité de produit ou bien de diminuer le volume. Dans nos expériences, nous avons décidé de travailler dans un volume d'environ 5µl afin de réduire les surfaces de contact (et limiter le phénomène d'adsorption) et de concentrer nos échantillons. Deux ratios trypsine/substrat ont été comparés, un ratio classique E/S de 1/20, qui est couramment utilisé en protéomique, et un ratio où la tendance est inversée E/S de 2/1 où la

## CONCLUSION GENERALE : Bilan des résultats et discussion

quantité de trypsine est deux fois supérieure à la quantité d'échantillon à digérer. Ce dernier ratio n'est jamais utilisé pour les grandes quantités de protéines, d'une part pour des raisons de coût mais aussi parce que les phénomènes d'autolyse de la trypsine génèreraient des peptides protéolytiques de la trypsine ultra-majoritaires qui risqueraient de saturer notre système nano LC-MS/MS. Mais appliqué à de petites quantités d'échantillon, nous avons observé une amélioration des identifications protéiques et peptidiques pour le ratio 2/1 en comparaison au ratio 1/20. Aucune autolyse induisant une perte d'activité tryptique n'a été détectée. Un seul pic d'autolyse issu de l'incubation trypsine-Lys C a été identifié, et est marqué comme contaminant lors des traitements de données (par le logiciel MaxQuant). Le ratio trypsine/substrat 2/1 a donc été adopté pour la suite du projet.

Le diméthylsulfoxyde (DMSO) est un solvant aprotique et polaire miscible à l'eau, l'acétonitrile ou les alcools avec une très haute solubilité protéique. C'est un agent cryoprotecteur et dénaturant, pur ou à haute concentration. Avec un point d'ébullition élevé (190.5°C) et une pression de vapeur saturante faible (0.59 mbar à 20°C) comparé à l'eau (23.4mbar à 20°C), il n'est pas volatile, sa présence même en petit pourcentage ( $\leq 5\%$ ) permet de limiter l'évaporation et faciliter la lyse des cellules. De plus sa présence permet d'améliorer l'efficacité d'ionisation des peptides en nano ESI. Nous avons donc évalué l'effet de son ajout dans le tampon sur la qualité des analyses en termes de nombre de protéines. Les échantillons ont été traités en présence de 2% DMSO final ou 5% DMSO final. L'ajout de DMSO à l'étape de récupération des peptides n'a eu aucun effet. En revanche, son ajout à l'étape de digestion a permis de limiter l'évaporation. En effet, pour les deux concentrations, une différence significative du volume final a pu être observée en comparaison avec l'échantillon contrôle en solution dans l'eau, en faveur des échantillons dilués dans le DMSO, qui ne s'évaporent que très peu. Bien que les concentrations de DMSO soient proches, la comparaison des médianes du nombre d'identifications (réplicats de préparation) montre que l'analyse protéomique obtenue à partir des échantillons à 5% DMSO final permet d'identifier plus de protéines de façon reproductible. Par la suite, ce taux de 5% DMSO a été choisi pour le reste des études. La seule modification observée au niveau de la chromatographie nano LC en présence de DMSO est un épaulement sur le profil de pression de la colonne en début de pression.

### 8.B Conditions de séparation chromatographique

La protéomique *bottom-up* met en œuvre dans la majeure partie des études une configuration nano LC ESI-MS/MS. La chromatographie en phase liquide utilise par défaut dans le cas des mélanges protéolytiques tryptiques des colonnes à polarité de phase inversée de type silice greffée octadécyle C18 de porosité 100-300 Å, de granulométrie 1.6-2.5µm de longueur 15-60 cm et de diamètre interne 75µm. À perte de charge comparable (et dépendant de la tolérance de l'instrumentation utilisée), les paramètres physiques de la colonne vont imposer un débit de phase mobile entre 200 et 350 nl/min. La nature de la phase et la longueur de la colonne sont des paramètres très importants pour le pouvoir résolutif mais la sensibilité de l'analyse va être en particulier influencée par le diamètre interne de la colonne, car en jouant sur la concentration locale des solutés au sein de la colonne et sur le débit d'élution et donc sur les conditions d'ionisation nano ESI, il est théoriquement possible d'augmenter le rapport signal/bruit et la détectabilité des espèces. D'un point de vue commercial, seules des colonnes de 75µm et 50µm d.i. sont disponibles, et sur un nombre restreint de phases stationnaires. Deux colonnes de diamètre interne différent (75µm et

## CONCLUSION GENERALE : Bilan des résultats et discussion

50 $\mu$ m) ont été comparées sur un Q Exactive HF. La configuration en directe injection en combinaison avec une colonne de petit diamètre interne a donné les meilleures identifications protéiques sur une gamme de quantités croissantes de digestats peptidiques commerciaux. Pour étudier des colonnes de diamètre plus faible, des colonnes fabriquées en laboratoire ont été testées et comparées à une colonne de diamètre interne standard (75 $\mu$ m) sur différentes configurations. Les identifications obtenues avec une colonne fabriquée au laboratoire *homemade* de diamètre interne 75 $\mu$ m sont comparables à celles obtenues avec une colonne commerciale du même diamètre interne. Les performances augmentent lorsque le diamètre interne diminue, avec un plus grand nombre d'identifications de protéines, atteignant environ 2000 protéines pour 10ng de digestat de HeLa. Un stage au laboratoire a permis de fabriquer des colonnes de 20 $\mu$ m d.i. Les données n'ont pas été détaillées dans ce manuscrit. En effet certains problèmes se sont révélés au cours de la mise en œuvre de ces colonnes ce qui a provoqué un manque de reproductibilité. Ainsi la perte de charge sur des colonnes de 20 $\mu$ m d.i. augmente comme l'inverse du carré du rayon, ce qui implique, pour rester dans les spécifications de nos pompes chromatographiques, d'augmenter la température du four à colonne à 65°C, sachant que l'acétonitrile a un point d'ébullition à 82°C. De plus, pour conserver une vélocité linéaire constante, le débit doit être fortement diminué et passe en deçà des spécifications de la majeure partie des pompes nano HPLC (21.3 nl/min vs. une limite de fonctionnement instrumental en élution en mode gradient de 50 nl/min). Une solution serait de réaliser un *split* sur notre pompe nano. Dans ce cas, il faudra aussi travailler afin que le débit reste suffisant pour l'établissement d'un spray stable, ce qui a pu être obtenu pour un débit de minimum de 50nL/min. Enfin la diminution du volume équivalent colonne rend le volume de la connectique surdimensionnée, ce qui va amplifier l'influence des volumes morts sur la séparation et notamment les volumes en sortie de colonne, ce qui fait chuter la résolution de la séparation mais aussi sa sélectivité. Malgré tout, les quelques essais dans ces conditions défavorables sont cohérents avec l'efficacité de la miniaturisation sur la recherche de sensibilité.

Une particularité de l'analyse de mélanges biologiques complexes en protéomique, malgré le travail réalisé pour limiter les volumes de travail, est le décalage entre le volume de l'échantillon à injecter et le volume du système de séparation chromatographique. Ainsi le volume équivalent colonne d'une colonne de 75 $\mu$ m d.i. et 50 cm de long sera d'environ 1.6 $\mu$ L, et le volume total de l'échantillon à analyser peut typiquement être de l'ordre de quelques  $\mu$ L. Il ne s'agit plus de condition d'injection ponctuelle. De plus pour injecter en injection directe sur la colonne 3 $\mu$ L d'échantillon à un débit de 200nL/min, il est nécessaire de charger l'échantillon pendant 15 minutes. Pour injecter plus rapidement, il faut augmenter le débit d'injection. Et cela peut se faire en ajoutant une cartouche de préconcentration de même nature que la colonne analytique mais de diamètre plus large et pouvant travailler à plus haut débit de 10 à 40 $\mu$ L/min. L'injection en préconcentration permet d'injecter un grand volume d'échantillon dilué, de piéger les espèces sur une cartouche intermédiaire puis de réduire le volume d'injection effectif sur colonne par concentration sur la pré-colonne tout en éliminant les espèces les plus hydrophiles comme les sels. Cependant, expérimentalement nous avons constaté que la configuration en préconcentration avec une colonne de petit diamètre interne ne permet pas d'atteindre les identifications observées en injection directe avec la même colonne. Quelques hypothèses peuvent être émises. Tout d'abord ce principe de préconcentration augmente les surfaces de contact de notre échantillon et des contenants,



## CONCLUSION GENERALE : Bilan des résultats et discussion

et favorise la perte d'espèce par adsorption, sur l'ensemble du chemin fluidique. De plus, le piégeage sur la pré-colonne n'est pas nécessairement efficace pour toutes les espèces en présence. Il peut aussi être quasi irréversible et peut donc créer un biais analytique.

À quelques exceptions près, pour s'affranchir de ces aspects et éviter les variations au niveau chromatographique, la plupart des expériences dans ce projet ont été réalisées sur un montage normalisé en injection directe intégrant un Tims TOF Pro (Bruker) avec des colonnes Aurora (Ion Opticks) de 75µm de diamètre interne x 25cm, en phase C18 de granulométrie 1.6µm, équipées d'un nébuliseur intégré car leur *tip* intégré permettait d'éviter les volumes morts post colonnes et était conçu spécialement pour la source nano ESI fournie par Bruker. Nous (et plusieurs autres équipes) avons pu constater une baisse de la qualité de ces colonnes lors d'un changement de lot de fabrication des colonnes en milieu de thèse. La reproductibilité de nos standards de digestats peptidiques a été affectée par cette baisse de qualité notamment sur les petites quantités. Ce point a été pris en compte : nous ne pouvons directement comparer les résultats ensemble, mais les rapporter à un standard analysé extemporanément et servant de référence entre campagnes d'acquisition.

### 8.C Les surfaces de contact : un vrai piège à peptides ?

Comme nous l'avons évoqué à plusieurs reprises, l'adsorption des peptides sur les surfaces de contact est une cause majeure de la perte d'échantillon tout au long des analyses. Tous les groupes travaillant sur la problématique de la protéomique sur cellule unique sont confrontés à ce problème et le traitent dans la littérature de façon plus ou moins explicite.

#### 8.C.1. Un peu de géométrie

Pour limiter les contacts, l'approche directe est de travailler sur la géométrie des surfaces. Ainsi, de façon théorique, le contact entre une sphère (la goutte d'échantillon) et une surface (le support de l'échantillon) est ponctuel. Cela va sans tenir compte de la mouillabilité des surfaces et des interactions moléculaires qui vont modifier cette interface. Quoi qu'il en soit, c'est la configuration qui doit minimiser la surface de contact. C'est dans cette idée qu'a été proposée la plaque de NanoPOTs<sup>136</sup>. Cette solution nécessite la micro-fabrication de supports en verre dédiés. Nous avons choisi de travailler avec un support plan disponible commercialement et dont les dimensions sont parfaitement normalisées avec les systèmes analytiques des laboratoires : les plaques de cibles MALDI, parfaitement planes et en matériaux inerte compatible avec la spectrométrie de masse. Un essai de support commercial avec des plaques MALDI non traitées (surface inox poli) et traitées (*AnchorChip*) a été comparé à une analyse classique en tube polypropylène *low binding*. En l'absence de DMSO, l'évaporation observée sur les plaques MALDI à 37°C pour des dépôts de faible volume (de l'ordre du µL) ne permet pas de travailler sans contrôle de la pression, de la vapeur saturante de l'eau dans l'environnement immédiat des dépôts tests des peptides de l'hémoglobine ou des digestats de protéines extraites de cellules de levure. En effet, la digestion de la trypsine reste incomplète puisque la solution s'évapore au bout de quelques minutes seulement. Le traitement de surface de l'*AnchorChip*, qui présente des ancrs hydrophiles permettant une concentration de la goutte sur une faible surface entourée d'un environnement hydrophobe est cependant très intéressant. Pour limiter davantage

## CONCLUSION GENERALE : Bilan des résultats et discussion

l'évaporation, la digestion des protéines extraites de cellules de levure a été réalisée à 20°C et comparée à la condition en tubes. Comme attendu, une digestion partielle est observée dans les tubes à 20°C, vu que la trypsine n'est pas à sa température optimale de fonctionnement. Néanmoins, sur la plaque *AnchorChip*, le nombre d'identifications a augmenté et permet d'obtenir des résultats très comparables, malgré l'absence de système de fermeture, confirmant que plus l'évaporation est retardée, plus la trypsine reste en solution et plus le rendement de la réaction est amélioré. N'ayant pas accès à une enceinte thermostatée avec contrôle d'hygrométrie, et le développement d'une enceinte de contrôle dédiée sortant du périmètre du projet, nous n'avons pas pu mettre en œuvre cette solution simple. Nous concluons donc ici que même dans des conditions très défavorables, l'utilisation d'une surface plane et sans rebords minimisant les interfaces échantillon/contenant grâce à un double traitement de surface polie hydrophobe exposant des ancras hydrophiles pour immobiliser des gouttes avec un angle de contact le plus élevé possible est une solution très prometteuse. Cette étude met en évidence l'effet limitant majeur de l'évaporation dans la manipulation des petits volumes pour les analyses miniaturisées.

Afin de travailler dans des enceintes fermées une alternative est d'utiliser des formats en puits avec couvercle, ce qui est la solution la plus démocratisée. À nouveau, peu de solutions commerciales de puits de faible dimension existent et aucune n'est dédiée à nos applications. Deux modèles s'en approchent avec des puits de respectivement 120µL et 58µL, qui sont encore bien trop grands au regard des volumes d'échantillon initial. Afin de travailler sur le prototypage, la recherche d'un support amélioré a continué sous un format de plaque multi-puits afin de pouvoir limiter l'évaporation avec des films. Plusieurs designs ont été créés pour fabriquer un prototype de support en PDMS avec des puits de volume plus petits afin de diminuer la surface de contact.

Le premier prototype était une chambre de réaction avec un canal de sortie permettant de pousser les échantillons par pression douce à la seringue. Il a été créé par moulage du PDMS sur un *wafers* de photolithographie sur lequel le dessin de la plaque avait été imprimé. Un manque de régularité a été rencontré lors de la création des coordonnées lors du poinçonnage. Un deuxième prototype a été créé par moulage de PDMS dans un moule en laiton usiné pour former le reflet négatif de la plaque voulue. Cette fois-ci les coordonnées étaient reproductibles. Cependant, la forme des puits (fond rond) et la taille des diamètres ne correspondaient pas aux dimensions d'une plaque multi-puits standard. Cette plaque micro-puits ne pouvait pas être combinée à l'utilisation d'appareils comme le trieur de cellules. Un troisième prototype a donc été créé par moulage de PDMS dans un nouveau moule en laiton. Les nouvelles dimensions des micro-puits correspondent à la taille d'une plaque 384 (diamètre des puits, disposition sur la plaque). Un fond en V a été ajouté afin de limiter la surface de contact entre l'échantillon et le support en PDMS. Cette plaque micro-puits a été testée en combinaison avec le trieur de cellules pour la récupération de cellules Jurkat et a montré sa compatibilité.

Les performances de cette plaque en nombre d'identification protéique à partir de 50ng d'extraits protéiques humains et de levure sont satisfaisantes puisque nous avons obtenu des identifications significativement différentes entre notre support en PDMS 10µl et une plaque standard en polypropylène à 120µl en faveur du support micro-puits. À nouveau, et en se basant sur une géométrie en puits différente des surfaces planes métalliques précédentes, la

## CONCLUSION GENERALE : Bilan des résultats et discussion

diminution des surfaces de contact permet d'obtenir une amélioration des performances analytiques globales.

### 8.C.2. L'impact de la nature du matériau de surface

Le PDMS est très souple, d'où son utilisation pour la micro-fabrication. En revanche, ses caractéristiques hydrophobes en font un mauvais support pour la manipulation des peptides et des protéines qui interagissent avec leur support. À l'instar des plaques MALDI, il semble donc intéressant d'étudier comment modifier la chimie de surface pour minimiser ces interactions. Un essai préliminaire de traitement de surface au PVP, sur une puce spécialement conçue pour ce test, a été mené. Aucune signature polymérique n'a été observée lors de l'étude des chromatogrammes. Seule la présence du PEG est observée, un contaminant présent dans tous les échantillons de protéomique. Sa présence est due aux plastiques utilisés durant les expérimentations et entre en compétition lors de l'ionisation avec les ions des peptides d'intérêt créant des effets de suppression de signal. Le traitement de surface au PVP montre une tendance à l'amélioration mais sans réelle validation d'un effet significatif d'un point de vue statistique pour un digestat commercial issu de cellules HeLa peu importe les quantités analysées. Considérant que ce premier prototype de puce simple offrait un volume total de seulement 100nl, les surfaces de contact étaient trop faibles pour que l'effet soit directement détectable. Pour maximiser la surface de contact et étudier un traitement de surface, un second prototype a été conçu incluant deux canaux et un réservoir afin de maximiser la surface de contact à 20µl par rapport au prototype précédent. Au laboratoire, plusieurs traitements de surface ont été testés après avoir observé que le passage des peptides dans le PDMS non traité causait la perte de plus de la moitié des peptides par rapport à un échantillon n'ayant jamais traversé le PDMS. De plus, la reproductibilité des échantillons est grandement altérée après leur passage dans le PDMS. Cependant, nous avons observé que le passage dans le PDMS privilégiait l'identification de peptides plus hydrophobes que l'échantillon n'ayant jamais traversé le PDMS, ce qui était contre-intuitif. Huit traitements de surface différents ont été comparés repris de la littérature ou créés en laboratoire. Tous les traitements ont permis une meilleure identification peptidique comparés à l'échantillon ayant traversé le PDMS non traité. Les meilleurs résultats sont observés pour les traitements au PVP, PVA et polydopamine. Vu que les traitements à la polydopamine sont lourds à mettre en œuvre et que le traitement au PVA nécessite de travailler dans des conditions très acides, le choix de l'étude s'est porté sur le traitement de surface au PVP pour rendre la surface du PDMS plus hydrophile. Sa reproductibilité et la facilité de sa mise en place ont aidé à décider d'une étude plus approfondie de ce traitement de surface sur nos échantillons d'intérêt sur nos supports d'intérêt pour vérifier son apport aux différentes réactions du protocole de protéomique.

L'apport du traitement de surface au PVP est étudié pendant l'étape de digestion sur des extraits protéiques commerciaux issus de levures. Nous observons une grande variabilité des données surtout dans la condition où le PDMS n'est pas traité. Une différence significative indique que le traitement au PVP permet l'amélioration de l'identification peptidique et protéique pendant l'étape de digestion.

Cette même étude a été refaite avec un protocole amélioré (en présence de DMSO et avec un ratio trypsine/substrat amélioré). Encore une fois, une grande variabilité est observée

## CONCLUSION GENERALE : Bilan des résultats et discussion

notamment pour les échantillons dans la condition contrôle. Les mêmes résultats que précédemment sont retrouvés pour des extraits protéiques humains et de levure avec une amélioration de l'identification en présence d'une surface traitée au PVP quel que soit le support utilisé.

### **8.D La manipulation des cellules : une étape très en amont à fort impact**

#### **8.D.1. Sélection et comptage**

L'isolement des cellules est une étape cruciale de la préparation des échantillons en cellule unique. Une inspiration par un design commercial a influencé la fabrication de puces en PDMS pour l'isolement des cellules. Plusieurs modèles (en flocons de neige ou linéaires) ont été créés et certains modèles ont montré une bonne diffusion par capillarité des colorants. Ces modèles ont ensuite été testés avec des billes fluorescentes puis des concentrations décroissantes de cellules. Le tri ne s'est pas correctement réalisé. De plus, plusieurs problèmes ont été rencontrés. La taille des puces était très importante et la largeur des canaux utilisés pour la séparation était trop grande (200 à 600 $\mu$ m de large), et n'était donc pas adaptée à une séparation par capillarité car les cellules sédimentaient avant d'arriver aux puits récepteurs. Une étude plus poussée aurait permis la fabrication de puces de taille plus petite et des canaux ayant une largeur de la taille d'une cellule (environ 10 à 20 $\mu$ m) afin de ne permettre le passage que d'une cellule à la fois. Il aurait également été possible de fabriquer des canaux avec des largeurs qui rétrécissent afin de pousser la diffusion dans le sens puits central vers les puits récepteurs. D'autres soucis rencontrés incluent les déchirures dans le PDMS dues au poinçonnage ainsi que les contaminations par les particules de poussière dès que les puces étaient mises au contact de l'air même sous une hotte. Ces expériences auraient pu être continuées en salle propre afin d'éviter ces contaminations.

Le tri et le comptage des cellules est une opération qui a de nombreuses applications et nous avons rapidement décidé de ne pas consacrer de temps à l'étude de cette étape en nous reposant sur l'instrumentation commerciale dont nous disposons et pour laquelle les développements précédents ont été rendus compatibles. Le trieur de cellules utilisé était robuste et reproductible. De plus, cet appareil permettait de séparer un nombre exact de cellules ce qui nous autorisait à travailler sur quelques centaines puis quelques dizaines de cellules afin d'optimiser nos protocoles et notre support.

#### **8.D.2. Lyse cellulaire**

Nous avons ensuite comparé notre support en micro-puits PDMS à une plaque commerciale standard en présence du traitement PVP pour l'étude du traitement de cellules entières. Cela signifie qu'on ajoute les étapes de distribution des cellules dans le support par le trieur de cellules et une étape de lyse thermique.

Étonnamment, les résultats observés pour 40, 10 et 2 cellules Jurkat et THP-1 analysées sont contradictoires avec les résultats précédents. Une trop grande variabilité dans les identifications protéiques et peptidiques obtenues sur les deux supports pour les deux types cellulaires rend difficile une conclusion. Pour les cellules THP-1, nous avons même obtenu un

## CONCLUSION GENERALE : Bilan des résultats et discussion

résultat montrant que la plaque commerciale standard en polypropylène permettait de meilleures identifications que notre support micro-puits en PDMS.

Notre hypothèse s'est basée sur la nature des matériaux. Cette comparaison des supports se repose sur deux paramètres, la nature du matériau (PDMS ou PP) et la taille des puits (120 $\mu$ l ou 10 $\mu$ l). Or, le PDMS a une très grande stabilité thermique. Notre hypothèse est donc que la lyse thermique n'est pas la meilleure option pour lyser les cellules dans notre support micro-puits.

Nous avons donc testé une lyse mécanique par sonication de cellules THP-1. Une comparaison sur 3 supports traités au PVP a été effectuée pour vérifier si la nature et la taille des puits influençait les identifications obtenues en combinaison avec la lyse mécanique. Une plaque commerciale avec les plus petits puits disponibles (58 $\mu$ l par puits) a été comparée à une plaque micro-puits en PDMS 10 $\mu$ l et une plaque en PP de 10 $\mu$ l par puits imprimée en laboratoire. Une variabilité importante entre les réplicats est observée d'autant plus quand la quantité de cellules était faible. La nature du support (PDMS ou PP) semble exercer une influence importante sur le nombre d'identifications protéiques obtenues. Cette lyse par sonication semble plus adaptée pour les supports en PDMS très isolants. Un plus grand nombre de réplicats et une étude de la stabilité du traitement PVP seraient nécessaires pour tirer de meilleures conclusions sur cette lyse mécanique.

Pour les supports en PP, nous avons remarqué un nombre d'identifications similaires entre la plaque commerciale de 58 $\mu$ l par puits et la plaque imprimée en laboratoire de 10 $\mu$ l par puits. Les résultats précédents avaient pourtant indiqué que la miniaturisation du support améliorait l'identification. La différence entre ces deux supports en PP est leur mode de fabrication, par moulage pour la plaque commerciale et par impression 3D pour la plaque micro-puits. Nous les avons comparés sur des échantillons de cellules THP-1 prétraitées en *bulk*. Nous avons observé une meilleure identification pour la plaque commerciale. La différence de fabrication influence l'état de surface du support. L'impression 3D des plaques micro-puits propose une plus petite surface de contact mais avec plus d'aspérités qu'une plaque industrielle moulée.

### 8.E Tentatives d'automatisation

Une solution pour pallier le manque de reproductibilité serait d'automatiser le protocole. Nous avons fait quelques tentatives qui n'ont pas été présentées dans ce projet. Nous avons tenté la reconfiguration d'un Probot (Dionex) pour la distribution de petites quantités de réactifs et l'aspiration de nos échantillons, mais plusieurs soucis ont été rencontrés. Le Probot est un appareil de liaison entre la LC et la source MALDI. En temps normal, il sert à distribuer en parallèle, les peptides séparés en LC et une matrice sur une plaque MALDI pour former des gouttes à analyser par MALDI TOF. Le problème est qu'en analyse MALDI, les gouttes sont séchées avant l'analyse, ce qui signifie qu'il n'y a aucun environnement dans le Probot permettant de maintenir les gouttes, nous avons donc fait face de nouveau à un problème d'évaporation précoce. De plus, l'implémentation d'un capillaire pour l'aspiration de nos échantillons a été un échec causé par l'aspiration d'air par manque d'étanchéité du capillaire au lieu de l'aspiration de l'échantillon d'intérêt.

## CONCLUSION GENERALE : Bilan des résultats et discussion

Une autre solution qui nous a été offerte pendant un court temps était un manipulateur de liquide. L'ECHO 525 (Beckman Coulter) est un manipulateur acoustique. Brièvement, un transducteur concentre des ondes ultrasons dans une fontaine d'eau. Cette fontaine passe sous des puits d'une plaque source et une goutte de 25nl de liquide (réactif à distribuer) est formée et envoyée vers une plaque réceptrice située au-dessus de la première. L'utilisation de cet appareil permettait d'éviter les problèmes d'adsorption liés aux étapes de transfert de l'échantillon et d'éviter les erreurs d'approximation des volumes obtenus par la pipette. De plus, le très petit volume de liquide transféré permettait de réaliser nos réactions enzymatiques à des concentrations optimales pour les petites quantités de cellules analysées. Ainsi, un protocole dédié à l'ECHO 525 a été conçu sur des plaques commerciales 384-puits dont les dimensions étaient adaptées aux instruments utilisés après distribution de cellules par le trieur de cellules. Les réactions de réduction, alkylolation, digestion et même un marquage au TMT ont été incluses dans ce protocole. Les marqueurs TMT étant repris dans l'acétonitrile, un solvant très volatile non géré par l'appareil, l'ajout des marqueurs a été fait à la pipette. Ce protocole méritait d'être reproduit et amélioré pour une étude plus approfondie mais l'inaccessibilité de l'appareil ne nous a pas permis de le faire. Un autre manipulateur de liquide de la même marque l'ECHO 505 qui permet la création de plus petites gouttes et la gestion de liquides volatiles aurait été le bienvenu.

## Chapitre 9 Conclusion et perspectives

En conclusion, ce projet de thèse a permis l'exploration de plusieurs pistes prometteuses pour l'amélioration du protocole de l'analyse du protéome total de quelques cellules. Tout d'abord plusieurs étapes du protocole de protéomique ont été étudiées et améliorées par ajout de DMSO, adaptation du ratio enzyme/substrat, du temps de digestion pour l'étude de petites quantités de matériel de départ.

Nous nous sommes également concentrés sur la création d'un modèle de support miniaturisé en plaque afin de montrer que la miniaturisation de la surface de contact permet d'augmenter l'identification peptidique et protéique. Nous avons également vérifié l'efficacité et l'apport d'un traitement de surface sur la récupération peptidique puis sur les différentes étapes du protocole. Enfin, nous avons exploré l'étape de lyse cellulaire en fonction de la nature de nos supports et avons découvert qu'en plus de la miniaturisation des supports, leur nature et leur état de surface influaient aussi sur les identifications protéiques obtenues à partir des cellules.

Nous avons optimisé le protocole de préparation des cellules et le support avec lequel ces cellules sont en contact mais des problèmes de reproductibilité sont toujours rencontrés, plus spécialement lorsque nous travaillons sur de très petites quantités d'échantillons.

Plusieurs options de développement de ce protocole pourraient pallier les problèmes rencontrés et le projet pourrait être continué avec une automatisation de la distribution qui aurait limité les erreurs de reproductibilité, un *multiplexing* des échantillons grâce à des marqueurs de taille afin d'augmenter le débit de nos analyses, une amélioration des conditions de séparation en chromatographie liquide en testant d'autres colonnes de séparation comme les colonnes à micro-piliers  $\mu$ PAC plus adaptés à la protéomique à faible débit ou encore une amélioration des conditions d'acquisition en spectrométrie de masse par une analyse en DIA au lieu de l'acquisition stochastique du DDA.

Enfin, il est possible de changer de paradigme et d'adopter une stratégie inverse. En effet tout au long de ce projet, nous avons travaillé à minimiser l'adsorption des protéines et des peptides sur les contenants. Une stratégie différente consisterait à favoriser cette adsorption sur un support à forte affinité pouvant servir de point de concentration et de véhicule des peptides vers le système nano LC-MS/MS. Cette idée a déjà été mise en œuvre avec l'utilisation de billes fonctionnalisées pour l'extraction des peptides, mais aussi de purification de sous-classes de macro-biomolécules en exploitant des caractéristiques d'affinité particulières ou des approches de *click-chemistry*. Le transfert des molécules d'intérêt peut notamment se faire en utilisant des supports magnétiques (billes SP3 ou Dynabeads®), ou des micro-cartouches pour accumuler ces billes à l'image des micro-colonnes de piégeage (Preomics ou Evosep). Combiner nos travaux avec cette nouvelle stratégie s'apparente alors à utiliser des supports dont la surface minimise les interactions pour manipuler des volumes incluant ces supports véhicules maximisant les interactions. Finalement, il nous semble important de garder à l'esprit les stratégies de microfluidique de goutte qui permettent d'encapsuler les cellules afin d'éliminer les contacts aux surfaces, bien qu'actuellement l'efficacité d'encapsulation est confrontée à d'autres phénomènes rédhibitoires comme les fuites de composés par diffusion hors des gouttes et le manque de compatibilité de ces systèmes biphasiques avec la nano LC telle qu'elle est utilisée actuellement.

## V. BIBLIOGRAPHIE

- 1 Hooke, R. *Micrographia: or some physiological descriptions of minute bodies made by magnifying glasses, with observations and inquiries thereupon.* (Courier Corporation, 2003).
- 2 Karp, G. *Cell and molecular biology: concepts and experiments.* (John Wiley & Sons, 2009).
- 3 Alberts, B. *Molecular biology of the cell.* (WW Norton & Company, 2017).
- 4 Maton, A. *et al. Cells: Building blocks of life.* (Pearson Prentice Hall, 1997).
- 5 Elsasser, W. M. Outline of a theory of cellular heterogeneity. *Proceedings of the National Academy of Sciences* **81**, 5126-5129 (1984).
- 6 Rubin, H. The significance of biological heterogeneity. *Cancer and Metastasis Reviews* **9**, 1-20 (1990).
- 7 Altschuler, S. J. & Wu, L. F. Cellular heterogeneity: do differences make a difference? *Cell* **141**, 559-563 (2010).
- 8 Brock, A., Chang, H. & Huang, S. Non-genetic heterogeneity—a mutation-independent driving force for the somatic evolution of tumours. *Nature Reviews Genetics* **10**, 336-342 (2009).
- 9 Aird, W. C. Endothelial cell heterogeneity. *Cold Spring Harbor perspectives in medicine* **2**, a006429 (2012).
- 10 Wilson, A. *et al.* Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* **135**, 1118-1129 (2008).
- 11 Heppner, G. H. Tumor heterogeneity. *Cancer research* **44**, 2259-2265 (1984).
- 12 Bugianesi, E. Non-alcoholic steatohepatitis and cancer. *Clinics in liver disease* **11**, 191-207 (2007).
- 13 Gay, L., Baker, A.-M. & Graham, T. A. Tumour cell heterogeneity. *F1000Research* **5** (2016).
- 14 Junttila, M. R. & De Sauvage, F. J. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature* **501**, 346-354 (2013).
- 15 Qian, M., Wang, D. C., Chen, H. & Cheng, Y. in *Seminars in Cell & Developmental Biology*. 143-149 (Elsevier).
- 16 Cohen, S. J. *et al.* Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer. (2008).
- 17 Campbell, L. L. & Polyak, K. Breast tumor heterogeneity: cancer stem cells or clonal evolution? *Cell cycle* **6**, 2332-2338 (2007).
- 18 Gascoigne, K. E. & Taylor, S. S. Cancer cells display profound intra-and interline variation following prolonged exposure to antimetabolic drugs. *Cancer cell* **14**, 111-122 (2008).
- 19 Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183-1186 (2002).
- 20 Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & Van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nature genetics* **31**, 69-73 (2002).
- 21 Raser, J. M. & O'Shea, E. K. Control of stochasticity in eukaryotic gene expression. *science* **304**, 1811-1814 (2004).
- 22 Losick, R. & Desplan, C. Stochasticity and cell fate. *science* **320**, 65-68 (2008).



## BIBLIOGRAPHIE : Conclusion et perspectives

- 23 Eliaš, J., Dimitrio, L., Clairambault, J. & Natalini, R. The p53 protein and its molecular network: modelling a missing link between DNA damage and cell fate. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **1844**, 232-247 (2014).
- 24 Aldridge, S. & Teichmann, S. A. Single cell transcriptomics comes of age. *Nature Communications* **11**, 1-4 (2020).
- 25 Sánchez Barea, J., Lee, J. & Kang, D.-K. Recent advances in droplet-based microfluidic technologies for biochemistry and molecular biology. *Micromachines* **10**, 412 (2019).
- 26 Salomon, R. *et al.* Droplet-based single cell RNAseq tools: a practical guide. *Lab on a Chip* **19**, 1706-1727 (2019).
- 27 Arias-Hidalgo, C. *et al.* Single-Cell Proteomics: The Critical Role of Nanotechnology. *International Journal of Molecular Sciences* **23**, 6707 (2022).
- 28 Papalexí, E. & Satija, R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nature Reviews Immunology* **18**, 35-45 (2018).
- 29 Haque, A., Engel, J., Teichmann, S. A. & Lönnberg, T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome medicine* **9**, 1-12 (2017).
- 30 Kivioja, T. *et al.* Counting absolute numbers of molecules using unique molecular identifiers. *Nature methods* **9**, 72-74 (2012).
- 31 Ding, Y., Howes, P. D. & deMello, A. J. Recent advances in droplet microfluidics. *Analytical chemistry* **92**, 132-149 (2019).
- 32 Zhang, X. *et al.* Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-seq systems. *Molecular cell* **73**, 130-142. e135 (2019).
- 33 Jaitin, D. A. *et al.* Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776-779 (2014).
- 34 MacParland, S. A. *et al.* Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nature communications* **9**, 1-21 (2018).
- 35 Yates, J. R., Ruse, C. I. & Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annual review of biomedical engineering* **11**, 49-79 (2009).
- 36 Ramazi, S. & Zahiri, J. Post-translational modifications in proteins: resources, tools and prediction methods. *Database* **2021** (2021).
- 37 Chen, J. & Zheng, N. Accelerating protein biomarker discovery and translation from proteomics research for clinical utility. *Bioanalysis* **12**, 1469-1481 (2020).
- 38 Yarmush, M. L. & Jayaraman, A. Advances in proteomic technologies. *Annual review of biomedical engineering* **4**, 349-373 (2002).
- 39 Saah, A. J. & Hoover, D. R. "Sensitivity" and "specificity" reconsidered: the meaning of these terms in analytical and diagnostic settings. *Ann Intern Med* **126**, 91-94, doi:10.7326/0003-4819-126-1-199701010-00026 (1997).
- 40 Wilson, R. Sensitivity and specificity: twin goals of proteomics assays. Can they be combined? *Expert review of proteomics* **10**, 135-149 (2013).
- 41 Anderson, N. L. *et al.* The human plasma proteome: a nonredundant list developed by combination of four separate sources. *Molecular & Cellular Proteomics* **3**, 311-326 (2004).

## BIBLIOGRAPHIE : Conclusion et perspectives

- 42 Rogers, J. C. & Bomgardner, R. D. in *Modern Proteomics—Sample Preparation, Analysis and Practical Applications* 43-62 (Springer, 2016).
- 43 Lössl, P., van de Waterbeemd, M. & Heck, A. J. The diverse and expanding role of mass spectrometry in structural and molecular biology. *The EMBO journal* **35**, 2634-2657 (2016).
- 44 Shehadul Islam, M., Aryasomayajula, A. & Selvaganapathy, P. R. A review on macroscale and microscale cell lysis methods. *Micromachines* **8**, 83 (2017).
- 45 Johnson, B. H. & Hecht, M. H. Recombinant proteins can be isolated from E. coli cells by repeated cycles of freezing and thawing. *Bio/Technology* **12**, 1357-1360 (1994).
- 46 Watson, J., Cumming, R., Street, G. & Tuffnell, J. 105-109 (Ellis Horwood, London, 1987).
- 47 Harrison, S. T. Bacterial cell disruption: a key unit operation in the recovery of intracellular products. *Biotechnology advances* **9**, 217-240 (1991).
- 48 Lilly, M. & Dunnill, P. in *Fermentation Advances* 225-248 (Academic Press, London, 1969).
- 49 Cañas, B., Piñeiro, C., Calvo, E., López-Ferrer, D. & Gallardo, J. M. Trends in sample preparation for classical and second generation proteomics. *Journal of chromatography A* **1153**, 235-258 (2007).
- 50 Sharma, R. *et al.* Coupling a detergent lysis/cleanup methodology with intact protein fractionation for enhanced proteome characterization. *Journal of proteome research* **11**, 6008-6018 (2012).
- 51 Brownridge, P. & Beynon, R. J. The importance of the digest: proteolysis and absolute quantification in proteomics. *Methods* **54**, 351-360 (2011).
- 52 Kota, U. & Stolowitz, M. L. Improving proteome coverage by reducing sample complexity via chromatography. *Modern Proteomics—Sample Preparation, Analysis and Practical Applications*, 83-143 (2016).
- 53 Kelly, R. T. Single-cell proteomics: progress and prospects. *Molecular & Cellular Proteomics* **19**, 1739-1748 (2020).
- 54 Xie, F., Smith, R. D. & Shen, Y. Advanced proteomic liquid chromatography. *Journal of Chromatography A* **1261**, 78-90 (2012).
- 55 Wilson, S. R., Olsen, C. & Lundanes, E. Nano liquid chromatography columns. *Analyst* **144**, 7090-7104 (2019).
- 56 Swartz, M. E. & Murphy, B. New frontiers in chromatography. *American laboratory, February*, 22-27 (2005).
- 57 Shen, Y. & Smith, R. D. Proteomics based on high-efficiency capillary separations. *Electrophoresis* **23**, 3106-3124 (2002).
- 58 Reuhs, B. L. in *Food analysis* 213-226 (Springer, 2017).
- 59 Glish, G. L. & Vachet, R. W. The basics of mass spectrometry in the twenty-first century. *Nature reviews drug discovery* **2**, 140-150 (2003).
- 60 Yamashita, M. & Fenn, J. B. Electrospray ion source. Another variation on the free-jet theme. *The Journal of Physical Chemistry* **88**, 4451-4459 (1984).
- 61 Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64-71 (1989).

## BIBLIOGRAPHIE : Conclusion et perspectives

- 62 Alexandrov, M. *et al.* Extraction of ions from solutions under atmospheric pressure as a method for mass spectrometric analysis of bioorganic compounds. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry* **22**, 267-270 (2008).
- 63 Bruins, A. P. Mechanistic aspects of electrospray ionization. *Journal of Chromatography A* **794**, 345-357 (1998).
- 64 Vaidyanathan, S., Kell, D. B. & Goodacre, R. Selective detection of proteins in mixtures using electrospray ionization mass spectrometry: influence of instrumental settings and implications for proteomics. *Analytical chemistry* **76**, 5024-5032 (2004).
- 65 Levy, E. & Slavov, N. Single cell protein analysis for systems biology. *Essays in biochemistry* **62**, 595-605 (2018).
- 66 Wells, J. M. & McLuckey, S. A. Collision-induced dissociation (CID) of peptides and proteins. *Methods in enzymology* **402**, 148-185 (2005).
- 67 Jedrychowski, M. P. *et al.* Evaluation of HCD-and CID-type fragmentation within their respective detection platforms for murine phosphoproteomics. *Molecular & Cellular Proteomics* **10** (2011).
- 68 Roepstorff, P. Proposal for a nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spec.* **11**, 60 (1984).
- 69 TUINMAN, A. A. & PETTIT, G. R. Analysis of the amino acid sequence of peptides by mass spectrometry An ion notation proposal. *International Journal of Peptide and Protein Research* **36**, 331-334 (1990).
- 70 Sajic, T., Liu, Y. & Aebersold, R. Using data-independent, high-resolution mass spectrometry in protein biomarker research: perspectives and clinical applications. *PROTEOMICS—Clinical Applications* **9**, 307-321 (2015).
- 71 Bruderer, R. *et al.* New targeted approaches for the quantification of data-independent acquisition mass spectrometry. *Proteomics* **17**, 1700021 (2017).
- 72 Simbürger, J. M., Dettmer, K., Oefner, P. J. & Reinders, J. Optimizing the SWATH-MS-workflow for label-free proteomics. *Journal of proteomics* **145**, 137-140 (2016).
- 73 Hu, A., Noble, W. S. & Wolf-Yadlin, A. Technical advances in proteomics: new developments in data-independent acquisition. *F1000Research* **5** (2016).
- 74 Parker, S. J. *et al.* in *Manual of Cardiovascular Proteomics* 227-245 (Springer, 2016).
- 75 Zhang, F., Ge, W., Ruan, G., Cai, X. & Guo, T. Data-independent acquisition mass spectrometry-based proteomics and software tools: a glimpse in 2020. *Proteomics* **20**, 1900276 (2020).
- 76 Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nature methods* **17**, 41-44 (2020).
- 77 Hu, Q. *et al.* The Orbitrap: a new mass spectrometer. *Journal of mass spectrometry* **40**, 430-443 (2005).
- 78 Scigelova, M. & Makarov, A. Orbitrap mass analyzer—overview and applications in proteomics. *Proteomics* **6**, 16-21 (2006).

## BIBLIOGRAPHIE : Conclusion et perspectives

- 79 Gabelica, V. & Marklund, E. Fundamentals of ion mobility spectrometry. *Current opinion in chemical biology* **42**, 51-59 (2018).
- 80 Cumeras, R., Figueras, E., Davis, C. E., Baumbach, J. I. & Gracia, I. Review on ion mobility spectrometry. Part 1: current instrumentation. *Analyst* **140**, 1376-1390 (2015).
- 81 Liu, H., Sadygov, R. G. & Yates, J. R. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Analytical chemistry* **76**, 4193-4201 (2004).
- 82 Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. & Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Analytical and bioanalytical chemistry* **389**, 1017-1031 (2007).
- 83 Bondarenko, P. V., Chelius, D. & Shaler, T. A. Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography– tandem mass spectrometry. *Analytical chemistry* **74**, 4741-4749 (2002).
- 84 Chelius, D. & Bondarenko, P. V. Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *Journal of proteome research* **1**, 317-323 (2002).
- 85 Megger, D. A., Bracht, T., Meyer, H. E. & Sitek, B. Label-free quantification in clinical proteomics. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **1834**, 1581-1590 (2013).
- 86 Pan, S. & Aebersold, R. in *Mass spectrometry data analysis in proteomics* 209-218 (Springer, 2007).
- 87 Zhang, L. & Elias, J. E. in *Proteomics* 185-198 (Springer, 2017).
- 88 O'Connell, J. D., Paulo, J. A., O'Brien, J. J. & Gygi, S. P. Proteome-wide evaluation of two common protein quantification methods. *Journal of proteome research* **17**, 1934-1942 (2018).
- 89 Brenes, A., Hukelmann, J., Bensaddek, D. & Lamond, A. I. Multibatch TMT reveals false positives, batch effects and missing values. *Molecular & Cellular Proteomics* **18**, 1967-1980 (2019).
- 90 Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *ELECTROPHORESIS: An International Journal* **20**, 3551-3567 (1999).
- 91 Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature protocols* **11**, 2301-2319 (2016).
- 92 Milo, R. What is the total number of protein molecules per cell volume? A call to rethink some published values. *Bioessays* **35**, 1050-1055 (2013).
- 93 Milo, R. & Phillips, R. *Cell biology by the numbers*. (Garland Science, 2015).
- 94 Sa, C., Ganesan, N. & Lokeswari, T. Comparison of differentiation to macrophages in isolated monocytes from human peripheral blood and THP1 cells. *Sri Ramachandra Journal of Medicine* **7** (2014).
- 95 Lodish, H. *et al.* in *Molecular Cell Biology. 4th edition* (WH Freeman, 2000).
- 96 Krueve, A. *et al.* Tutorial review on validation of liquid chromatography–mass spectrometry methods: Part I. *Analytica chimica acta* **870**, 29-44 (2015).
- 97 Lee, M. C. G. & Sun, B. Quantitation of nonspecific protein adsorption at solid–liquid interfaces for single-cell proteomics. *Canadian Journal of Chemistry* **96**, 102-109 (2018).

## BIBLIOGRAPHIE : Conclusion et perspectives

- 98 Bennike, T., Birkelund, S., Stensballe, A. & Andersen, V. Biomarkers in inflammatory bowel diseases: current status and proteomics identification strategies. *World Journal of Gastroenterology: WJG* **20**, 3231 (2014).
- 99 Rardin, M. J. Rapid assessment of contaminants and interferences in mass spectrometry data using skyline. *Journal of The American Society for Mass Spectrometry* **29**, 1327-1330 (2018).
- 100 Keller, B. O., Sui, J., Young, A. B. & Whittall, R. M. Interferences and contaminants encountered in modern mass spectrometry. *Analytica chimica acta* **627**, 71-81 (2008).
- 101 Lehninger, A. Nelson DL, and Cox MM. Principles of Biochemistry. *New York: Worth* (1993).
- 102 Keil-Dlouha, V., Zylber, N., Imhoff, J.-M., Tong, N.-T. & Keil, B. Proteolytic activity of pseudotrypsin. *FEBS letters* **16**, 291-295 (1971).
- 103 Sun, B. & Kumar, S. Protein Adsorption Loss— The Bottleneck of Single-Cell Proteomics. *Journal of Proteome Research* **21**, 1808-1815 (2022).
- 104 Shen, L. & Zhu, J. Heterogeneous surfaces to repel proteins. *Advances in Colloid and Interface Science* **228**, 40-54 (2016).
- 105 Slavov, N. Driving single cell proteomics forward with innovation. *Journal of proteome research* **20**, 4915-4918 (2021).
- 106 Gonzalez de Castro, D., Clarke, P., Al-Lazikani, B. & Workman, P. Personalized cancer medicine: molecular diagnostics, predictive biomarkers, and drug resistance. *Clinical Pharmacology & Therapeutics* **93**, 252-259 (2013).
- 107 Jayamohan, H. *et al.* Advances in microfluidics and lab-on-a-chip technologies. *Molecular Diagnostics*, 197-217 (2017).
- 108 Chao, T. C. & Hansmeier, N. Microfluidic devices for high-throughput proteome analyses. *Proteomics* **13**, 467-479 (2013).
- 109 Groisman, A. *et al.* A microfluidic chemostat for experiments with bacterial and yeast cells. *Nature methods* **2**, 685-689 (2005).
- 110 Irimia, D., Tompkins, R. G. & Toner, M. Single-cell chemical lysis in picoliter-scale closed volumes using a microfabricated device. *Analytical chemistry* **76**, 6137-6143 (2004).
- 111 Jen, C.-P., Hsiao, J.-H. & Maslov, N. A. Single-cell chemical lysis on microfluidic chips with arrays of microwells. *Sensors* **12**, 347-358 (2011).
- 112 Sethu, P., Anahtar, M., Moldawer, L. L., Tompkins, R. G. & Toner, M. Continuous flow microfluidic device for rapid erythrocyte lysis. *Analytical chemistry* **76**, 6247-6253 (2004).
- 113 Tandiono, T. *et al.* Sonolysis of Escherichia coli and Pichia pastoris in microfluidics. *Lab on a Chip* **12**, 780-786 (2012).
- 114 Quinto-Su, P. A. *et al.* Examination of laser microbeam cell lysis in a PDMS microfluidic channel using time-resolved imaging. *Lab on a Chip* **8**, 408-414 (2008).
- 115 Kecskemeti, A. & Gaspar, A. Preparation and characterization of a packed bead immobilized trypsin reactor integrated into a PDMS microfluidic chip for rapid protein digestion. *Talanta* **166**, 275-283 (2017).
- 116 Qin, W. *et al.* Trypsin immobilization on hairy polymer chains hybrid magnetic nanoparticles for ultra fast, highly efficient proteome digestion, facile 18O labeling and absolute protein quantification. *Analytical chemistry* **84**, 3138-3144 (2012).

## BIBLIOGRAPHIE : Conclusion et perspectives

- 117 Cheng, G., Hao, S.-J., Yu, X. & Zheng, S.-Y. Nanostructured microfluidic digestion system for rapid high-performance proteolysis. *Lab on a Chip* **15**, 650-654 (2015).
- 118 Ji, J. *et al.* Proteolysis in microfluidic droplets: an approach to interface protein separation and peptide mass spectrometry. *Lab on a Chip* **12**, 2625-2629 (2012).
- 119 Luk, V. N. & Wheeler, A. R. A digital microfluidic approach to proteomic sample processing. *Analytical chemistry* **81**, 4524-4530 (2009).
- 120 Nelson, W. C. *et al.* Incubated protein reduction and digestion on an electrowetting-on-dielectric digital microfluidic chip for MALDI-MS. *Analytical chemistry* **82**, 9932-9937 (2010).
- 121 Ramsey, R. & Ramsey, J. Generating electrospray from microchip devices using electroosmotic pumping. *Analytical Chemistry* **69**, 1174-1178 (1997).
- 122 Chan, J. H., Timperman, A. T., Qin, D. & Aebersold, R. Microfabricated polymer devices for automated sample delivery of peptides for analysis by electrospray ionization tandem mass spectrometry. *Analytical chemistry* **71**, 4437-4444 (1999).
- 123 Gebreyesus, S. T. *et al.* Streamlined single-cell proteomics by an integrated microfluidic chip and data-independent acquisition mass spectrometry. *Nature communications* **13**, 1-13 (2022).
- 124 Petrosius, V. & Schoof, E. M. Recent advances in the field of single-cell proteomics. *Translational Oncology* **27**, 101556 (2023).
- 125 Chaudhari, G. A. & Sarje, S. H. Clean Room Classification for Pharmaceutical Industry. *Int. J. Eng. Tech. Res*, 241-244 (2015).
- 126 Picó, Y., Farré, M., Kantiani, L. & Barceló, D. Microfluidic Devices: Biosensors. *Chemical Analysis of Food: Techniques and Applications*, 177-217 (2012).
- 127 Voldman, J., Gray, M. L. & Schmidt, M. A. Microfabrication in biology and medicine. *Annual review of biomedical engineering* **1**, 401-425 (1999).
- 128 Goebel-Stengel, M., Stengel, A., Taché, Y. & Reeve Jr, J. R. The importance of using the optimal plasticware and glassware in studies involving peptides. *Analytical biochemistry* **414**, 38-46 (2011).
- 129 Bodas, D., Rauch, J.-Y. & Khan-Malek, C. Surface modification and aging studies of addition-curing silicone rubbers by oxygen plasma. *European Polymer Journal* **44**, 2130-2139 (2008).
- 130 Li, N. *et al.* Versatile antifouling coatings based on self-assembled oligopeptides for engineering and biological materials. *Journal of Materials Chemistry B* **7**, 2242-2246 (2019).
- 131 Hemmilä, S., Cauich-Rodríguez, J. V., Kreutzer, J. & Kallio, P. Rapid, simple, and cost-effective treatments to achieve long-term hydrophilic PDMS surfaces. *Applied Surface Science* **258**, 9864-9875 (2012).
- 132 Li, N., Schwartz, M. & Ionescu-Zanetti, C. PDMS compound adsorption in context. *SLAS Discovery* **14**, 194-202 (2000).
- 133 Boscaini, E., Alexander, M. L., Prazeller, P. & Märk, T. D. Investigation of fundamental physical properties of a polydimethylsiloxane (PDMS) membrane using a proton transfer reaction mass spectrometer (PTRMS). *International Journal of Mass Spectrometry* **239**, 179-186 (2004).

## BIBLIOGRAPHIE : Conclusion et perspectives

- 134 Whitesides, G. M., Ostuni, E., Takayama, S., Jiang, X. & Ingber, D. E. Soft lithography in biology and biochemistry. *Annual review of biomedical engineering* **3**, 335-373 (2001).
- 135 Specht, H. *et al.* Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. *Genome biology* **22**, 1-27 (2021).
- 136 Zhu, Y. *et al.* Nanodroplet processing platform for deep and quantitative proteome profiling of 10–100 mammalian cells. *Nature communications* **9**, 1-10 (2018).
- 137 Trantidou, T., Elani, Y., Parsons, E. & Ces, O. Hydrophilic surface modification of PDMS for droplet microfluidics using a simple, quick, and robust method via PVA deposition. *Microsystems & nanoengineering* **3**, 1-9 (2017).
- 138 Wong, I. & Ho, C.-M. Surface molecular property modifications for poly (dimethylsiloxane)(PDMS) based microfluidic devices. *Microfluidics and nanofluidics* **7**, 291-306 (2009).
- 139 Harris, J. M. *Poly (ethylene glycol) chemistry: biotechnical and biomedical applications*. (Springer Science & Business Media, 1992).
- 140 Zalipsky, S. & Harris, J. M. (ACS Publications, 1997).
- 141 Doherty, E. A., Meagher, R. J., Albarghouthi, M. N. & Barron, A. E. Microchannel wall coatings for protein separations by capillary and chip electrophoresis. *Electrophoresis* **24**, 34-54 (2003).
- 142 Dolník, V. Wall coating for capillary electrophoresis on microchips. *Electrophoresis* **25**, 3589-3601 (2004).
- 143 Hollahan, J. R. & Bell, A. T. Techniques and applications of plasma chemistry. (1974).
- 144 Owen, M. J. & Smith, P. J. Plasma treatment of polydimethylsiloxane. *Journal of adhesion science and technology* **8**, 1063-1075 (1994).
- 145 Hillborg, H. *et al.* Crosslinked polydimethylsiloxane exposed to oxygen plasma studied by neutron reflectometry and other surface specific techniques. *Polymer* **41**, 6851-6863 (2000).
- 146 Tan, S. H., Nguyen, N.-T., Chua, Y. C. & Kang, T. G. Oxygen plasma treatment for reducing hydrophobicity of a sealed polydimethylsiloxane microchannel. *Biomicrofluidics* **4**, 032204 (2010).
- 147 Vickers, J. A., Caulum, M. M. & Henry, C. S. Generation of hydrophilic poly (dimethylsiloxane) for high-performance microchip electrophoresis. *Analytical chemistry* **78**, 7446-7452 (2006).
- 148 Brittain, W. J., Advincula, R. C., Caster, K. C. & R  he, J. *Polymer Brushes: Synthesis, Characterization and Applications*. (John Wiley & Sons, 2006).
- 149 Xiao, Y. & Isaacs, S. N. Enzyme-linked immunosorbent assay (ELISA) and blocking with bovine serum albumin (BSA)—not all BSAs are alike. *Journal of immunological methods* **384**, 148-151 (2012).
- 150 Jitsukawa, T., Nakajima, S., Sugawara, I. & Watanabe, H. Increased coating efficiency of antigens and preservation of original antigenic structure after coating in ELISA. *Journal of immunological methods* **116**, 251-257 (1989).
- 151 Bratcher, P. E. & Gaggar, A. Characterization and prevention of the adsorption of surfactant protein D to polypropylene. *PLoS One* **8**, e73467 (2013).

## BIBLIOGRAPHIE : Conclusion et perspectives

- 152 Mohan, T. *et al.* Highly Protein Repellent and Antiadhesive Polysaccharide Biomaterial Coating for Urinary Catheter Applications. *ACS Biomaterials Science & Engineering* **5**, 5825-5832 (2019).
- 153 Granville, A. M. & Brittain, W. J. Recent advances in polymer brush synthesis. *Polymer Brushes: Synthesis, Characterization, Applications*, 33-50 (2004).
- 154 Minko, S. in *Polymer surfaces and interfaces* 215-234 (Springer, 2008).
- 155 Macchione, M. A., Biglione, C. & Strumia, M. Design, synthesis and architectures of hybrid nanomaterials for therapy and diagnosis applications. *Polymers* **10**, 527 (2018).
- 156 Suresh, D., Goh, P. S., Ismail, A. F. & Hilal, N. Surface Design of Liquid Separation Membrane through Graft Polymerization: A State of the Art Review. *Membranes* **11**, 832 (2021).
- 157 Carneiro, L. B., Ferreira, J., Santos, M. J., Monteiro, J. P. & Giroto, E. M. A new approach to immobilize poly (vinyl alcohol) on poly (dimethylsiloxane) resulting in low protein adsorption. *Applied surface science* **257**, 10514-10519 (2011).
- 158 Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome biology* **19**, 1-12 (2018).
- 159 Shahi, P., Kim, S. C., Haliburton, J. R., Gartner, Z. J. & Abate, A. R. Abseq: Ultrahigh-throughput single cell protein profiling with droplet microfluidic barcoding. *Scientific reports* **7**, 1-12 (2017).
- 160 Ornatsky, O. *et al.* Study of cell antigens and intracellular DNA by identification of element-containing labels and metallointercalators using inductively coupled plasma mass spectrometry. *Analytical chemistry* **80**, 2539-2547 (2008).
- 161 Tracey, L. J., An, Y. & Justice, M. J. CyTOF: An Emerging Technology for Single-Cell Proteomics in the Mouse. *Current Protocols* **1**, e118 (2021).
- 162 Bandura, D. R. *et al.* Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Analytical chemistry* **81**, 6813-6822 (2009).
- 163 Ornatsky, O. *et al.* Highly multiparametric analysis by mass cytometry. *Journal of immunological methods* **361**, 1-20 (2010).
- 164 Broome, A.-M. *Cancer Nanotechnology*. (Academic Press, 2018).
- 165 Marx, V. A dream of single-cell proteomics. *Nature Methods* **16**, 809-812 (2019).
- 166 Hughes, A. J. *et al.* Single-cell western blotting. *Nature methods* **11**, 749-755 (2014).
- 167 Wu, D., Milutinovic, M. D. & Walt, D. R. Single molecule array (Simoa) assay with optimal antibody pairs for cytokine detection in human serum samples. *Analyst* **140**, 6277-6282 (2015).
- 168 Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202-1214 (2015).
- 169 Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nature methods* **14**, 865-868 (2017).
- 170 Shao, X. *et al.* Integrated proteome analysis device for fast single-cell protein profiling. *Analytical chemistry* **90**, 14003-14010 (2018).



## BIBLIOGRAPHIE : Conclusion et perspectives

- 171 Li, Z.-Y. *et al.* Nanoliter-scale oil-air-droplet chip-based single cell proteomic analysis. *Analytical chemistry* **90**, 5430-5438 (2018).
- 172 Zarini, S., Barkley, R. M., Gijón, M. A. & Murphy, R. C. in *High-Throughput Metabolomics* 81-105 (Springer, 2019).
- 173 Hájek, R., Jirásko, R., Lída, M., Cífková, E. & Holčápek, M. Hydrophilic interaction liquid chromatography–mass spectrometry characterization of gangliosides in biological samples. *Analytical Chemistry* **89**, 12425-12432 (2017).
- 174 Hutchins, P. M., Barkley, R. M. & Murphy, R. C. Separation of cellular nonpolar neutral lipids by normal-phase chromatography and analysis by electrospray ionization mass spectrometry. *Journal of lipid research* **49**, 804-813 (2008).
- 175 Chen, A. T., Franks, A. & Slavov, N. DART-ID increases single-cell proteome coverage. *PLoS computational biology* **15**, e1007082 (2019).
- 176 Leduc, A., Huffman, R. G., Cantlon, J., Khan, S. & Slavov, N. Exploring functional protein covariation across single cells using nPOP. *bioRxiv*, 2021.2004. 2024.441211 (2022).
- 177 Derks, J. *et al.* Increasing the throughput of sensitive proteomics by plexDIA. *Nature Biotechnology*, 1-10 (2022).
- 178 Woo, J. *et al.* High-throughput and high-efficiency sample preparation for single-cell proteomics using a nested nanowell chip. *Nature communications* **12**, 1-11 (2021).
- 179 Hartlmayr, D. *et al.* An automated workflow for label-free and multiplexed single cell proteomics sample preparation at unprecedented sensitivity. *bioRxiv* **10**, 14.439828 (2021).
- 180 Lee, Y. *et al.* ProteoChip: A highly sensitive protein microarray prepared by a novel method of protein immobilization for application of protein-protein interaction studies. *Proteomics* **3**, 2289-2304 (2003).
- 181 Mund, A., Brunner, A.-D. & Mann, M. Unbiased spatial proteomics with single-cell resolution in tissues. *Molecular Cell* **82**, 2335-2349 (2022).
- 182 Ctorteka, C. *et al.* An automated workflow for multiplexed single-cell proteomics sample preparation at unprecedented sensitivity. (2021).
- 183 Cupp-Sutton, K. A., Fang, M. & Wu, S. Separation methods in single-cell proteomics: RPLC or CE? *International journal of mass spectrometry* **481**, 116920 (2022).
- 184 Nie, S. *et al.* Maximizing hydrophobic peptide recovery in proteomics and antibody development using a mass spectrometry compatible surfactant. *Analytical Biochemistry* **658**, 114924 (2022).
- 185 Shen, Y. *et al.* High-efficiency nanoscale liquid chromatography coupled on-line with mass spectrometry using nanoelectrospray ionization for proteomics. *Analytical chemistry* **74**, 4235-4249 (2002).
- 186 Zhu, Y. *et al.* Subnanogram proteomics: Impact of LC column selection, MS instrumentation and data analysis strategy on proteome coverage for trace samples. *International journal of mass spectrometry* **427**, 4-10 (2018).
- 187 Cong, Y. *et al.* Improved single-cell proteome coverage using narrow-bore packed NanoLC columns and ultrasensitive mass spectrometry. *Analytical chemistry* **92**, 2665-2671 (2020).
- 188 Choi, S. B., Polter, A. M. & Nemes, P. Patch-Clamp Proteomics of Single Neurons in Tissue Using Electrophysiology and Subcellular Capillary Electrophoresis Mass Spectrometry. *Analytical chemistry* **94**, 1637-1644 (2021).

## BIBLIOGRAPHIE : Conclusion et perspectives

- 189 Lombard-Banek, C., Moody, S. A., Manzini, M. C. & Nemes, P. Microsampling capillary electrophoresis mass spectrometry enables single-cell proteomics in complex tissues: developing cell clones in live *Xenopus laevis* and zebrafish embryos. *Analytical chemistry* **91**, 4797-4805 (2019).
- 190 Phlairaharn, T. *et al.* High sensitivity limited material proteomics empowered by data-independent acquisition on linear ion traps. *Journal of Proteome Research* (2022).
- 191 Cui, M., Cheng, C. & Zhang, L. High-throughput proteomics: a methodological mini-review. *Laboratory Investigation*, 1-12 (2022).
- 192 Virant-Klun, I., Leicht, S., Hughes, C. & Krijgsveld, J. Identification of maturation-specific proteins by single-cell proteomics of human oocytes. *Molecular & Cellular Proteomics* **15**, 2616-2627 (2016).
- 193 Brunner, A. D. *et al.* Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *Molecular systems biology* **18**, e10798 (2022).
- 194 Orsburn, B. C., Yuan, Y. & Bumpus, N. N. Insights into protein post-translational modification landscapes of individual human cells by trapped ion mobility time-of-flight mass spectrometry. *Nature communications* **13**, 1-14 (2022).
- 195 Slavov, N. Scaling up single-cell proteomics. *Molecular & Cellular Proteomics* **21** (2022).
- 196 Specht, H. & Slavov, N. Transformative opportunities for single-cell proteomics. *Journal of Proteome Research* **17**, 2565-2571 (2018).
- 197 Agard, N. J., Maltby, D. & Wells, J. A. Inflammatory stimuli regulate caspase substrate profiles. *Molecular & Cellular Proteomics* **9**, 880-893 (2010).
- 198 Yang, L., Liu, J., Reddy, R. & Wang, J. Single-Cell Proteomics Study of Immune Cell Diversity by Quantitating 465 Proteins. *bioRxiv* (2022).
- 199 Weiss, A., Wiskocil, R. & Stobo, J. The role of T3 surface molecules in the activation of human T cells: a two-stimulus requirement for IL 2 production reflects events occurring at a pre-translational level. *The Journal of Immunology* **133**, 123-128 (1984).
- 200 Ren, X. *et al.* Blockade of the immunosuppressive KIR2DL5/PVR pathway elicits potent human NK cell-mediated antitumor immunity. *Journal of Clinical Investigation* **132**, e163620 (2022).
- 201 Ford, E. S. *et al.* CD8+ T cell clonotypes from prior SARS-CoV-2 infection predominate during the cellular immune response to mRNA vaccination. (2022).
- 202 Gross, A. *et al.* Technologies for single-cell isolation. *International journal of molecular sciences* **16**, 16897-16919 (2015).
- 203 Cossarizza, A. *et al.* Guidelines for the use of flow cytometry and cell sorting in immunological studies. *European journal of immunology* **49**, 1457-1973 (2019).
- 204 Mouradian, S. Lab-on-a-chip: applications in proteomics. *Current opinion in chemical biology* **6**, 51-56 (2002).
- 205 Hansen, K. C. *et al.* An in-solution ultrasonication-assisted digestion method for improved extracellular matrix proteome coverage. *Molecular & Cellular Proteomics* **8**, 1648-1657 (2009).
- 206 Strzelecka, D., Holman, S. W. & Eyers, C. E. Evaluation of dimethyl sulfoxide (DMSO) as a mobile phase additive during top 3 label-free quantitative proteomics. *International journal of mass spectrometry* **391**, 157-160 (2015).

## BIBLIOGRAPHIE : Conclusion et perspectives

- 207 Haaf, F., Sanner, A. & Straub, F. Polymers of N-vinylpyrrolidone: synthesis, characterization and uses. *Polymer Journal* **17**, 143-152 (1985).
- 208 compounds are defined by IUPAC, M. as “dipolar five-(possibly six-) membered heterocyclic compounds in which both the negative and the positive charge are delocalized, for which a totally covalent structure cannot be written, and which cannot be represented satisfactorily by any one polar structure. The formal positive charge is associated with the ring atoms, and the formal negative charge is associated with ring atoms or an exocyclic nitrogen or chalcogen atom”. IUPAC. Compendium of Chemical Terminology, 2nd edn. Gold Book Compiled by AD McNaught and A. Wilkinson. *Compendium of Chemical Terminology, 2nd Ed. Gold Book Compiled* (1997).
- 209 McNaught, A. D. & Wilkinson, A. *Compendium of chemical terminology*. Vol. 1669 (Blackwell Science Oxford, 1997).
- 210 Zhou, W., Zhang, W., Liu, Y., Yu, X. & Chen, Z. Polydopamine-functionalized poly (ether ether ketone) tube for capillary electrophoresis-mass spectrometry. *Analytica chimica acta* **987**, 64-71 (2017).
- 211 Zhu, Z., Chen, P., Liu, K. & Escobedo, C. A versatile bonding method for PDMS and SU-8 and its application towards a multifunctional microfluidic device. *Micromachines* **7**, 230 (2016).
- 212 Xue, P. *et al.* Surface modification of poly (dimethylsiloxane) with polydopamine and hyaluronic acid to enhance hemocompatibility for potential applications in medical implants or devices. *ACS applied materials & interfaces* **9**, 33632-33644 (2017).
- 213 Cruciani, G. *et al.* Peptide studies by means of principal properties of amino acids derived from MIF descriptors. *Journal of Chemometrics* **18**, 146-155 (2004).
- 214 Osorio, D., Rondón-Villarreal, P. & Torres, R. Peptides: a package for data mining of antimicrobial peptides. *Small* **12**, 44-444 (2015).
- 215 Ariati, R., Sales, F., Souza, A., Lima, R. A. & Ribeiro, J. Polydimethylsiloxane composites characterization and its applications: A review. *Polymers* **13**, 4258 (2021).
- 216 Hong, S., Wang, R., Huang, X. & Liu, H. Facile one-step fabrication of PHC/PDMS anti-icing coatings with mechanical properties and good durability. *Progress in Organic Coatings* **135**, 263-269 (2019).
- 217 Kameya, Y. & Takada, Y. Prolonged water freezing on a superhydrophobic micropillar polydimethylsiloxane film. *Results in Materials* **14**, 100274 (2022).

## VI. ANNEXES

### Chapitre 10 Matériels et méthodes

#### *Matériel*

Le polydiméthylsiloxane (PDMS) Sylgard 184 provient de Dow Corning, le photorésist SU-2000 de Microchem et la plaquette de silicium d'Entegris. Le digest de HeLa provient de Pierce (Thermo Fisher Scientific). Les digests de levure et les extraits protéiques de levure et d'humain proviennent de Promega. Les cellules proviennent de l'ATCC. Le PBS et l'eau servant à la culture cellulaire viennent de Gibco. Les flasques de culture cellulaire viennent de Corning. L'iodoacétamide (IAM), dithiothréitol (DTT), la protéine standard BSA et les solvants (TFA, FA et BCA) provenaient de Sigma Aldrich. Le mix de trypsine (gold) et Lys C provient de Promega. L'eau Milli-Q a été produite par un IQ 7000 de Millipore.

Les plaques 384-puits en PP de 120 $\mu$ l par puits proviennent de Greiner Bio. Les plaques 384-puits en PP de 58 $\mu$ l par puits proviennent de ThermoFisher Scientific). Les plaques MALDI Opti TOF proviennent de AB Sciex et les plaques *AnchorChip* de Bruker. Le fil de polypropylène provient de Volumic.

Tous les dispositifs microfluidiques, utilisés dans cette étude, ont été fabriqués au sein de la plateforme technologique de l'Institut Pierre Gilles de Gennes pour la microfluidique (IPGG).

#### 10.A MICRO-FABRICATION POUR LES ESSAIS DE TRAITEMENTS DE SURFACE

##### 10.A.1. Fabrication du moule en résine

Pour fabriquer les puces voulues il faut d'abord un moule (ou *wafer*) sur lequel on coulera le PDMS. Nous avons utilisé la lithographie douce pour créer les deux moules. La première étape consiste à préparer un masque correspondant au dessin des puces voulues à partir de film photosensible. Les motifs sont imprimés sur le film à l'aide d'un laser (on noircit les zones où passe le laser) puis on développe le masque dans deux bains. Une fois en possession du masque, on imprime les motifs en 3D sur une plaque de silicone. On coule sur cette plaque une couche de résine uniforme de 60 $\mu$ m d'épaisseur obtenue par mise en rotation de la plaque à 2500rpm pour la résine SU8-2050. La résine a été choisie pour pouvoir atteindre la bonne épaisseur à un nombre raisonnable de rpm. On place le tout sur plaque chauffante à 65°C 1min30s (préchauffage) puis à 95°C 7min. On place ensuite sous lumière UV à une puissance de 23,3W pendant 8s. Le masque est placé entre la lampe et la couche de résine pour la polymérisation de cette dernière ne soit activée que dans les zones où le masque laisse passer la lumière UV. On place sur plaque chauffante à 65°C 1min puis à 95°C 7min. On plonge la plaque dans la solution de développement et on agite pendant 6min. On finalise en laissant le *wafer* sur une plaque chauffante à 200°C 5min.

##### 10.A.2. Fabrication des puces en PDMS

Mélanger 50g de PDMS 1:10 (réticulant : base) de type Sylgard184.

Centrifuger 4min à 4000rpm.

Entourer le *wafer* de papier aluminium puis couler le PDMS.

## ANNEXES : Matériels et méthodes

Placer sous vide 1h minimum.

Placer ensuite au four à 70°C pour 1h minimum.

La plaque de PDMS obtenue est ensuite découpée et les puces perforées pour créer les entrées des canaux. Elles sont placées sous plasma d'oxygène 1 min, collées à des lamelles de verres et passées au four à 90°C 15min.

### 10.A.3. Vérification des canaux

Pour vérifier la bonne forme et l'absence de fuite dans les puces nous avons fait passer des colorants à l'intérieur.

### 10.A.4. Protocoles de traitement sur PDMS

#### (a) Dopamine tamponnée (pH = 8.8) sans oxydation

Préparer 50mL de tampon Tris-HCl à 50mM et pH = 8.8 (dissoudre 310mg, pH rectifié au HCl 10%).

Y dissoudre la dopamine (250mg) pour obtenir une solution à 5mg/ml.

Nettoyer les canaux à l'HCl 10% et sécher à l'azote.

Charger la solution de dopamine tamponnée et renouveler toutes les 2h pendant 72h.

Rincer abondamment à l'eau milli-Q et sécher à l'azote.

#### (b) Dopamine tamponnée (pH = 8.8) avec oxydation préliminaire

Préparer 50mL de tampon Tris-HCl à 50mM et pH = 8.8 (dissoudre 310mg, pH rectifié au HCl 10%).

En prélever 10mL et dissoudre la dopamine (50mg) pour obtenir une solution à 5mg/ml (opération à répéter).

Préparer 12mL de solution d'oxydation  $H_2O$  milli-Q :  $H_2O_2$  30% : HCl 37% à 5 : 1 : 1 en volume.

Nettoyer les canaux à l'HCl 10% et sécher à l'azote.

Charger la solution d'oxydation et laisser réagir 45min en renouvelant deux fois.

Charger la solution de dopamine tamponnée et renouveler toutes les 2h pendant 72h.

Rincer à l'eau milli-Q (2mL + bain de 48h à 37°C) et sécher à l'azote.

#### (c) Dopamine tamponnée (pH = 5) avec oxydation

Préparer 50mL de tampon AcONa à 50mM et pH = 5 (dissoudre 340mg de  $AcONa(H_2O)_3$ , pH rectifié avec AcOH(l)).

En prélever 10mL et dissoudre la dopamine (20mg) ainsi que le  $NaIO_4(s)$  (0,50mg) pour obtenir une solution à 2mg/ml en dopamine et à 20mM en  $NaIO_4$ .

Nettoyer les canaux à l'HCl 10% et sécher à l'azote.

Charger la solution et renouveler toutes les 2h pendant 24h.

Rincer à l'eau milli-Q (3mL) et sécher à l'azote.

## ANNEXES : Matériels et méthodes

### (d) Traitements par le PVA

Dissoudre 1g de PVA dans 100mL d'eau Milli-Q.

Placer sous agitation à 25°C 40 min.

Placer sous agitation à 100°C 40 min.

Placer sous agitation à 65°C 3h.

Laver les canaux à l'HCl 10%

Introduire la solution d'oxydation  $H_2O$  Milli-Q :  $H_2O_2$  30% :  $HCl$  37% à 5 :1 :1 en volume.

Laisser poser 35 min avec un renouvellement.

Introduire la solution de PVA 1% et laisser poser 10 min à 25°C.

Rincer à l'eau milli-Q (1mL) et sécher à l'azote.

Passer à 110°C pendant 15 min sur une plaque chauffante.

### (e) Traitements par le PVP

Préparer une solution de PVP à 0.2%.

Introduire la solution dans la puce tout juste sortie de 10min au four 90°C après collage plasma et laisser poser 10 min.

Rincer à l'eau milli-Q (1mL) et sécher à l'azote.

### (f) Traitements par l'APTES

Introduire la solution d'oxydation  $H_2O$  milli-Q :  $H_2O_2$  30% :  $HCl$  37% à 5 :1 :1 en volume.

Laisser poser 20 min.

Introduire (très doucement) la solution d'APTES 5% dans le chloroforme.

Laisser poser 24h en empêchant l'évaporation au maximum.

Rincer à l'éthanol (2mL) puis à l'eau milli-Q (2mL).

### (g) Traitements par le $\beta$ - octyl

Introduire la solution d'oxydation  $H_2O$  milli-Q :  $H_2O_2$  30% :  $HCl$  37% à 5 :1 :1 en volume.

Laisser poser 20 min.

Sécher à l'azote.

Introduire la solution de  $\beta$  - octyl 1% dans le tampon Tris-HCl 50mM, pH = 8.8

Laisser poser 24h.

Rincer à l'eau milli-Q (1mL) et sécher à l'azote.

### (h) Traitements par l'acide hyaluronique (HA)

Introduire la solution d'oxydation  $H_2O$  milli-Q :  $H_2O_2$  30% :  $HCl$  37% à 5 :1 :1 en volume.

Laisser poser 20 min.

Sécher à l'azote.

Introduire une première solution d'HA 1% dans le tampon Tris-HCl 50mM, pH = 8.8

Laisser poser 20 min.

Introduire une seconde solution d'HA 0.1% dans le tampon Tris-HCl 50mM, pH = 8.8

Laisser poser 24h.

Rincer à l'eau milli-Q (1mL) et sécher à l'azote.

## 10.B MICRO-FABRICATION DES SUPPORTS

### 10.B.1. Par photolithographie

Un design de puce a été dessiné sur le logiciel AutoCAD (*student* version, 2020). Ces dessins ont été enregistrés au format « .DXF » entourés d'encoches pour aider à repérer les motifs plus facilement et disposés dans un carré de 6cmx 6cm qui représente le *wafer* de photolithographie. Puis, le fichier est converti en « .CIF » avec le logiciel Klayout (version 0.26.3) pour délimiter les designs puis en « .GBR » avec le logiciel Clewin5 pour l'impression du masque de photolithographie par l'ingénieur de la plateforme de l'IPGG, Kevin Phan.

Après impression du masque, il faut préparer le *wafer* pour le processus de lithographie. Après avoir chauffé le *wafer* à 200°C sur une plaque chauffante pendant 15 minutes. Un film sec (*dry film*) photosensible de 50µm d'épaisseur est laminé grâce à une lamineuse sur le *wafer*. Sur l'appareil de lithographie, le masque est apposé sur le *wafer* laminé et le processus de lithographie est lancé selon les paramètres suivants pour un programme continu : énergie à 95mJ pendant 4 secondes, épaisseur : 350µm et distance entre le masque et le *wafer* : 0 cm. Selon les dessins sur le masque, la résine négative du film va photo-réticuler là où elle est éclairée. Le *wafer* est ensuite chauffé pendant 1 minute à 65°C et 5 minutes à 100°C pour arrêter le processus de photo-réticulation. C'est ce qu'on appelle le « *post exposure bake* ». Pour le film sec choisi, le développeur utilisé est le cyclohexane pendant 10min puis rincé à l'isopropanol. Enfin, afin d'améliorer l'adhésion avec le substrat on recuit le *wafer* à 200°C pendant 2minutes. C'est ce qu'on appelle le « *hard bake* ».

Une fois que le *wafer* est prêt, il peut être réutilisé indéfiniment pour la fabrication de puces. Il est entouré de papier aluminium et 50g de PDMS (Sylgard 184) et son réticulant en ratio 10 :1 sont coulés sur le *wafer*. Le *wafer* est placé dans une pompe sous vide pendant 1 heure afin de retirer toutes les bulles d'air et enfin, placé à 70°C pendant 2 heures jusqu'à durcissement du PDMS.

Le PDMS durci est retiré du *wafer*. Les puces sont découpées et à l'aide de poinçons, des trous de la taille requise sont poinçonnés selon les demandes des designs pour former les puces. Les puces sont ensuite traitées au plasma à oxygène en même temps que des lames en verre afin d'activer les surfaces et les coller ensemble de façon permanente. Les puces sont ensuite traitées avec une solution de PVP à 0.2% (v/v) pendant 10 minutes afin de les rendre plus hydrophiles et lavées à l'eau.

### 10.B.2. Par micro-usinage

Le micro-usinage des moules a été fait sur la plateforme de l'Institut Pierre Gilles de Gennes (IPGG) par les ingénieurs Dr. Guillaume Votte (1<sup>er</sup> design de moule en laiton) et Dr. Audric Jan (2<sup>ème</sup> design de moule en laiton) à partir d'un moule en laiton de 8cm x 8cm vierge.

Le dessin des moules a été fait sur le logiciel Fusion 360 puis transféré grâce à un logiciel sur l'usineuse Mini-Mill/GX (Minitex Machinery, USA). Une fraise de 3mm de diamètre et 6mm de longueur a été utilisée pour la première partie de l'usinage puis une fraise de 0.5mmx0.5mm a été utilisée pour fabriquer les rampes et la finition des plots.

Une fois le moule en laiton prêt, il peut être réutilisé indéfiniment pour la fabrication de puces. Il est entouré de papier aluminium et 50g de PDMS (Sylgard 184) et son réticulant en ratio 10 :1 sont coulés sur le moule. Il est placé dans une pompe sous vide pendant 1 heure afin de retirer toutes les bulles d'air et enfin, placé à 70°C pendant 2 heures jusqu'à durcissement du PDMS. Le PDMS durci est retiré du moule. Les micro-puits sont ensuite traités au plasma à oxygène afin d'activer les surfaces puis avec une solution de PVP à 0.2% (v/v) pendant 10 minutes afin de les rendre plus hydrophiles et lavés à l'eau.

### 10.C PREPARATION DES ECHANTILLONS

#### 10.C.1. CULTURE CELLULAIRE

Les cellules sont mises en culture dans un milieu RPMI enrichi (RPMO 1640, Sigma Aldrich) avec du bicarbonate de sodium, sans L-glutamine dans lequel sont ajoutés 10% de sérum fœtal bovin (FBS, Sigma), 1% de pyruvate de sodium 100mM (Gibco), 91µl de beta mercaptoéthanol 1000X à 55mM dans du DPBS (Gibco), 1% de L-glutamine 200mM (Gibco) et 1% de pénicilline streptomycine. Les cellules sont cultivées dans des flasques de 75µm<sup>2</sup> (Corning) et sont repassées à 0.3 millions de cellules par ml tous les 3 jours ou quand une confluence de 1.5 millions de cellules par ml est atteinte.

#### 10.C.2. VOLUME DES REACTIONS

Afin de limiter les ratios volumes/nombre de cellules et d'augmenter l'efficacité des réactions étudiées, toutes les expérimentations sont faites dans un volume inférieur à 10µl (environ 5µl). Le système LC utilise soit un UDP-1 µl (pour *user defined program*) injectant 1µl de l'échantillon dans une boucle de 20µl pour l'U-3000 connectée au Q Exactive HF ou une « *full loop* » d'1µl prélevant environ 5µl et injectant 1µl de l'échantillon provenant de la boucle d'1µl dans la colonne analytique pour l'U-3000 connectée au Tims-TOF Pro. Ainsi, pour 50ng de protéines déposées, l'équivalent de 10ng est analysé. De la même façon, pour 200 cellules déposées dans un puits, l'équivalent de 40 cellules est analysé.

#### 10.C.3. PREPARATION DES ECHANTILLONS

La lyse thermique des cellules consiste en un cycle de congélation à -80°C pendant une nuit puis une décongélation à température ambiante.

Les différentes méthodes dans la partie résultats décrivent les échantillons analysés et l'ajout des réactifs dont le ratio de trypsine/substrat utilisé. Lorsque ce n'est pas indiqué, le temps de digestion est de 3 heures à 37°C.

Les protocoles ERAD et ED sont décrits ici. Le protocole ERAD consiste en 1µl d'extrait protéique à 50ng/µl réduit par 1µl de dithiothréitol (DTT) à 55mM pendant 15 minutes à 37°C, alkylé par 1µl d'iodoacétamide (IAA) à 30mM pendant 15 minutes à température ambiante dans l'obscurité puis digérés par 1µl de trypsine – Lys C (ratio 1:20 (E/S)) pendant 1h à 37°C. Le protocole ED consiste en 1µl d'extrait protéique à 50ng/µl auquel est ajouté 2µl d'ammonium bicarbonate (BCA à 50mM) qui est digéré par 1µl de trypsine – Lys C (ratio 1:20 (E/S)) pendant 1h à 37°C.



Pour toutes nos expériences, de l'acide formique à 0.2% *qsp* 5µl a été ajouté pour arrêter la réaction de digestion et récupérer les peptides. Puis les peptides sont analysés en nano LC-MS/MS.

#### 10.D FABRICATION DE LA COLONNE *HOMEMADE*

La colonne *homemade* a été fabriquée à l'institut Pasteur. Cette fabrication comporte plusieurs étapes dont la première consiste à choisir un capillaire vide de diamètre interne déterminé (75 µm dans notre cas) et une phase (C18, taille des particules : 1.9 µm). Tout d'abord, un mélange de Kasil et de formamide formant le fritté est déposé sur un papier Wattman puis placé en bout de capillaire afin de se solidifier et d'empêcher la fuite de la phase. Puis, la phase est insérée dans la colonne grâce à la pression d'un gaz (ici l'argon) délivré par une pompe fixée à 60bar jusqu'à la longueur souhaitée (ici 15.7cm sont remplies en 30 à 45min). L'étape suivante est facultative et consiste à étirer le capillaire de la colonne grâce à une étireuse laser pour en faire un embout de moins de 10µm de diamètre interne aussi appelé *tip*. Il ne reste plus qu'à sertir avec une férule et une connectique (en PEEK) correspondant au diamètre externe du capillaire (ici 360 µm). Dans notre cas, la colonne maison (*homemade*) est connectée comme les colonnes commerciales au Q Exactive HF avec un capillaire et un tip externe (Figure 127).

#### 10.E ANALYSE LC-MS

##### 10.E.1. Configurations utilisées

Les diverses acquisitions ont été effectuées :

- Sur le Q Exactive HF : sur une U-3000 par une colonne de séparation par une colonne de séparation nanoEase (Waters) en phase BEH+C18 (avec des particules de 1.7µm de diamètre) de 75µm d.i. x 25 cm avec un débit de 220nl/min et un gradient d'élution de 30min ou 60min de 2 à 25% de tampon d'élution (H<sub>2</sub>O/ACN 20/80, AF 0.1%) ou par une colonne de séparation par une colonne Acclaim PepMap (ThermoFisher Scientific) en phase C18 (avec des particules de 2µm de diamètre) de 50µm d.i. x 50cm avec un débit de 115nl/min et un gradient d'élution de 30min ou 60min de 2 à 25% de tampon d'élution (H<sub>2</sub>O/ACN 20/80, AF 0.1%).
- Sur le Tims TOF Flex : sur une Nano-Elute par une colonne de séparation Aurora C18 (Ion Opticks) (75µm d.i. x 25cm) en directe injection avec un débit de 220nl/min et un gradient de 30min et 2 à 25% de tampon d'élution (H<sub>2</sub>O/ACN 20/80, AF 0.1%).
- Sur l'Eclipse Tribid : sur une U-3000 connectée à l'Eclipse Tribid avec une colonne nanoEase HSS C18 T3 (Waters) de 75µm d.i. et 25cm, avec un gradient de 37 minutes, passant de 2.5 à 44% de tampon B (H<sub>2</sub>O/ACN 20 :80 (v/v), AF 0.1%).
- Sur le Tims TOF Pro : sur une U-3000 par une colonne de séparation Aurora C18 (Ion Opticks) (75µm d.i x 25cm) en directe injection avec un débit de 220nl/min et un gradient de 30min de 5 à 30% de tampon d'élution (H<sub>2</sub>O/ACN 20/80, AF 0.1%).

##### 10.E.2. Paramètres MS

Les paramètres de la méthode MS sont les paramètres de base avec une gamme entre 100 et 1700 m/z en mode positif, en DDA. Sur le Tims TOF Pro, le Tims est activé et la méthode *low* décrite dans la partie résultats est utilisée avec le PASEF activé pour une analyse MS/MS.

## 10.F ANALYSE DES DONNEES

### 10.F.1. Analyse des données brutes

La plupart des analyses a été faite avec les logiciels disponibles dans les laboratoires de Sanofi ou de l'ESPCI. Les analyses de protéomique ont été faites sur le serveur Mascot dans un premier temps (données non montrées), puis avec le logiciel MaxQuant 1.6.17.0.

Les fichiers « .Raw » générés par l'analyse par le Q Exactive HF ou l'Eclipse Tribrid sont transformés en fichiers « .MGF » (*Mascot Generic Flatfiles*) grâce au logiciel Proteome Discoverer 2.4 (Thermo Scientific). Les fichiers « .D » générés par l'analyse par le Tims TOF Pro sont transformés en fichiers « .MGF » grâce à une méthode de post processing par le logiciel *Data Analysis* (Bruker). Ces fichiers « .MGF » sont utilisés pour la recherche dans les banques de données via le logiciel Mascot (Matrix Science).

Le serveur MASCOT 2.8.1 (Matrix Sciences) a été utilisé afin obtenir des résultats préliminaires nous indiquant les identifications en nombre de protéines et de peptides identifiés dans nos différentes acquisitions. La banque de données commune UniProt/SwissProt est utilisée comme banque de référence. La taxonomie choisie est adaptée à l'échantillon étudié : « *Homo Sapiens* » pour les échantillons de cellules humaines, « *Saccharomyces cerevisiae* » pour les échantillons de levure... L'interrogation de cette banque de données se fait avec les critères suivants : l'enzyme utilisée est la trypsine/P et jusqu'à 2 clivages manqués sont autorisés. D'autre part, les modifications fixes et variables sont modifiées en fonction du traitement subit par l'échantillon : la carbamidométhylation des cystéines est activée comme modification fixe lorsque les échantillons ont subi une réduction au DTT suivie d'une alkylation à l'iodoacétamide ou variable lorsque ce n'est pas le cas ; l'acétylation en N-terminal de la protéine ainsi que l'oxydation des méthionines sont marquées comme des modifications variables. La masse mono-isotopique est sélectionnée.

Pour les données générées par le Tims TOF, la tolérance de masse des peptides parents est de 15ppm et celle des peptides fragments est de 0.02Da. La présence d'un carbone 13 est supposée. La charge des peptides recherchée est de 2+, 3+ et 4+. L'instrument sélectionné est l'ESI-QUAD-TOF.

Pour les données générées par les appareils Orbitrap (le Q Exactive HF), la tolérance de masse des peptides précurseurs est de 5ppm et celle des peptides fragments est de 20mmu. La charge des peptides recherchée est de 2+ et 3+. L'instrument sélectionné est ESI-TRAP.

Enfin, une fois l'analyse terminée, un filtre à 1% de FDR est appliqué au nombre de protéines identifiées.

Pour les études de protéomique avec le logiciel MaxQuant 1.6.17.0 (*Max Plank Institute of Biochemistry, Department of Proteomics and Signal Transduction, Munich*), les paramètres de base sont conservés sauf mention contraire. L'enzyme de digestion est la trypsine/P avec 2 clivages manqués autorisés. Les modifications fixes et variables sont modifiées en fonction du traitement subit par l'échantillon : la carbamidométhylation des cystéines est activée comme modification fixe lorsque les échantillons ont subi une réduction au DTT suivie d'une alkylation à l'iodoacétamide ou variable lorsque ce n'est pas le cas ; l'acétylation en N-terminal de la protéine ainsi que l'oxydation des méthionines sont marquées comme des modifications variables. La quantification LFQ est inactivée. Pour les données brutes provenant d'une analyse par le Tims TOF Pro (Bruker), les paramètres d'instrument suivants sont modifiés :

« seuil d'intensité MS1 » et « seuil d'intensité MS2 » sont égaux à 0. Un filtre à 0.01 FDR est appliqué pour l'identification et l'option MBR est activée (*Match between runs*). Un peptide minimum est requis pour la quantification des protéines.

Les données de l'expérience analysée sur l'Eclipse Tribrid ont été traitées par le logiciel Proteome Discoverer 2.4 avec les mêmes paramètres que ceux utilisés pour le logiciel MaxQuant décrits précédemment.

### 10.F.2. Analyse des données traitées

#### (a) Les logiciels utilisés :

- Les diagrammes de Venn ont été faits grâce au logiciel Venn diagram Plotter (version 1.5.4798).
- Un script R rédigé au laboratoire incluant le package Kyte Doolittle a été utilisé pour les calculs des indices d'hydrophobicité et le tracé des graphiques en découlant.
- Les données de protéomique obtenues après une analyse par le logiciel MaxQuant sont analysées sur Excel pour éliminer les protéines contaminantes, *decoy* et celles identifiées seulement par leur site.
- Les logiciels GraphPad Prism (version 9.3.1) et Excel ont été utilisés pour le retraitement des données MaxQuant, l'analyse statistique et les graphiques.

#### (b) Tests statistiques :

Sur le logiciel GraphPad Prism :

- Si 2 conditions sont à comparer : un t-test non apparié a été effectué.
- Si 3 conditions sont à comparer : un test Anova à un facteur avec des comparaisons multiples a été effectué pour comparer les moyennes des conditions deux à deux.

## Chapitre 11 Figures annexes

### (a) Photographie du système d'évaporation

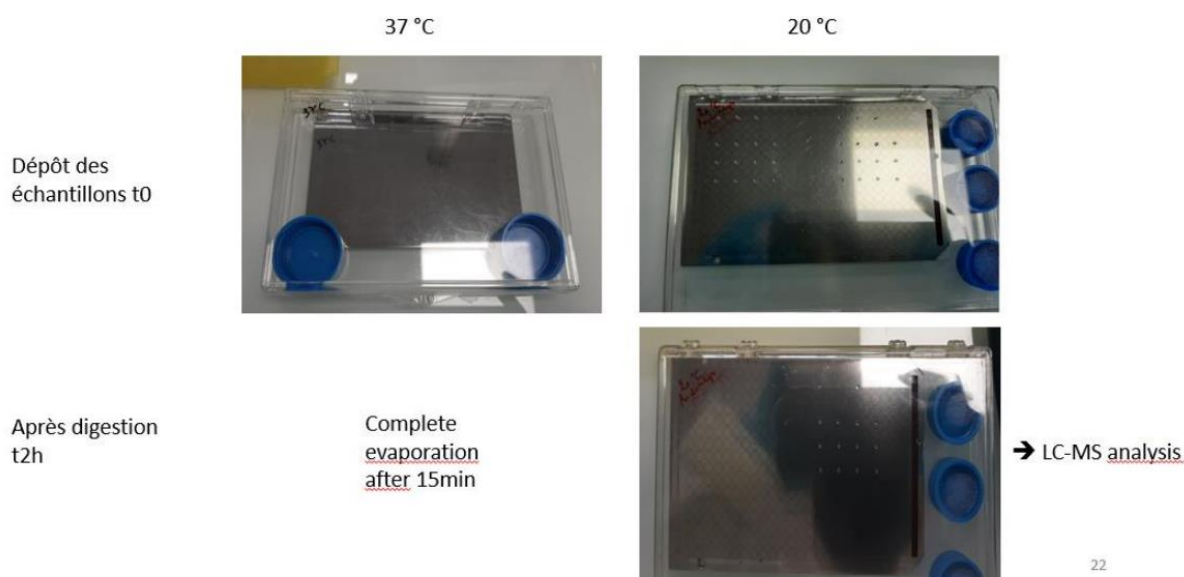


Figure 122 : photographie à t0 et à t2h de la plaque MALDI et *AnchorChip* utilisées pour la digestion des échantillons d'extraits protéiques de levure à 37°C et à 20°C avec un système *homemade* pour limiter l'évaporation consistant en une boîte fermée incluant des bouchons d'eau à 100°C.

### (b) YEAST DIGESTS ANALYSES SUR UN QEX HF

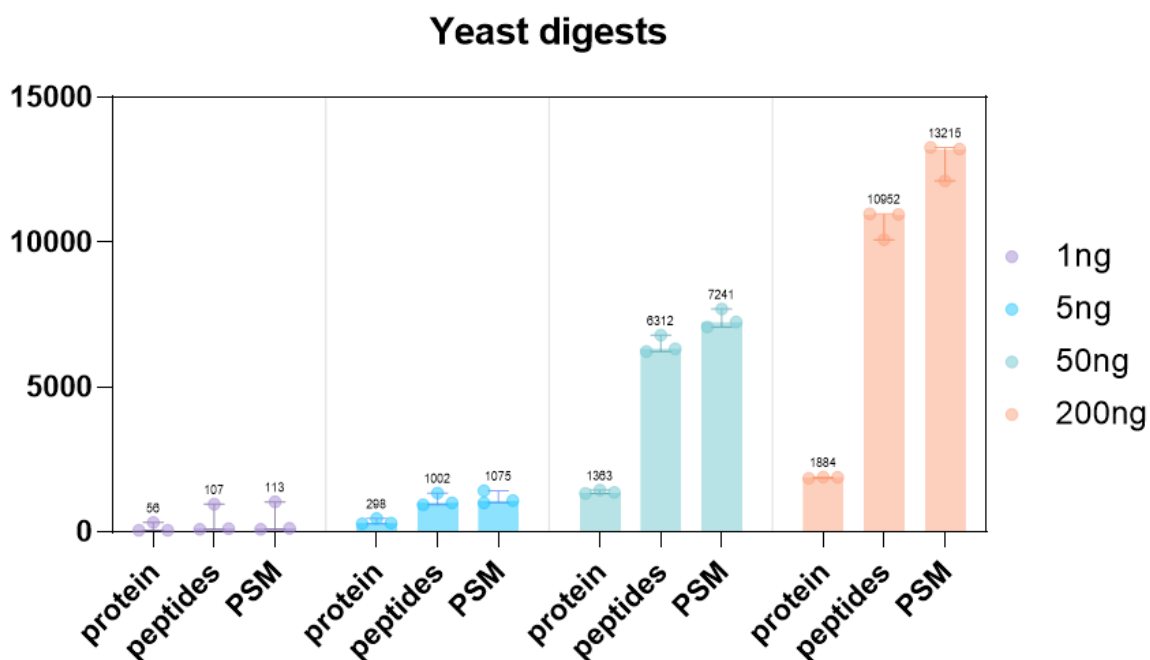


Figure 123 : Nombre de protéines, peptides et PSM identifiés en LC-MS/MS sur un Q Exactive HF à partir de différentes quantités d'un digestat de levure *Saccharomyces cerevisiae*. La moyenne des 3 réplicats est annotée au-dessus des barres d'écart type standard.

ANNEXES : Figures annexes

(c) SPECTRES THEORIQUES POUR ANALYSE DES POLYMERES

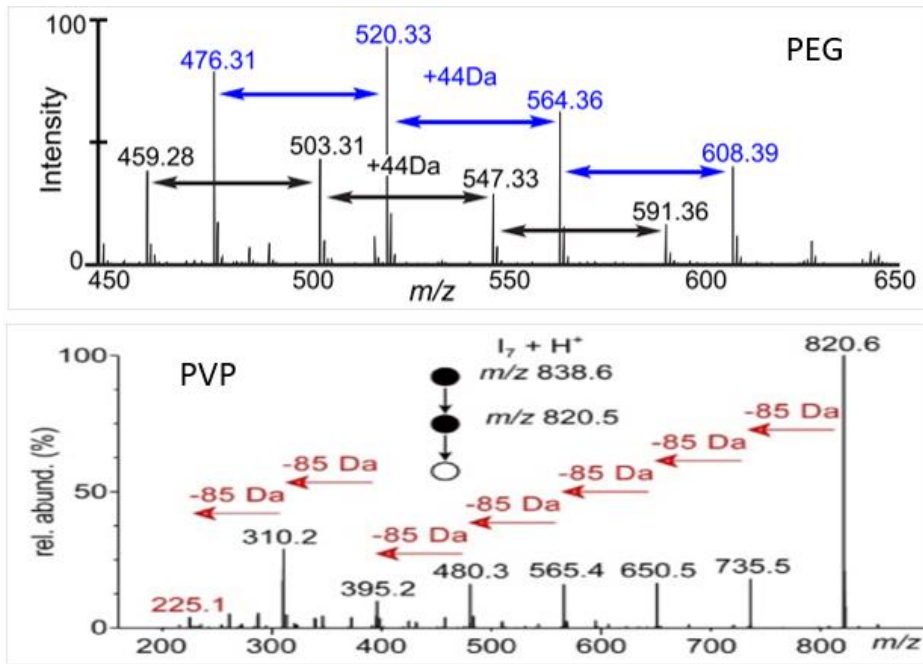
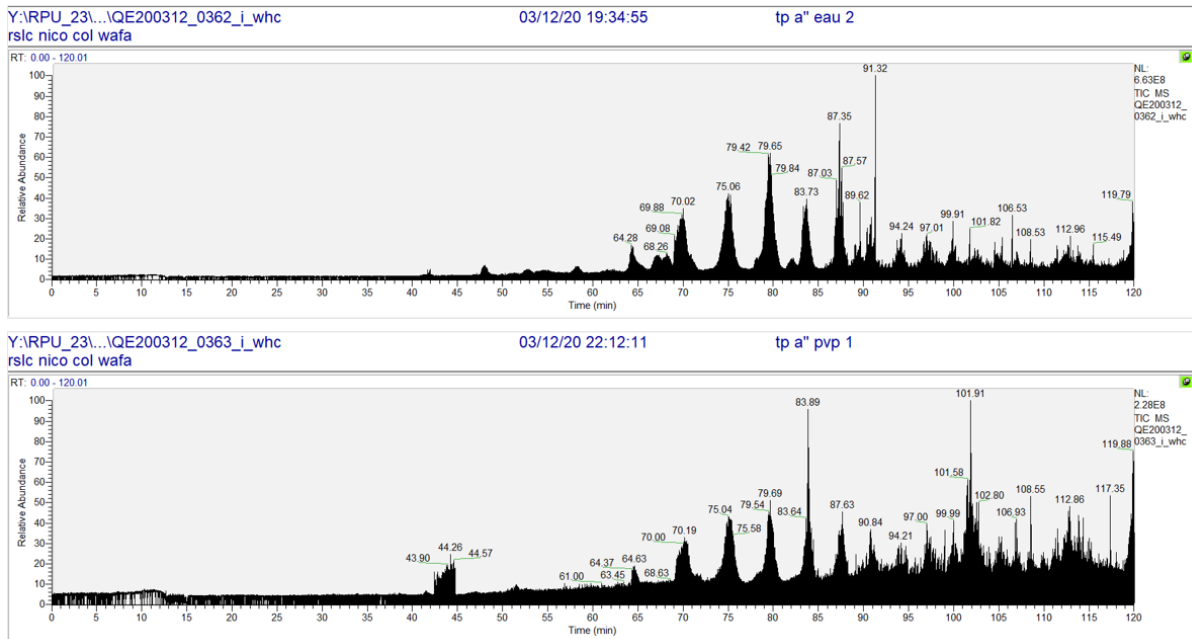


Figure 124 : spectres théoriques de polymères (en haut) PEG et (en bas) PVP

(d) TICS POUR ANALYSE DES POLYMERES



ANNEXES : Figures annexes

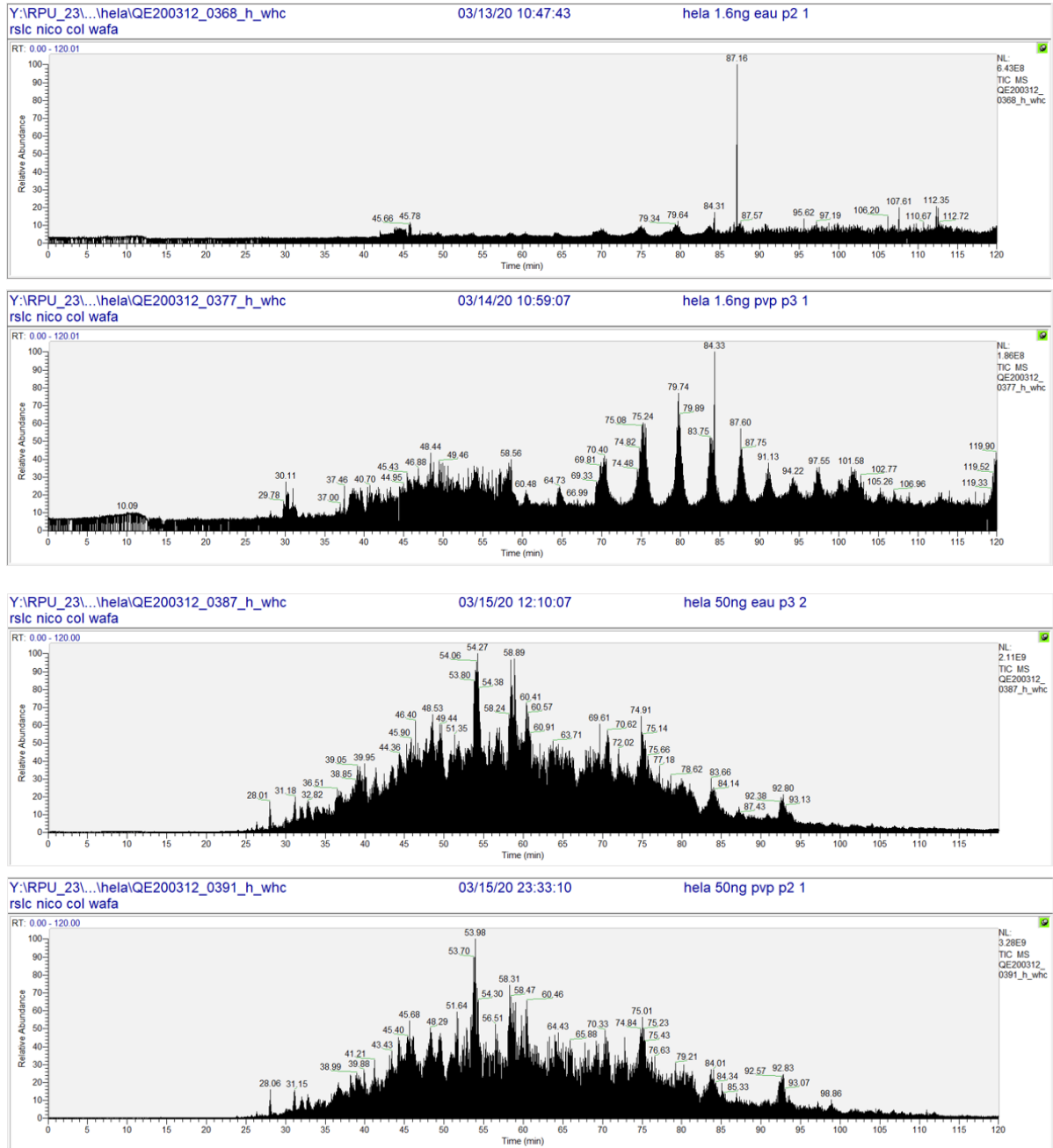


Figure 125 : TICs des échantillons, dans l'ordre de haut en bas, un échantillon blanc ayant traversé la puce non traitée, un échantillon blanc ayant traversé la puce traitée au PVP, un échantillon de digestat de HeLa 1.6ng ayant traversé la puce non traitée, un échantillon de digestat de HeLa 1.6ng ayant traversé la puce traitée au PVP, un échantillon de digestat de HeLa 50ng ayant traversé la puce non traitée, un échantillon de digestat de HeLa 50ng ayant traversé la puce traitée au PVP.

ANNEXES : Figures annexes

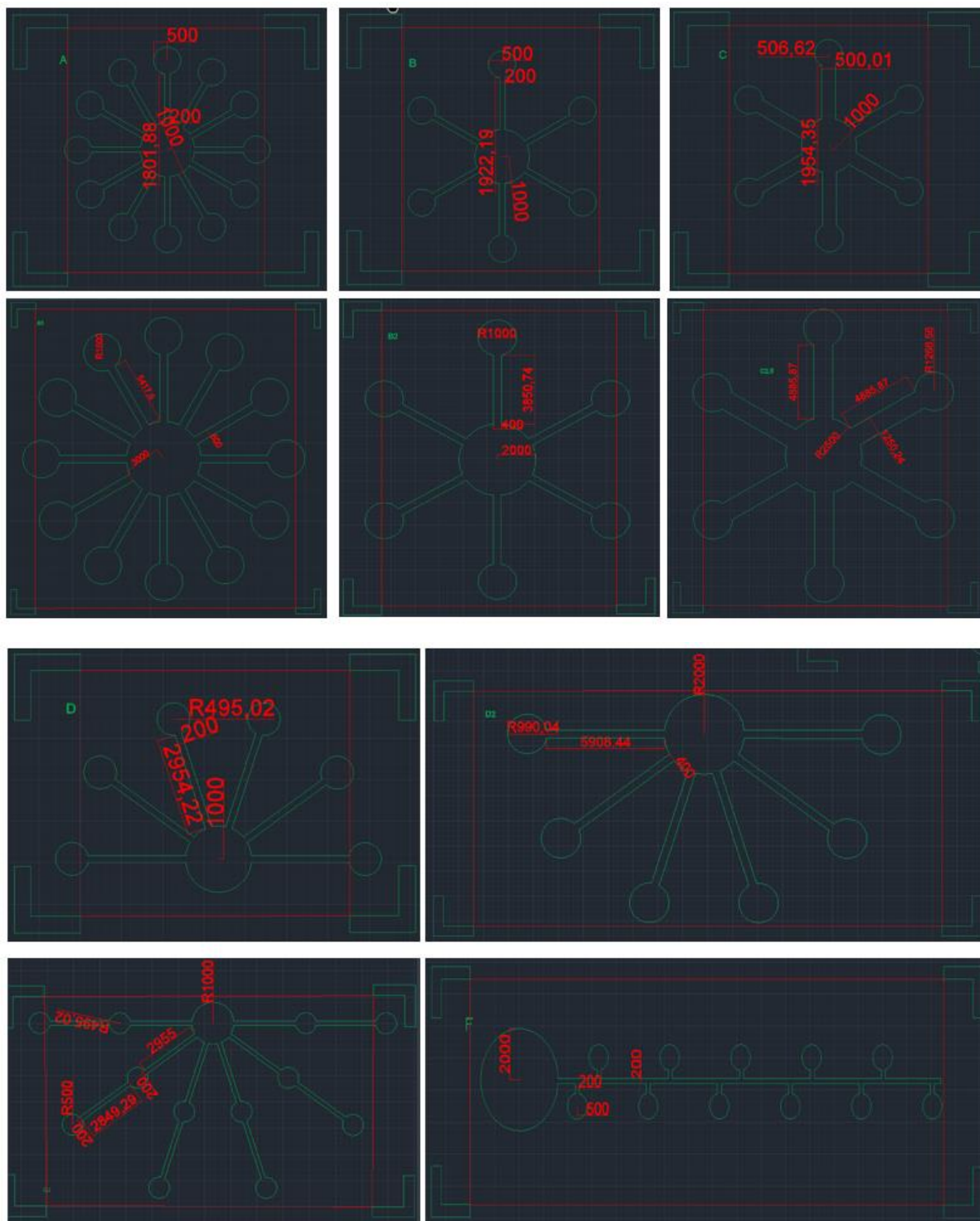


Figure 126 : Dimensions des différents designs de puce d’isolation.

ANNEXES : Figures annexes

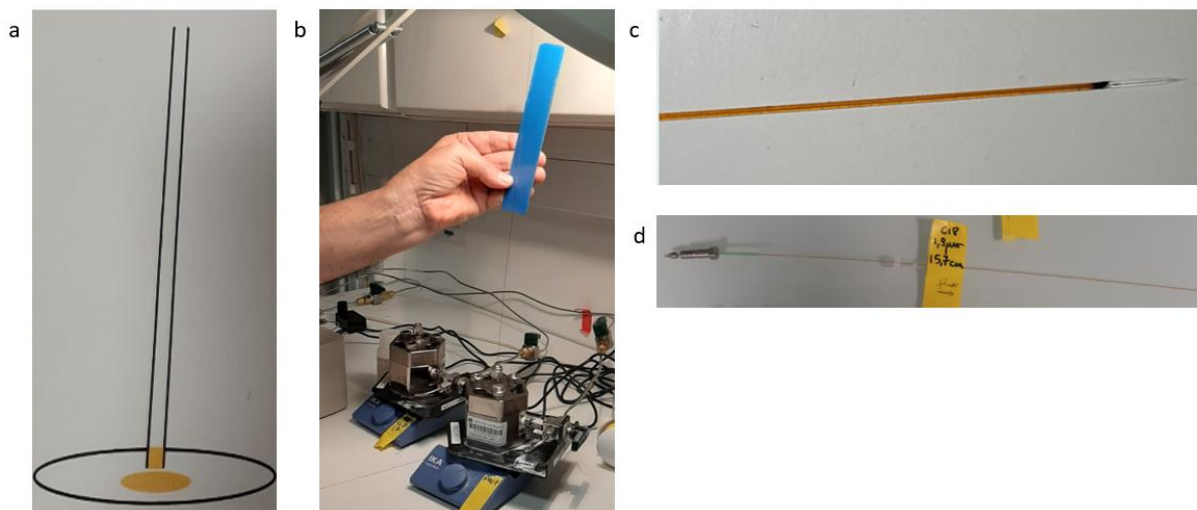


Figure 127 : Étapes de la fabrication d'une colonne en laboratoire. (a) installation du fritté (b) insertion de la phase dans le capillaire grâce à une pompe de gaz (c) création d'un embout pour la connexion avec la MS (d) sertissage de la colonne pour la connexion avec la LC.



Figure 128 : schéma de fonctionnement de l'Echo (*Acoustic Liquid Handler*). Le transducteur passe sous la plaque source et émet une énergie acoustique répétée qui produit un ménisque



## RÉSUMÉ

La cellule est une unité fondamentale de structure et de fonction en biologie. Les systèmes cellulaires tels que les cultures cellulaires, les tissus et les cancers se composent d'une variété de cellules aux propriétés moléculaires et fonctionnelles distinctes. La caractérisation de ces différences cellulaires est essentielle pour comprendre la physiologie ou les pathologies. Pour relever de nombreux défis médicaux et biologiques, l'analyse protéomique du protéome d'une cellule unique est devenue essentielle car elle fournit des informations sur les changements dynamiques du protéome ou l'hétérogénéité cellulaire. Cependant, le protocole classique utilisé en protéomique n'est pas adapté à une cellule unique ou même à des quantités minimales de matériel biologique. Il est donc essentiel de le miniaturiser pour l'adapter à une quantité minimale de matériel. De l'isolement des cellules à l'analyse LC-MS/MS, chaque étape du processus de préparation des échantillons nécessite un développement minutieux pour augmenter la sensibilité de l'analyse et réduire les pertes d'échantillons avant l'arrivée de l'échantillon devant le détecteur du spectromètre de masse. L'étude des surfaces des supports en contact avec les échantillons est souvent négligée et pourtant très importante. Un des phénomènes observés avec les protéines représentant un réel défi en protéomique unicellulaire est l'adsorption non spécifique des molécules d'intérêt sur les surfaces. Ainsi, l'adsorption des protéines et des peptides limite le rendement de récupération et est délétère pour identifier un maximum de peptides. Cet effet est d'autant plus visible que la quantité d'échantillons est faible. Dans le présent travail, nous avons étudié la nature de la surface afin d'optimiser le rendement de préparation des échantillons. Nous avons également étudié plusieurs étapes du protocole afin de les améliorer. Nos résultats ont confirmé qu'un protocole minimaliste ainsi que la réduction et le traitement des surfaces de contact sont des paramètres cruciaux pour améliorer la robustesse de la récupération et de l'identification des protéines et la sensibilité de la détection des peptides lors des analyses en LC-MS/MS de faibles quantités d'extraits de protéines ou un faible nombre de cellules humaines.

## MOTS CLÉS

Spectrométrie de masse, miniaturisation, protéomique, échantillons de trace, préparation d'échantillons

## ABSTRACT

The cell is a fundamental unit of structure and function in biology. Cellular systems such as cell cultures, tissues and cancers are composed of a variety of cells with distinct molecular and functional properties. Characterization of these cellular differences is essential for understanding normal physiology or pathology. To address many medical and biological challenges, proteomic analysis of the single cell proteome has become essential as it provides information on dynamic changes in the proteome or cellular heterogeneity. However, the classical workflow used in proteomics is not adapted to a single cell or even to minimal quantities of biological material. It is therefore essential to miniaturize it to adapt it to a minimal amount of material. From cell isolation to LC-MS/MS analysis, each step of the sample preparation process requires careful development to increase the sensitivity of the analysis and reduce sample loss before the sample arrives at the mass spectrometer detector. The study of the surfaces of the containers in contact with the samples is often neglected and yet very important. One of the phenomena observed with proteins and representing a real challenge in single-cell proteomics is the non-specific adsorption of molecules of interest on surfaces. Thus, the adsorption of proteins and peptides limits the recovery yield and is deleterious to identify a maximum of peptides. This effect is even more visible when the amount of samples is small. In the present work, we studied the nature of the surface to optimize the sample preparation yield. We also studied several steps of the protocol in order to improve them. Our results confirmed that a minimalist protocol and the reduction and treatment of contact surfaces are crucial parameters to improve the robustness of protein recovery and identification and the sensitivity of peptide detection when working in LC-MS/MS with low amounts of protein extracts or low numbers of human cells.

## KEYWORDS

Mass spectrometry, miniaturization, single cell proteomics, trace samples, sample preparation