N°d'ordre NNT : xxx

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
L'Université Claude Bernard Lyon 1

Ecole Doctorale N° 206
Chimie

Spécialité de doctorat : Chimie
Discipline : Sciences Analytiques

Soutenue publiquement le 29/11/2017, par :
Lény GARCIA

Développement d'une méthodologie couplant la photodissociation laser et la spectrométrie de masse haute résolution pour l'identification de nouveaux biomarqueurs.

Devant le jury composé de :

ENJALBAL Christine, Professeure des Universités, Université Montpellier 2	Rapporteur
ROLANDO Christian, Directeur de recherche CNRS, Université de Lille 1	Rapporteur
CARAPITO Christine, Chargée de Recherche CNRS, Université de Strasbourg	Examinatrice
DUGOURD Philippe, Directeur de recherche CNRS, Université Lyon 1	Examinateur
LEMOINE Jérôme, Professeur des Universités, Université Lyon 1	Directeur de thèse
GIROD Marion, Chargée de Recherche CNRS, Université Lyon 1	Co-Directrice de thèse

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	**M. le Professeur Frédéric FLEURY**
Président du Conseil Académique	M. le Professeur Hamda BEN HADID
Vice-président du Conseil d'Administration	M. le Professeur Didier REVEL
Vice-président du Conseil Formation et Vie Universitaire	M. le Professeur Philippe CHEVALIER
Vice-président de la Commission Recherche	M. Fabrice VALLÉE
Directrice Générale des Services	Mme Dominique MARCHAND

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur G.RODE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. X. PERROT
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : Mme la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. THEVENARD
Département Chimie Biochimie	Directeur : Mme C. FELIX
Département GEP	Directeur : M. Hassan HAMMOURI
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur G. TOMANOV
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : M. le Professeur J-C PLENET

UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. Y.VANPOULLE
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. le Professeur E.PERRIN
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. le Professeur C. VITON
Ecole Supérieure du Professorat et de l'Education	Directeur : M. le Professeur A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Directeur : M. N. LEBOISNE

« Toute création véritable implique une certaine surdité à l'appel d'autres valeurs, pouvant aller jusqu'à leur refus, sinon même à leur négation. Car on ne peut à la fois se fondre dans la jouissance de l'autre, s'identifier à lui et se maintenir différent. Pleinement réussie la communication intégrale avec l'autre condamne, à plus ou moins brève échéance, l'originalité de sa et de ma création. Les grandes époques créatrices furent celles où la communication était devenue suffisante pour que des partenaires éloignés se stimulent, sans être cependant assez fréquente et rapide pour que les obstacles, indispensables entre les individus comme entre les groupes, s'amenuisent au point que des échanges trop faciles égalisent et confondent la diversité »

« Une formule à laquelle nous avons fait un aussi grand sort que "l'enfer, c'est les autres" ne constitue pas une proposition philosophique, mais un témoignage ethnographique sur une civilisation. Car on nous a habitués dès l'enfance à craindre l'impureté du dehors. Quand ils proclament au contraire que "l'enfer, c'est nous-mêmes", les peuples sauvages nous donnent une leçon de modestie qu'on voudrait croire que nous sommes encore capables d'entendre. En ce siècle où l'homme s'acharne à détruire d'innombrables formes vivantes, après tant de sociétés dont la richesse et la diversité constituaient de temps immémorial le plus clair de son patrimoine, jamais sans doute il n'a été plus nécessaire de dire, comme font les mythes, qu'un humanisme bien ordonné ne commence pas par soi-même, mais place le monde avant la vie, la vie avant l'homme, le respect des autres êtres avant l'amour-propre ; et que même un séjour d'un ou deux millions d'années sur cette terre, puisque de toute façon il connaîtra un terme, ne saurait servir d'excuse à une espèce quelconque, fût-ce la nôtre, de se l'approprier comme une chose et s'y conduire sans pudeur ni discrétion. Et pourtant, aujourd'hui, l'être humain est posé comme valeur ultime à laquelle tout, dans l'univers, doit lui être subordonné ou sacrifié. Les rappels à la modestie, lancés sous des formes différentes, par Copernic, par Darwin et par Freud, les opérations de décentrement auxquelles ont procédé, chacun à sa manière, ces trois savants, sont restés inutiles. À la fin du XIXe siècle, notamment au travers des idéologies du progrès, l'homme revendique désormais une royauté sans partage et sans limites. »

Claude Lévi-Strauss par Emmanuel Terray

Remerciements

Je tiens tout d'abord à remercier Monsieur Jérôme Lemoine pour m'avoir donné la chance de réaliser cette thèse au sein de son équipe ainsi que pour la liberté de travail qui m'a été accordée. A travers celle-ci, j'ai pu m'enrichir sur bien d'autres domaines que la spectrométrie de masse, me permettant, aujourd'hui, une solidité et une ouverture que je n'aurais pu acquérir autrement.

Je tiens également à remercier considérablement Madame GIROD Marion pour tout son accompagnement, sa disponibilité et son travail pour me soutenir et mener à bien ce projet durant ces 3 années. A travers ces échanges j'ai pu grandir scientifiquement et ouvrir mon regard à la pluralité des manières d'aborder un point de vue. En résulte aujourd'hui, beaucoup d'incertitudes sur toutes nos certitudes, ce qui, je l'espère, me permettra un regard et un geste scientifique à la hauteur de leur définition.

Je tiens à remercier pareillement toutes les personnes avec qui j'ai pu mener à bien ces projets. Tout d'abord je remercie sérieusement Madame Christine CARAPITO et Madame Magali ROMPAIS pour leur collaboration. Sans la construction de cette bibliothèque de spectres, cette thèse n'aurait, tout simplement, pas pu être. Pareillement, je remercie Monsieur Luke MACALEESE pour tout le travail en commun effectué en rapport à la C-trap-LID et pour tous les échanges scientifiques riches en résultant. Je termine par remercier Madame Joëlle SAULNIER et Madame Diane CALA pour leur gentillesse, leur aide matérielle et manuelle concernant les dosages de Bradford et la préparation des échantillons E.Coli.

Je remercie les membres du Jury d'avoir étudié ce manuscrit : Madame Christine ENJALBAL et Monsieur Christian ROLANDO pour avoir accepté d'être rapporteurs de ce travail de thèse, Madame Christine CARAPITO et Monsieur Philippe DUGOURD pour avoir accepté d'être présents en tant qu'examinateurs.

Je tiens, d'une manière moins informelle mais tout aussi sérieuse, à remercier toute l'équipe permanente du laboratoire. Avec, tout d'abord, une mention spéciale à Monsieur Romain CARRIERE pour son café de haute volée, ses discours politiques toujours aux petits oignons et pour sa passion « ultra no limit » à l'égard du football et du tennis. Je tiens également à remercier fortement Madame Catherine FONBONNE pour sa gentillesse incroyable, sûrement au-delà du réel, et pour sa gérance rigoureuse du laboratoire. Il en va de même pour Madame Sophie ACYRIEX tout en espérant que le projet « lipides » pourra se développer comme elle le désire. Je termine par remercier Monsieur Arnaud SALVADOR, pour tous ses conseils relatifs au cinéma et au jogging. Je lui souhaite bon courage dans sa quête de réalisation des meilleurs cours possibles pour ses étudiants.

Je remercie d'une manière moins informelle mais encore moins sérieuse tous les doctorants avec qui j'ai pu échanger au cours de ces trois ans. Tout d'abord au laboratoire : un grand merci à Antonin ! A travers nos discussions de barbus tachées de bières et de voyages sudistes, nous

avons pu entrapercevoir une nouvelle métaphysique qui se vit essentiellement entre 19h et 04h du matin à 1 g et pas moins ! Un grand merci également à Nour pour sa folie quotidienne, son café trop léger et nos discussions essentiellement axées sur rien, et puis sur tout aussi, mais seulement quand on est fatigué, c'est-à-dire souvent. Et puis un merci à Blandine, pour sa gentillesse et la transmission de son savoir durant ces deux années. J'espère qu'elle ne s'attardera pas trop en Angleterre et qu'elle concrétisera enfin son rêve d'éleveuse de chèvres.

Je remercie tous les doctorants d'ISATOPE que j'ai croisé ou avec qui j'ai travaillé. Principalement un grand merci à Marie, Léa, Marion, Maria, Elodie, Fériel, Pierrick, Emmanuelle, Gabriel, Juan et Selim pour s'être impliqués dans cette association et l'avoir faite vivre.

Je remercie également le personnel d'ANAQUANT et en particulier Chloé pour son calme légendaire et sa zénitude face à l'adversité spectrométrique.

Je termine par remercier toute ma famille, principalement mes parents et mes grands-parents à qui je dédie ce travail car ils sont la raison de ce que je suis et de ce que je tends à être. Je les remercie de m'avoir offert la chance de pouvoir devenir librement.

Un grand merci à tous les potes de Montpellier, d'Ardèche, de Lyon, de Barcelone, de Paris, de Bruxelles, de la Napa Valley, de Bretagne, de Blois et d'Orléans : je vous remercie pour tous les délires passés et à venir.

Viva Carnival of Sins !!

Table des matières

Remerciements 5
Abréviations 11
Liste des figures 13
Liste des tableaux 17

Introduction générale 19

Chapitre I : Identification et quantification de biomarqueurs protéines par spectrométrie de masse – état de l'art 22
Introduction 23
I) Les biomarqueurs 24
I.1) Définition 24
I.2) La découverte de nouveaux biomarqueurs 24
I.3) La validation de nouveaux biomarqueurs 26
I.4) Les protéines kinases humaines 27
I.5) Difficultés actuelles liées à l'identification et à la validation de nouveaux biomarqueurs 31
II) Principes de l'analyse de protéines en spectrométrie de masse 33
II.1) Préparation d'échantillons lors d'une analyse bottom-up 33
II.2) La séparation chromatographique en phase liquide (HPLC) 35
II.3) L'analyse de peptides par spectrométrie de masse 36
III) Différentes méthodes d'activation des ions en phase gazeuse pour l'analyse de peptides et de protéines en spectrométrie de masse 47
III.1) Nomenclature des ions fragments résultant de la fragmentation des peptides par spectrométrie de masse 47
III.2) La dissociation induite par collision (CID) 49
III.3) L'activation électronique 50
III.4) La photoactivation 54
Conclusion 63

Chapitre II : Développement de la méthodologie DIA-LID à 473 nm au sein d'un spectromètre de masse haute résolution pour l'identification et la quantification de protéines kinases humaines 64
Introduction 65

I) Matériels et Méthodes 66
I.1) Echantillons 66
I.2) Choix du chromophore 67
I.3) Préparation d'échantillons 68
I.4) Partie chromatographie 70
I.5) Spectrométrie de masse 72
I.6) Logiciels 85
II) Présentation générale de la dérivation et de la photo-fragmentation LID à 473 nm des peptides à cystéine dérivés 86
II.1) Dérivation des peptides à cystéine en solution 86
II.2) Etude de stabilité des peptides à cystéine dérivés 89
II.3) Photo-fragmentation laser à 473 nm des peptides à cystéine dérivés avec le DABCYL-C2 maléimide 91
II.4) Temps d'activation laser des peptides à cystéine dérivés avec le DABCYL-C2 maléimide à 473 nm 93
II.5) Taux de photo-fragmentation des peptides à cystéine dérivés avec le DABCYL-C2 maléimide à 473 nm 94
II.6) Comparaison entre la méthode d'activation LID à 473 nm et la fragmentation HCD sur des peptides à cystéine dérivés avec le DABCYL-C2 maléimide 96
III) Construction d'une bibliothèque de spectres LID à 473 nm spécifiques de peptides à cystéine dérivés rapporteur de protéines kinases humaines 98
III.1) Sélection des peptides à cystéine constitutifs de la bibliothèque de spectres 98
III.2) Construction de la bibliothèque de spectres LID à 473 nm 98
III.3) Contrôle de l'information structurale contenue dans la bibliothèque de spectres LID à 473 nm 102
III.4) Comparaison entre la DDA-LID et la DIA-LID à 473 nm 104
IV) Preuve de concept de l'utilisation de la DIA-LID à 473 nm et d'une bibliothèque de spectres pour identifier et quantifier des peptides à cystéine dérivés rapporteur de protéines humaines au sein de matrices biologiques 105
IV.1) Preuve de concept de l'utilisation de la DIA-LID à 473 nm et d'une bibliothèque de spectres pour identifier des peptides à cystéine dérivés au sein de matrices biologiques 105
IV.2) Preuve de concept de l'utilisation de la DIA-LID à 473 nm et d'une bibliothèque de spectres pour quantifier de manière répétable des peptides à cystéine dérivés au sein de matrices biologiques 115
V) Utilisation de la méthodologie DIA-LID à 473 nm et d'une bibliothèque de spectres pour identifier et quantifier des peptides à cystéine dérivés endogènes issus de protéines kinases humaines au sein d'un extrait cellulaire humain 120
VI) Perspective 124

VI.1) Dérivation des peptides à cystéine après l'étape de digestion à la trypsine 124
VI.2) Dérivation des peptides à cystéine à l'aide d'un nouveau chromophore 125
Conclusion 135

Chapitre III : Etude du modèle de fragmentation des peptides à cystéine dérivés avec le DABCYL C2-maléimide par LID à 473 nm 136
Introduction 137
III.I) Précision quant aux résultats présentés dans ce chapitre 138
III.II) Types de fragments générés lors de la photo-fragmentation laser de peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm 139
III.III) Nature de l'acide aminé fragmenté lors de la photo-fragmentation laser de peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm 142
III.IV) Etude du processus de redistribution d'énergie au niveau de la séquence peptidique lors de la photo-fragmentation laser de peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm . 145
III.V) Influence de la position de la cystéine lors de la photo-fragmentation laser de peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm 149
Conclusion 154

Chapitre IV : Mise en place d'une nouvelle méthodologie intitulée C-trap-LID 157
Introduction 158
IV.I) Matériels et Méthodes 159
IV.I.1) Sélection des échantillons 159
IV.I.2) Chromophore utilisé 159
IV.I.3) Préparation d'échantillons 159
IV.I.4) Couplage LC/MS/MS 159
IV.I.5) Logiciels 160
IV.II) Constatation de la possibilité d'effectuer la photo-fragmentation laser au sein de la C-Trap (C-trap-LID) 160
IV.III) Apport de la méthodologie C-trap-LID lors de l'analyse de peptides à cystéine dérivés 162
IV.IV) Mise en place de la C-trap-LID 164
IV.IV.1) Installation d'un obturateur 164
IV.IV.2) Réglage de l'obturateur 166
IV.V) Mise en application du concept C-Trap-LID à 473 nm 169
IV.V.1) Calibration avec le DABCYL-C2 maléimide 169
IV.V.2) Photo-fragmentation laser à 473 nm de peptides à cystéine dérivés avec le DABCYL C2-maléimide en utilisant la méthodologie C-Trap-LID 171

Conclusion 180

Conclusion générale 181

Bibliographie 184

Annexes 202

Abréviations

µL : Micro litre

AA : Acétate d'ammonium

AIF : All Ion Fragmentation

ACN : Acétonitrile

AGC : Automatic Gain Control

ATP : Adenosine Tri-Phosphate

CE : Collision Energy

CID : Collision Induced Dissociation

Cps : Coups par seconde

CV : Coefficient de variation

DDA : Data-Dependent Acquisition

DIA : Data-Independent Acquisition

E.Coli : Escherichia Coli

ECD : Electron Capture Dissociation

ETD : Electron Transfer Dissociation

(a-EPD) : Activated-Electron PhotoDetachment

ELISA : Enzyme-linked ImmunoSorbent Assay

ESI : Electrospray Ionization

eV : Electron volt

FA : Acide formique

FMS : Full MS

FTICR : Fourier-Transform Ion Cyclotron Resonance

HCD : Higher-energy Collisional Dissociation

HPLC : High Pressure Liquid Chromatography

IRMPD : Infrared Multiple Photon Dissociation

IT : Injection Time

LC/MS/MS : Couplage chromatographie en phase liquide et spectrométrie de masse en tandem

LID : Laser Induced Dissociation

LIT : Linear Ion Trap

LOQ : Limit Of Quantification

***m/z* :** rapport masse sur charge

MeOH : Méthanol

min : Minutes

mL : Millilitres

NCE : Normalized Collision Energy

nm : Nanomètres

PRM : Parallel Reaction Monitoring

PTM : Post-Translational Modification

R^2 : Coefficient de corrélation

RT : Retention Time

SPE : Solid Phase Extraction

SRM/MRM : Selected/Multiple Reaction Monitoring

SWATH : Sequential Window Acquisition of all Theoretical Spectra

TOF : Time Of Flight

UVPD : UltraViolet PhotoDissociation

UV : UltraViolet

V.rep : Volume de reprise

VUV : Ultraviolet du vide

TIC : Total Ion Current

XIC : eXtracted-Ion Chromatogram

λ : Longueur d'onde

Liste des figures

Figure 1 : Phosphorylation d'un acide aminé tyrosine par le biais d'une tyrosine kinase et de l'ATP[32] 28
Figure 2 : Représentation en 3 dimensions du domaine catalytique des protéines kinases[42] 29
Figure 3 : Evolution du coût financier (en millions de dollars) pour identifier et valider un biomarqueur par un laboratoire pharmaceutique à travers 4 époques différentes[65] 32
Figure 4 : Liaison peptidique entre 2 acides aminés 34
Figure 5 : Schéma récapitulatif de l'approche bottom-up pour l'analyse de protéines en LC/MS/MS 36
Figure 6 : Fonctionnement du mode SRM sur un triple quadrupole[83] 38
Figure 7 : Comparaison du fonctionnement du mode SRM et du mode PRM[85] 39
Figure 8 : Fonctionnement du mode DDA : 1 scan FMS est acquis suivi d'une sélection individuelle successive de N ions précurseurs en vue d'une étape de fragmentation par CID/HCD 41
Figure 9 : Procédé utilisé pour identifier des peptides en mode DDA 41
Figure 10 : Fonctionnement du mode DIA : 1 scan FMS est acquis suivi d'une fragmentation de la totalité de la plage de m/z par l'intermédiaire de larges fenêtres successives de plusieurs Th 44
Figure 11 : Complexité des spectres de (co)-fragmentation obtenu en DIA 46
Figure 12 : Nomenclature des ions fragments issus de peptides[69] 48
Figure 13 : Activation par CID des ions en phase gazeuse[147] 49
Figure 14 : Principe de l'activation ETD[158] 51
Figure 15 : Ions principalement produits lors de l'activation de peptides par ETD/ECD[157] 52
Figure 16 : Spectres de fragmentation du peptide NEREEHAMR dichargé en ECD et trichargé en ETD[155]. 53
Figure 17 : Spectres de fragmentation du peptide DLYANTVLSGGTTMYPGIADR sous forme dichargé après activation en UPVD à 193 nm (a) et en CID (b)[184] 56
Figure 18 : Transition d'un état stable E_0 à un état excité E_1 57
Figure 19 : Spectre de photo-fragmentation après irradiation à 351 nm d'un peptide non dérivé avec le chromophore nitroazobenzène (a) et du peptide KPLLIIAEDVEGEY dérivé avec ce chromophore sur l'acide aminé Y (b)[202] 59
Figure 20 : Spectre de photo-fragmentation après irradiation à 351 nm du peptide à cystéine YCASGGELGFPY dérivé avec le chromophore Alexa Fluor sur l'acide aminé cystéine[203] 60
Figure 21 : Absence d'interférence dans le transition SRM suivi du peptide rapporteur de la C4b binding protéine en photo-SRM (a) alors que présente en SRM classique (c)[14] 61
Figure 22 : Evolution du taux de photo-fragmentation du peptide à cystéine IVGGWECEK (2+) après irradiation à 473 nm en fonction du temps d'activation[9] 61
Figure 23 : Structure chimique du DABCYL C2-maléimide 67
Figure 24 : Spectre d'absorption du DABCYL C2-maléimide en solution 68
Figure 25 : Schéma détaillé d'un Qexactive[205] 72
Figure 26 : Principe de fonctionnement d'une source ESI[206] 73
Figure 27 : Schéma d'un quadripôle avec les potentiels appliqués sur les 4 barreaux[210] 75
Figure 28 : Zones stables dans un quadripole d'un ion donné suivant l'axe x ou y (a et b). Les quatre zones de stabilités sont notées de A à D (c). La zone A est celle utilisée dans un quadripole (d). La partie hachurée correspond à un potentiel U positif[210] 77
Figure 29 : Diagramme de stabilité en fonction de U et V donné pour un ion[204] 78
Figure 30 : Schéma d'une orbitrap[215] 79
Figure 31 : Modification de la géométrie de la cellule entre une orbitrap standard et une orbitrap High-field[215] 81
Figure 32 : Schéma d'un Qexactive couplé avec un laser irradiant à 473 nm dans la cellule HCD 82

Figure 33 : Dérivation de la fonction thiol à l'aide d'un groupement maléimide 86
Figure 34 : Peptide à cystéine dérivé avec le chromophore DABCYL C2-maléimide sur l'acide aminé cystéine..... 86
Figure 35 : XIC du peptide à cystéine NILWSAENECFK non dérivé dichargé dans l'échantillon non traité (B) (a) et traité avec le DABCYL C2-maléimide (A) (c), dérivé trichargé dans l'échantillon traité avec le DABCYL C2-maléimide (A) (b) et non traité (B) (d) 87
Figure 36 : XIC du peptide à cystéine ACDALYPPGPPSNPK non dérivé dimérisé dichargé dans l'échantillon non traité (B) (a) et traité avec le DABCYL C2-maléimide (A) (c), dérivé trichargé dans l'échantillon traité avec le DABCYL C2-maléimide (A) (b) et non traité (B) (d)..... 88
Figure 37 : Etude de stabilité de l'ocytocine dérivée dans 3 milieux de conservations différents durant 3 jours..... 90
Figure 38 : Etude de stabilité de la vasopressine dérivée dans 3 milieux de conservations différents durant 3 jours 90
Figure 39 : Comportement de l'ion précurseur monochargé (m/z 392.17) du DABCYL C2-maléimide en PRM-LID à 473 nm lorsque le laser est éteint (a) ou allumé (temps d'activation 3 ms) (b) 91
Figure 40 : Structure du DABCYL C-2 maléimide avec localisation du site de fragmentation générant l'ion rapporteur m/z 252,11..... 92
Figure 41 : Spectre de photo-fragmentation de l'ion dichargé m/z 796,39 du peptide à cystéine dérivé YCSVTTQADGR obtenu en PRM LID à 473 nm avec un temps d'activation de 25 ms..... 93
Figure 42 : Spectre de photo-fragmentation de l'ion trichargé m/z 513,59 du peptide à cystéine dérivé LCDSGELVAIK en PRM LID à 473 nm à 4 temps d'activation différents 25 ms (a) 50 ms (b) 75 ms (c) et 100 ms (d)..... 94
Figure 43 : Spectres de fragmentation des ions dichargés m/z 879,90 et 796,39 des peptides à cystéine dérivés CTEALYAPAEGDK et YCSVTTQADGR en PRM-HCD (a,b) et en PRM-LID à 473 nm (c,d) 97
Figure 44 : XIC de l'ion quadrichargés du peptide à cystéine EGNIQYLFLEYCSGGELFDR sous forme non dérivée (a) et dérivée (b) lors de l'analyse DDA-LID à 473 nm..... 100
Figure 45 : Interface sous le logiciel Skyline de la bilbiothèque de spectres LID à 473 nm contenant les 354 peptides à cystéine dérivés rapporteurs de 354 protéines kinases humaines 101
Figure 46 : Couverture de séquence moyenne obtenue via les ions b et y pour les 354 peptides à cystéine dérivés de la bibliothèque de spectres LID à 473 nm 102
Figure 47 : Proportion moyenne, sur un spectre de photo-fragmentation, de l'intensité de l'ion rapporteur m/z 252,11 et des ions fragments b et y pour 40 peptides à cystéine dérivés de la bibliothèque de spectres LID à 473 nm..... 103
Figure 48 : Comparaison des spectres de photo-fragmentation du peptide à cystéine dérivé VMCDSNDLPK en DDA-LID à 473 nm (a) et en DIA-LID à 473 nm (b)..... 104
Figure 49 : XIC de 6 peptides à cystéine dérivés détectés en mode DIA-LID à 473 nm dans les échantillons d'E.Coli dopés avec les 354 peptides à cystéine dérivés..... 107
Figure 50 : XIC de 3 peptides à cystéine dérivés détectés en mode DIA-LID à 473 nm dans l'échantillon d'E.Coli dopé avec les 354 peptides à cystéine dérivés présentant une isomérie de conformation..... 108
Figure 51 : % de peptides à cystéine dérivés identifiés avec 6, 5, 4 ou 3 fragments 109
Figure 52 : XIC de 2 peptides à cystéine dérivés détectés en mode DIA-LID à 473 nm avec un dopt score inférieur à 0,85 111
Figure 53 : Spectre de co-fragmentation des peptides à cystéine dérivés NILWSAENECFK et NILVGENTLCK en mode DIA-LID à 473 nm (a et c) et en DIA-HCD (b et d). Les * correspondent à des ions précurseurs ne contenant pas de cystéine dérivée. 113

Figure 54 : XIC des 2 peptides à cystéine dérivés LYLAENYCFDSIPK et NILWSAENECFK en mode DIA-HCD (a et c) et DIA-LID à 473 nm (b et d) 114
Figure 55 : XIC des 3 réplicas du peptide à cystéine dérivé APEILLGLPFCEK obtenus en mode DIA-LID à 473 nm au point de gamme 0,001 117
Figure 56 : Linéarité obtenue en mode DIA-LID à 473 nm pour le peptide à cystéine dérivé VMCDSNDLPK ajouté dans 100 µg d'E.Coli 119
Figure 57 : XIC des 4 peptides à cystéine dérivés endogènes détectés en mode DIA-LID à 473 nm dans l'extrait cellulaire issus de protéines humaines kinases 121
Figure 58 : XIC des 7 peptides à cystéine dérivés endogènes détectés en mode DIA-LID à 473 nm dans l'extrait cellulaire issus de protéines humaines 123
Figure 59 : Structure du chromophore JCM 1425 A (a) et du DABCYL C2-maléimide (b) 126
Figure 60 : Structure du chromophore JCM 1425 A attendu à pH = 7 (a) et pH inférieur à 5 (b) 126
Figure 61 : Spectres FMS (a) et PRM LID à 473 nm (b) obtenus pour JCM 1425 A infusé à pH = 7 en matrice aqueuse 127
Figure 62 : Spectres FMS (a) et PRM LID à 473 nm (b) obtenus pour JCM 1425 A infusé à pH = 2,5 en matrice aqueuse 128
Figure 63 : Structure de JCM 1425 A' avec localisation du site de fragmentation générant l'ion rapporteur m/z 296,14 128
Figure 64 : Zoom sur le spectre FMS obtenu pour le peptide IVGGWECEK dans les gammes de m/z attendues des formes trichargées avec le chromophore non hydrolysé (a) et hydrolysé (b) 129
Figure 65 : Zoom sur le spectre FMS dans les gammes de m/z attendues des formes trichargées obtenu pour le peptide à cystéine dérivé IVGGWECEK avec le chromophore non hydrolysé (a) et hydrolysé (b) après chauffage à 40°C pendant 45min 130
Figure 66 : Zoom sur le spectre FMS dans les gammes de m/z attendues des formes trichargées pour le peptide à cystéine dérivé IVGGWECEK avec le chromophore non hydrolysé (a) et hydrolysé (b) après chauffage à 60°C pendant 6h 131
Figure 67 : Spectres de photo-fragmentation des formes trichargées du peptide IVGGWECEK dérivés avec le DABCYL C2-maléimide (a) et JCM 1425 A' (b) 132
Figure 68 : Spectre de photo-fragmentation des formes trichargées du peptide à cystéine dérivé IVGGWECEK avec JCM 1425 A' avec un temps d'activation de 25 ms (a) et 5 ms (b) 133
Figure 69 : Spectre de photo-fragmentation à 473 nm du peptide à cystéine dérivé ELGAALQGCLLR 139
Figure 70 : Types de fragments produits en fonction de l'acide aminé qui est fragmenté 141
Figure 71 : Rang d'intensité moyen des fragments générés par acide aminé 145
Figure 72 : Nombre de fragments détectés en fonction de la distance entre la cystéine et le site de fragmentation a été tracé pour chaque taille de peptide à cystéine 147
Figure 73 : Proportion du type de fragment généré au niveau des acides aminés P et G relativement à la position de la cystéine 151
Figure 74 : Schéma du Qexactive avec un alignement particulier du laser permettant une photo-fragmentation au sein de la C-trap 161
Figure 75 : Spectre FMS obtenu pour le DABCYL-C2 maléimide avec un alignement particulier du laser permettant une photo-fragmentation au sein de la C-trap lorsque le laser est allumé (a) ou éteint (b) 161
Figure 76 : Schéma du Qexactive avec un alignement particulier du laser permettant la C-trap-LID et un obturateur 164
Figure 77 : Signaux inconnus détectés et observés à l'oscilloscope 165
Figure 78 : Paramètres modifiables lors de l'utilisation du Qexactive en mode FMS 167

Figure 79 : Spectre de masse obtenu lors de l'infusion du DABCYL C2-maléimide lorsque l'obturateur est fermé (a) ou ouvert (b)..... 169
Figure 80 : XIC correspondant à l'ion rapporteur m/z 252,11 du DABCYL C2-maléimide (a) et à l'ion précurseur du DABCYL C2-maléimide (b)..... 170
Figure 81 : XIC (a) et spectre de masse obtenu pour le peptide à cystéine dérivé LTSVVGTILYSCPEVLK lorsque l'obturateur est fermé et ouvert (b et c) ainsi que pour le peptide à cystéine dérivé VLDANSCQSELHEK (d, e, f)..... 172
Figure 82 : XIC des peptides à cystéine dérivés LTSVVGTILYSCPEVLK (a) et LQGINCGPDFTPSFANLGR (b)173
Figure 83 : TIC extrait entre 16,6 et 17,8 min..... 173
Figure 84 : Soustraction des scans FMS et C-Trap-LID à 473 nm au RT correspondant à l'apex du pic chromatographique du peptide à cystéine SNCITENPAGPYGQR dans un digest de 3 protéines 178

Liste des tableaux

Tableau 1 : Liste non exhaustive de biomarqueurs protéiques associés à une pathologie particulière 25
Tableau 2 : Variation de la réponse en intensité de 10 ions précurseurs tri ou quadrichargés issus de 10 peptides à cystéine dérivés en fonction de la composition du solvant de reprise post évaporation..... 70
Tableau 3 : Caractéristique de la colonne chromatographique utilisée pendant tout ce chapitre 71
Tableau 4 : Variation de la composition en solvant B de la phase mobile en fonction du temps lors du gradient 1 71
Tableau 5 : Variation de la composition en solvant B de la phase mobile en fonction du temps lors du gradient 2 72
Tableau 6 : Temps d'analyse de l'orbitrap en fonction de la résolution..... 84
Tableau 7 : Paramétrage du Qexactive fixé en fonction du mode utilisé 85
Tableau 8 : Taux de photo-fragmentation obtenus pour 10 peptides à cystéine dérivés 95
Tableau 9 : Couvertures de séquence obtenues pour les 10 peptides à cystéine dérivés 96
Tableau 10 : Proportion (cumulée ou non) de chaque taille de fragment (nombre d'acides aminés) au sein des 6 fragments les plus intenses issus des 354 spectres de photo-fragmentation de la bibliothèque de spectres LID à 473 nm. 108
Tableau 11 : Nombre, pourcentage et pourcentage cumulé de peptides à cystéine dérivés identifiés en mode DIA-LID à 473 nm en fonction de la taille de la fenêtre en RT utilisée 110
Tableau 12 : Nombre, pourcentage et pourcentage cumulé de peptides à cystéine dérivés identifiés en fonction du dopt score 110
Tableau 13 : CV (%) obtenus sur 3 réplicas en mode DIA-LID à 473 nm pour les 65 peptides à cystéine dérivés ajoutés dans 100 µg d'E.Coli pour les 4 points de concentration 117
Tableau 14 : Aire et CV (%) obtenus pour 3 réplicas en mode DIA-LID à 473 nm pour le peptide à cystéine dérivé VMCDSNDLPK ajouté dans 100 µg d'E.Coli pour les 8 points de concentration..... 118
Tableau 15 : Liste des 4 peptides à cystéine dérivés endogènes détectés au sein de l'extrait cellulaire en mode DIA-LID à 473 nm. La liste des protéines kinases humaines desquels ils sont issus, leur meilleure transition et le CV correspondant sont également présentés 120
Tableau 16 : Liste des 7 peptides à cystéine dérivés endogènes détectés en mode DIA-LID à 473 nm au sein de l'extrait cellulaire. La liste des protéines humaines desquels ils sont issus, leur meilleure transition et le CV correspondant sont également présentés 122
Tableau 17 : Gain en signal de 10 ions précurseurs issus de 10 peptides à cystéine dérivés issus des 3 protéines digérées dû à une étape de dérivation réalisée post digestion..... 125
Tableau 18 : Taux de photo-fragmentation du peptide à cystéine dérivé IVGGWECEK avec le DABCYL C2-maléimide et JCM 1425 A'..... 132
Tableau 19 : % du type de fragment généré en moyenne pour les 401 peptides à cystéine dérivés de la bilbiothèque de spectres LID à 473 nm..... 140
Tableau 20 : % du type de fragments généré en fonction du rang du fragment..... 140
Tableau 21 : % de fragments généré par acide aminé en fonction de leur abondance dans la séquence de tous les peptides à cystéine de la bibliothèque de spectres LID à 473 nm..... 143
Tableau 22 : Position moyenne relativement à la position de la cystéine pour chaque acide aminé au sein des 401 peptides à cystéine de la bibliothèque de spectres LID à 473 nm..... 144
Tableau 23 : Nombre de peptides à cystéine dérivés issus de la bibliothèque LID à 473 nm utilisés pour chaque taille de séquence peptidique 148
Tableau 24 : Influence de la position de la cystéine sur les types d'ions fragments produits..... 150

Tableau 25 : % de fragments et d'ions y générés au niveau de P et G par rapport à la totalité des fragments et ions y générés au niveau de tous les acides aminés pour chaque quartile..... 152
Tableau 26 : Influence de la position de la cystéine sur le type d'ion fragment produit sans compter les acides aminés P et G..... 153
Tableau 27 : Tableau récapitulatif des observations relatives au modèle de fragmentation des peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm..... 156
Tableau 28 : Variation de la composition en solvant B de la phase mobile en fonction du temps lors du gradient 3 160
Tableau 29 : Variation de la composition en solvant B de la phase mobile en fonction du temps lors du gradient 4 160
Tableau 30 : Paramétrage du Qexactive 165
Tableau 31 : Taux de photo-fragmentation obtenus pour les 10 peptides à cystéine dérivés..... 174
Tableau 32 : Nombre de fragments détectés pour les 10 peptides à cystéine dérivés 176
Tableau 33 : Masses extraites correspondants à des peptides à cystéine dérivés putatifs..... 178

Introduction générale

De nos jours, la maladie et la durée de vie sont deux sujets de préoccupation majeurs au sein de notre système social. Il y est essentiel de minimiser la rencontre avec le premier et de maximiser l'étendue du second dans des conditions les plus acceptables possibles. Dans cette optique, une attention toute particulière a été portée sur la manière de détecter le plus rapidement possible la présence d'une maladie au sein d'un organisme humain. Selon ce projet, les laboratoires cliniques et médicaux procèdent à de nombreux tests dits de dépistages. Ces tests se basent principalement sur la détection et la quantification de biomarqueurs où les biomarqueurs se doivent de refléter l'état physiologique de l'organisme humain examiné. Dans le domaine médical, une multitude de symptômes ou de composés peuvent jouer le rôle de biomarqueurs dont les protéines[1]. Ces macromolécules, dosables dans les fluides biologiques, assurent des rôles essentiels nécessaires au maintien de l'homéostasie du corps humain et plus précisément au niveau de l'activité cellulaire. Elles peuvent y jouer un rôle structurel (comme l'actine), un rôle dans la mobilité (comme la myosine), un rôle catalytique (les enzymes), un rôle de régulation de la compaction de l'ADN (les histones) ou encore d'expression des gènes (les facteurs de transcription). En somme, l'immense majorité des fonctions cellulaires est assurée par des protéines. Qui plus est, le moindre impact maladif sur notre organisme va se solder par la variation du taux de certaines protéines au sein de celui-ci. De ce fait, pouvoir doser des protéines spécifiquement et sensiblement semble fondamental d'un point de vue médical.

De nombreux procédés ont été mis au point pour doser des biomarqueurs protéines putatifs[2,3]. Ce sont généralement des méthodes spectrophotométriques basées sur diverses caractéristiques spectrales ou réactionnelles des acides aminés constituant ces macromolécules biologiques. Le choix dépend des besoins et des caractéristiques recherchées : fiabilité, sensibilité, rapidité, taille des échantillons, possibilité de récupérer l'échantillon après dosage ou encore présence de substances interférentes dans l'échantillon. Mais toutes les méthodes anciennes ou actuelles sont plus ou moins limitées par l'une de ces caractéristiques : la sensibilité, les problèmes d'interférences ou encore des tests avec des réactivités très aléatoires selon les protéines[4]. A l'heure actuelle, les méthodes de références utilisées sont les tests immuno- enzymatiques, très sensibles, permettant de détecter la présence d'antigènes au sein d'un organisme humain grâce à des anticorps spécifiques. Cependant, les développements nécessaires à la mise en place de nouveaux dosages au moyen de ces tests sont extrêmement longs (étape de production et de validation d'anticorps spécifiques) et l'avenir se tournant vers le dosage multiplexé, ils ne pourront répondre que difficilement à cette attente.

C'est pourquoi, pour satisfaire la demande liée au multiplexage, une alternative reposant sur l'utilisation de la spectrométrie de masse s'est fortement développée ces dernières années[5]. Cette technique physique d'analyse permet de détecter et d'identifier des molécules d'intérêt par mesure de leurs masses, et de caractériser leurs structures chimiques. A travers cette technique, il est possible de réaliser des dosages de protéines multiplexés, rapidement et sans avoir recours à la production d'anticorps spécifiques. Cependant, son défaut principal réside dans sa sensibilité qui est bien inférieure aux tests immuno- enzymatiques (µg-ng/mL contre du pg/mL)[6,7]. Cela est surtout dû à la complexité des matrices biologiques et à la préparation des échantillons biologiques avant leur analyse en spectrométrie de masse où les protéines sont digérées en une multitude de peptides pour pouvoir être identifiées par leurs intermédiaires. Il s'est donc avéré nécessaire de focaliser la recherche sur l'amélioration de la sensibilité de cette technique. A dessein, différentes alternatives peuvent être entreprises dont la purification des échantillons biologiques avant analyse afin d'en diminuer la complexité. La déplétion de protéines majoritaires ainsi que l'immuno- enrichissement peuvent être des voies à suivre[8]. Mais il ne faut pas s'arrêter à la complexité de la matrice, car en cumulant plusieurs préparations d'échantillons il y a un risque réel de perte de l'analyte. En attendant le perfectionnement de l'appareillage utilisé dans les années à venir, il est également possible d'influer sur les méthodes de fragmentation utilisées. En effet, lors d'une analyse de peptides issus d'une matrice biologique par spectrométrie de masse, ceux-ci ont besoin d'être fragmentés pour être identifiés. Cependant, la méthode de fragmentation classiquement utilisée intitulée « Collision Induced Dissociation » (CID) n'est pas spécifique ce qui engendre des problèmes d'interférences et nuit à leur identification[9]. Il devient alors possible de faciliter la détection des peptides en se focalisant sur le développement et l'utilisation de nouvelles méthodes de fragmentation plus spécifiques.

Parmi les nouvelles méthodes de fragmentation en développement, la photo-fragmentation laser des composés biologiques (dont les peptides) est depuis plusieurs années utilisée[10–12]. Cette méthodologie repose sur l'activation de molécules biologiques ciblées en fonction de leurs propriétés optiques. En effet, pour qu'un édifice moléculaire puisse photo-fragmenter il doit absorber les photons émis par le laser à une longueur d'onde donnée : la palette de longueurs d'ondes accessibles étant très vaste et les composés biologiques n'absorbant pas à toutes les longueurs d'ondes, il devient possible d'obtenir une spécificité lors de l'étape de photo-fragmentation. Ainsi, différentes études relatives à la photo-fragmentation des peptides à différentes longueurs d'ondes (ultraviolet, infrarouge etc.…) ont été développées[13]. Certaines se sont axées sur une photo-fragmentation des peptides dans le domaine du visible. Les molécules biologiques et donc les peptides n'absorbant pas dans le visible, il est nécessaire de les dériver spécifiquement avec un groupement chromophore pour qu'ils puissent photo-fragmenter à la longueur d'onde d'émission du laser, le chromophore absorbant à cette énergie. Une méthodologie intitulée « Laser Induced Dissociation » (LID) à 473 nm et ciblant spécifiquement les peptides à cystéine à l'aide du

chromophore DABCYL-C2 maléimide a mis en évidence pragmatiquement son efficience à simplifier les spectres de masse obtenus et à identifier spécifiquement des peptides à cystéine et les protéines desquels ils étaient issus[9,14]. L'intérêt principal de cette technique réside dans son ciblage spécifique des peptides à cystéine. En effet, l'acide aminé cystéine est un acide aminé rare, il est présent dans 89 % des protéines humaines à raison d'un acide aminé par protéine[15]. Cela signifie qu'en suivant spécifiquement les peptides à cystéine il est possible d'être exhaustif en couvrant potentiellement tout le protéome humain et d'autre part de subvenir au problème de spécificité inhérent à la méthode de fragmentation CID. C'est ici qu'intervient mon travail de thèse. Ce travail s'est inscrit dans la continuité du développement et de l'application de cette méthode d'activation pour détecter spécifiquement des peptides à cystéine rapporteurs de biomarqueurs protéines putatifs au sein de matrices biologiques humaines.

Ce travail de thèse est organisé en quatre chapitres :

Le premier chapitre est une étude bibliographique qui définira la notion de biomarqueurs au sein du milieu médical ainsi que les différentes techniques analytiques existantes permettant leur suivi. Il en découlera une exposition des raisons qui font des protéines kinases des biomarqueurs protéines putatifs pertinents chez l'homme et la spectrométrie de masse une alternative crédible afin de les suivre. Une attention toute particulière sera portée sur les différents modes d'analyses et méthodes d'activation accessibles en spectrométrie de masse et les possibilités qu'ils sous-tendent. En découlera la raison pour laquelle le mode d'acquisition « Data-Independent Acquisition » (DIA) couplée à la méthode d'activation LID à 473 nm semble adapté pour suivre, identifier et quantifier des biomarqueurs protéines, dont les protéines kinases, par spectrométrie de masse au sein de fluides biologiques humain.

Le second chapitre présentera le développement, la mise en place et l'utilisation de la méthodologie DIA-LID à 473 nm sur un spectromètre de masse haute résolution Qexactive pour identifier et quantifier des peptides à cystéine rapporteurs de protéines kinases humaines au sein d'un extrait cellulaire mammaire cancéreux.

Le troisième chapitre se focalisera sur une première investigation concernant le modèle de fragmentation des peptides à cystéine dérivés avec le DABCYL C2-maléimide en LID à 473 nm. Il sera exposé les premières observations relatives aux mécanismes mis en jeu lors de cette étape de photo-fragmentation.

Le quatrième chapitre se concentrera sur l'introduction d'une nouvelle méthodologie intitulée C-trap-LID. Celle-ci repose sur l'exécution de l'étape de photo-fragmentation, non plus dans la cellule de collision, mais directement au sein de la C-trap du Qexactive.

Chapitre I : Identification et quantification de biomarqueurs protéines par spectrométrie de masse – état de l'art

Introduction

Ce chapitre a pour but de présenter les raisons qui nous ont amenées à vouloir suivre la régulation de biomarqueurs protéines putatifs au sein de matrices biologiques humaines par le biais de la spectrométrie de masse pour le dépistage de maladies chez l'homme.

Dans cette optique nous définirons tout d'abord la notion de biomarqueur au sein du milieu médical, les raisons de leurs études ainsi que les différentes techniques disponibles à l'heure actuelle permettant leur identification et leur validation. Subséquemment à l'exposition des raisons qui font de la spectrométrie de masse une des alternatives intéressantes, il sera présenté les arguments faisant des protéines kinases humaines des biomarqueurs putatifs pertinents pour le dépistage de maladies chez l'homme, ainsi que les premiers résultats encourageants quant à leur traçabilité à l'aide de cet outil analytique.

En suivra les différentes méthodologies et modes offerts par la spectrométrie de masse pour identifier et quantifier des biomarqueurs protéines dans des matrices biologiques. Une attention toute particulière sera portée sur les différents modes de fragmentation accessibles et les possibilités qu'ils sous-tendent. En découlera pourquoi le mode « Data-Independent Acquisition » (DIA) couplé à la méthode d'activation intitulée « Laser Induced Dissociation » à 473 nm (LID) nous ont semblé adapter au suivi des protéines kinases humaines au sein de matrices biologiques.

I) Les biomarqueurs

Pour remédier aux dérèglements de l'homéostasie de l'organisme humain, des processus de recherche et de développement au niveau pharmaceutique sont initiés. Ils auront pour but à la fois d'être capables de détecter ces altérations de la santé le plus précocement possible et de fournir des remèdes pouvant lutter contre elles. Dans cette optique, les processus de recherche et de développement se sont axés vers la découverte et la validation de biomarqueurs.

I.1) Définition

Selon le National Institute of Health (USA)[16,17], la définition proposée pour un biomarqueur est la suivante : "un biomarqueur représente une caractéristique biologique mesurée de façon objective et évaluée comme un indicateur soit de processus biologiques normaux ou pathologiques, soit de réponses pharmacologiques résultant d'une intervention thérapeutique". Il s'agit donc d'une corrélation entre un marqueur biologique et un résultat clinique observé (réponse à un traitement, contraction d'une maladie etc…)

I.2) La découverte de nouveaux biomarqueurs

Suivant cette définition énormément de paramètres peuvent donc jouer le rôle de biomarqueurs tel que la variation de la température corporelle qui est, par exemple, un indicateur de la contraction du paludisme lors d'un retour de voyage ou les nausées, symptomatiques d'une trop forte exposition au soleil ou encore d'une grossesse.

Ainsi, de nombreuses voies sont empruntées dans le but de mettre à jour des nouveaux biomarqueurs qui pourront être caractéristiques de perturbations via leur seul ciblage. Dans ce sens, un intérêt particulier s'est développé quant à l'étude des fluides biologiques (sang, urine, tissu) ceux-ci contenant, transportant et évacuant une multitude de composés présents au sein de l'organisme. Ils permettent alors de poser un regard sur l'évolution (présence, absence, variation en concentration) de ces composés au sein de l'organisme et de potentiellement les relier à un état pathologique particulier. Des études différentielles pourront ainsi être effectuées entre un patient considéré comme sain et un autre malade (ou plusieurs malades aux différents stades d'une maladie) en suivant les variations de l'expression de composés suspectés.

Différents composés sont envisagés dans les fluides biologiques tels que les sucres[18], les lipides (Alzheimer, Diabètes etc…)[19] ou encore les protéines qui font l'objet d'un réel intérêt [1]. Celles-ci, assurant énormément de fonction au sein de l'organisme humain, sont suspectées de pouvoir refléter ces perturbations homéostasiques[20]. En effet, en 1949 Linus Pauling démontra que l'altération de l'hémoglobine menait à la drépanocytose[21]. Dans le courant des années 1960, il a été démontré que la protéine AFP pouvait jouer le rôle de marqueur tumoral[22]. Ou encore en 1980 où la glycoprotéine CA 125 a été assimilée en tant qu'indicateur de cancers de l'ovaire[23]. Ces

quelques exemples parmi beaucoup d'autres (Tableau 1)[24] sont révélateurs du lien étroit qui lie la fonction et l'expression des protéines au phénotype.

Biomarqueur	Type de protéines	Matrice biologique	Type de cancer
α-Fetoprotein	Glycoprotein	Serum	Nonseminomatous testicular
Human chorionic gonadotropin-β	Glycoprotein	Serum	Testicular
CA19-9	Carbohydrate	Serum	Pancreatic
CA125	Glycoprotein	Serum	Ovarian
Pap smear	Cervical smear	Cervix	Cervical
CEA	Protein	Serum	Colon
Epidermal growth factor receptor	Protein	Colon	Colon
KIT	Protein (IHC)	Gastrointestinal tumour	GIST
Thyroglobulin	Protein	Serum	Thyroid
PSA (total)	Protein	Serum	Prostate
PSA (complex)	Protein	Serum	Prostate
PSA (free PSA%)	Protein	Serum	Prostate
CA15-3	Glycoprotein	Serum	Breast
CA27-29	Glycoprotein	Serum	Breast
Cytokeratins	Protein (IHC)	Breast tumour	Breast
Oestrogen receptor and progesterone receptor	Protein (IHC)	Breast tumour	Breast
HER2/NEU	Protein (IHC)	Breast tumour	Breast
HER2/NEU	Protein	Serum	Breast
HER2/NEU	DNA (FISH)	Breast tumour	Breast
Chromosomes 3, 7, 9 and 17	DNA (FISH)	Urine	Bladder
NMP22	Protein	Urine	Bladder
Fibrin/FDP	Protein	Urine	Bladder
BTA	Protein	Urine	Bladder
High molecular weight CEA and mucin	Protein (Immunofluorescence)	Urine	Bladder

Tableau 1 : Liste non exhaustive de biomarqueurs protéiques associés à une pathologie particulière

Au niveau de l'étude des biomarqueurs protéiques, leur structure 3D peut être déterminée expérimentalement ou par calculs de modélisation moléculaire[25] et leur régulation peut être suivie selon 3 voies principalement : les dosages après électrophorèse sur gel de polyacrylamide en deux dimensions, les tests immuno- enzymatiques et différentes approches en spectrométrie de masse.

Concernant la technique des dosages après électrophorèse sur gel de polyacrylamide en deux dimensions : celle-ci a été introduite dans les années 1970[3] et consiste en la séparation des protéines sur gel selon leur point isoélectrique (pH auquel la protéine est globalement neutre) dans une première dimension et selon leur masse dans une seconde dimension. Leur détection est ensuite effectuée par coloration de gels (bleu de Coomassie ou nitrate d'argent). Ces techniques ont largement démontré leur efficacité par le passé[26] mais présentent quelques limitations à travers leur étape de détection (la détection au bleu de Coomassie est peu sensible et celle au nitrate d'argent ne permet pas toujours l'obtention de réponse linéaire).

Dans le même temps les tests immuno- enzymatiques ont été développés[2]. Ceux-ci sont devenus les méthodes de références utilisées lors du développement de test de biomarqueurs et sont

mondialement connus sous le nom d'Enzyme-linked ImmunoSorbent Assay (ELISA). Leur procédé repose sur l'utilisation d'anticorps pour la détection spécifique d'antigène matérialisé par 3 grandes méthodes de dosage (ELISA indirect, ELISA en sandwich et ELISA par compétition). Les avantages majeurs proposés par ces techniques sont la simplicité d'utilisation, leur spécificité et la sensibilité importante atteignable. Cependant, ces techniques font aussi face à leurs propres limitations dont les plus importantes reste la difficulté de synthétiser les anticorps adaptés et le faible multiplexage proposé[27].

Pour répondre à la nécessité liée au multiplexage, la spectrométrie de masse se propose comme une alternative de plus en plus crédible malgré une sensibilité moindre que les tests ELISA. Le développement rapide et le multiplexage massif permis par la spectrométrie de masse[1] rendent possible l'observation des variations de régulation d'un grand nombre de protéines lors de perturbations subies par un organisme. Les processus dans lesquels ces protéines sont impliquées peuvent être déterminés ce qui peut permettre de remonter aux causes/conséquences générées par ces perturbations. C'est donc un regard exhaustif sur le contenu des fluides biologiques, essentiel lors d'une étude de découverte de biomarqueurs, qu'autorise cette technique. Celle-ci peut d'ailleurs être corrélée aux procédés de séparation sur gel de polyacrylamide en deux dimensions et se substituer à la détection par coloration de gels.

I.3) La validation de nouveaux biomarqueurs

Lorsque des biomarqueurs (protéines ou non) sont suspectés, ils doivent passer par une étape de validation où leur pertinence clinique ainsi que la possibilité de les analyser de manière reproductible doit être démontrée à travers des analyses robustes sur plusieurs milliers d'échantillons[28].

Suivant l'étape de validation pour les méthodes de dosages proposée par l'organisation internationale de la standardisation (IOS), il est impératif d'établir les possibilités intrinsèques de la méthode de dosage qui est retenue ou souhaitée en questionnant les valeurs cibles à atteindre pour le biomarqueur, la capacité de la méthode choisie à les atteindre et les performances techniques de la méthode en termes de robustesse etc...

Dans ce sens, l'Association américaine des sciences pharmaceutiques (AAPS) a proposé une méthodologie adaptée au concept de la validation de biomarqueurs[29] :

- Etape 1 : Définition de l'objectif et sélection de la méthode de dosage
- Etape 2 : Rédaction de la procédure de validation de la méthode de dosage
- Etape 3 : Validation technique de la méthode de dosage : évaluation des performances et de la capacité à refléter la réalité biologique du biomarqueur

-Etape 4 : Validation clinique du biomarqueur et de sa méthode de dosage
-Etape 5 : Contrôle qualité de la méthode de dosage lors de son utilisation en routine.

Durant le courant de ma thèse, nous ne nous sommes pas tournés vers la partie validation mais sur celle qui prend pour objet la découverte de biomarqueurs. À la vue des précédentes considérations à son sujet et dans l'optique de développer une méthodologie pour identifier des biomarqueurs au sein de matrices biologiques diverses, il a été choisi de se tourner vers la spectrométrie de masse grâce au potentiel lié à son multiplexage important et à son développement rapide.

I.4) Les protéines kinases humaines

Au niveau des biomarqueurs à suivre en spectrométrie de masse un choix arbitraire a été effectué : ce sont les protéines qui ont été ciblées. Ceci résulte à la fois des raisons présentées plus haut mais aussi des compétences acquises en protéomiques par l'équipe au sein du laboratoire. Deux approches se sont alors offertes à nous :

- Une approche où le protéome entier est analysé et aucune supposition n'est faite quant aux protéines pouvant être impliquées dans un état pathologique particulier. Tout est ouvert au regard pour répondre de l'état pathologique particulier.

- Une approche où des biomarqueurs potentiels sont déjà présélectionnés en fonction des connaissances du milieu scientifique sur l'état pathologique approché. Eux-seuls seront sondés.

L'approche choisie n'a pas été de sonder tout le protéome à l'aveugle mais de déjà présélectionner des biomarqueurs potentiels par rapport aux connaissances scientifiques acquises. Les protéines kinases ont été choisies. En vue d'exposer les raisons de leur ciblage, il sera présenté leur rôle physiologique, le fonctionnement et la régulation de leur activité et les méthodes de dosages actuelles adéquates pour les identifier et les quantifier.

I.4.a) Rôle physiologique

Les protéines kinases sont des enzymes qui appartiennent au groupe des transférases. Celles-ci ont pour fonction de catalyser le transfert d'un groupement fonctionnel d'une molécule donneuse à une molécule accepteuse. Pour les kinases il s'agit de catalyser, au sein des cellules, les réactions de phosphorylation (ajout d'un groupement phosphate) à partir des molécules donneuses d'adénosine triphosphate (ATP) sur des molécules cibles de différents types (lipides, sucres) dont des protéines (Figure 1). Au sein de ces protéines ciblées ce sont les groupements hydroxyles des chaines latérales des acides aminées sérine, thréonine, et tyrosine qui sont phosphorylés à des taux normalisés de 1000, 100 et 1 respectivement chez les eucaryotes. Ce processus de phosphorylation est connu depuis 1906[30] et la première description de la

phosphorylation d'une protéine date de 1954[31]. Par la suite, l'étude ne portera que sur les protéines kinases humaines.

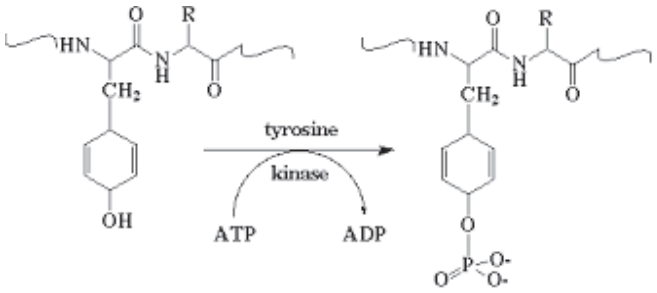


Figure 1 : Phosphorylation d'un acide aminé tyrosine par le biais d'une tyrosine kinase et de l'ATP[32]

Ces réactions de phosphorylations sont assimilées à des modifications post-traductionnelles (PTM) majeures car on estime que 1/3 du protéome humain total est touché par ce type de modification[33]. Leur intense activité est un marqueur de leur poids conséquent au niveau de la régulation cellulaire (différenciation, division, prolifération, apoptose…) : celles-ci interviennent dans la régulation de la plupart des voies métaboliques et des processus cellulaires de l'organisme humain[34]. Les premiers à avoir mis en avant leur fonctionnalité sont Krebs et Fischer en 1956 lors de leur travail sur l'activation de la phosphorylase b dans le muscle squelettique[35]. En effet, la transmission d'un groupement PO_3^{2-} va inhiber ou amplifier l'action d'un substrat protéique, et ainsi moduler et réguler des processus vitaux de la cellule par le biais de réponses spécifiques. Etablir une liste exhaustive de l'ensemble des processus de signalisation cellulaires où les kinases interviennent est trop conséquent. Néanmoins, pour donner un aperçu de l'étendue et de l'importance de leur activité, quelques-uns des processus associés à leurs actions peuvent être listés :

- La régulation de la transcription et la différenciation des cellules nerveuses par les cyclin-dependent kinases (CDK)[36].
- La régulation du taux de Ca^{2+} dans les cellules du cœur (Protein Kinase A)[37].
- La régulation de la sécrétion des fluides intestinaux (Protein Kinase G)[38].
- La Transcription de l'ADN (Casein kinase 1)[39].
- La réponse immunitaire (Glycogen synthase 3)[40].
- La suppression des tumeurs dans les mélanomes[41].

A travers ces quelques exemples nous pouvons entrapercevoir que les kinases sont impliquées dans des processus régulant l'homéostasie du cœur, du système cérébral, intestinal, des réponses immunitaires et donc de l'homéostasie de l'organisme entier. Il est donc possible qu'elles puissent

réellement jouer le rôle de marqueurs quant à l'état de santé de celui-ci comme nous le verrons par la suite.

I.4.b) Mécanisme de l'activité catalytique des kinases

La partie principale de la structure des protéines kinases se situent au niveau de leur domaine catalytique puisque c'est à cet endroit qu'il y a fixation de la molécule d'ATP, du substrat et transfert du groupement phosphate. La structure de cette zone et ses mécanismes sont détaillés ci-dessous.

I.4.b.1) Structure du domaine catalytique

A l'heure actuelle, 518 protéines kinases (1,7% du protéome humain) ont été recensées dans le protéome humain et chacune possède un domaine catalytique spécifique[34]. Il est néanmoins possible de les regrouper en 9 sous-groupes en fonction de leur similarité de séquence et de structure 3D[34] : AGC, CAMK, CMGC, CK1, STE, TK, TKL, RGC et atypiques. Voici une représentation du domaine catalytique (Figure 2) :

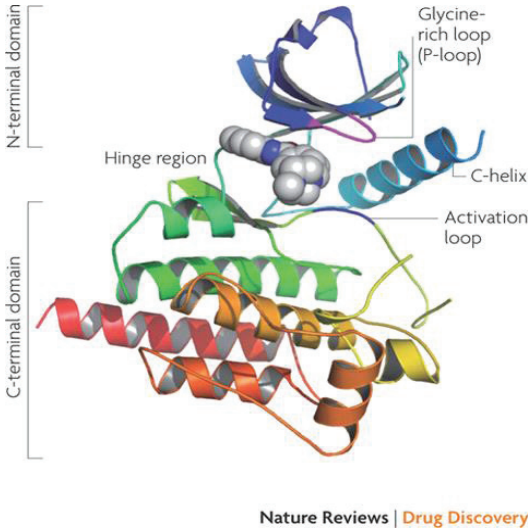


Figure 2 : Représentation en 3 dimensions du domaine catalytique des protéines kinases[42]

Ce domaine catalytique est constitué d'environ 270 acides aminés répartis en 2 lobes :

- le lobe N-terminal est constitué d'un feuillet antiparallèle à cinq brins β et d'une hélice C_α comprenant 80 acides aminés. La boucle P relie les deux premiers brins β.

- le lobe C-terminal, comprenant 190 acides aminés, est essentiellement constitué d'hélice α, de deux brins β et d'une boucle d'activation.

- Les deux lobes sont reliés par la région charnière qui joue sur la flexibilité respective attribuée aux deux lobes[43].

L'étape de phosphorylation s'effectue au niveau du site actif à l'intersection des 2 lobes où 2 poches accolées vont accueillir respectivement le substrat et l'ATP. La fixation des 2 composés se fait de manière indépendante[44]. Le site de reconnaissance du substrat se situe au niveau de la boucle d'activation et par sa variabilité inter-kinases (longueur, séquence etc..) va contribuer à la spécificité et au rôle de chacune : une kinase pour un substrat particulier. Le substrat sera phosphorylé via un transfert du groupement phosphate en position γ de l'ATP sur l'hydroxyle de la sérine, thréonine ou tyrosine. Pour ce faire, l'ATP doit être dans un positionnement particulier : le groupement ribose est stabilisé par des liaisons hydrogènes avec des résidus du lobe C-terminal, les 3 phosphates sont alignés en partie par leur interaction avec la boucle P et la partie adénine est orientée par des liaisons hydrogènes dans une poche formée de résidus hydrophobes au sein des deux lobes.

I.4.b.2) Régulation de l'activité catalytique

La fonction catalytique des kinases est régulée principalement selon 2 manières : (i) soit la kinase est dite sous forme active mais son environnement biologique empêche la fixation du substrat ou de l'ATP sur leur site respectif (ii) soit la kinase est dite sous forme inactive en raison d'une modification de sa conformation spatiale par régulation allostérique. Il devient alors impossible à tout substrat et/ou ATP de venir se fixer : l'activité catalytique est inhibée[43,45]. Le premier mode de régulation peut être réalisée directement par la kinase elle-même, on parle alors d'auto-inhibition (une partie externe au domaine catalytique de la kinase vient bloquer l'accès au site de fixation). Par ailleurs une séquestration de la kinase dans une région particulière de la cellule où le substrat n'est pas présent permet aussi une régulation de son activité. La seconde régulation se fait, comme la définition de la régulation allostérique l'indique, via la fixation d'un composé particulier qui va engendrer un changement de la conformation de la kinase pour l'activé ou non. A noter que lorsque la régulation d'une protéine kinase nécessite l'intervention d'un composé externe, bien souvent il s'agit d'une autre protéine kinase qui l'active par phosphorylation, on parle alors d'activation en cascades des protéines kinases[46].

Si la régulation des kinases et leur activité sont essentielles à l'homéostasie de l'organisme humain, la dérégulation des mécanismes qui les contrôlent (non activation, activation non désirée, sur ou sous activation) a été corrélée à de nombreuses maladies allant du cancer[47–49] aux maladies cardiovasculaires, neurologiques ou encore métaboliques[50–52]. Qui plus est, une étude a démontré que dans 210 formes humaines de cancer, les séquences codantes de 120 kinases présentaient des mutations susceptibles d'être responsables de la maladie[53]. Les protéines kinases peuvent donc bien jouer le rôle de biomarqueurs dans des approches pharmacologiques liés à des états pathologiques particuliers et ont déjà été établies comme telles.

I.4.c) Dosage des kinases

Plusieurs approches ont été développées en vue d'identifier et de quantifier des kinases dans des matrices biologiques humaines. Tout d'abord, les méthodes immuno- enzymatiques utilisées

pour mesurer l'abondance et l'activité des kinases individuellement[54]. Cependant, il a rapidement été remarqué que cela n'était pas suffisant en raison de la vision non exhaustive de la régulation de la signalisation cellulaire et donc des relations inter-kinases[55,56]. Des études sur le kinome entier ont alors été entreprises utilisant la spectrométrie de masse couplée à la chromatographie en phase liquide (LC/MS/MS)[57–59]. Malheureusement, l'importante gamme dynamique au niveau de l'expression des kinases entre elles et d'un type de cellule a un autre ont été un frein nécessitant le développement et l'utilisation de techniques d'enrichissement particulières au préalable[57]. Différentes méthodologies ont ainsi été testées : l'IMAC (immobilized metal ion chromtaography : enrichissement des protéines kinases via leur intéraction avec des groupements métalliques type Fe^{III} par exemple)[60,61], la MOC (metal oxyde chromatography où les peptides ou protéines phosphorylés sont retenus par des particules d'oxyde métalliques en raison de l'affinité des groupements phosphates)[62,63], l'immuno-précipitation[64] (précipitation spécifique des protéines kinases en solution via leur intéraction avec des anticorps type Py par exemple) ainsi que des analogues de l'ATP[57]. Les résultats, quant à la détection des kinases par le biais de ces techniques d'enrichissement, corrélés à l'augmentation des performances technologiques se sont avérés satisfaisants et prometteurs. En effet, 261 kinases humaines ont ainsi pu être identifiées au sein d'extrait cellulaire[59] et 212 quantifiées dans 6 extraits cellulaires cancéreux différents[57].

Ces résultats positifs quant à la détection et au dosage des kinases par LC/MS/MS appel à persévérer dans cette voie. Qui plus est, la nécessité pharmacologique quant au développement d'une méthodologie couvrant le kinome entier nous a décidé à nous focaliser essentiellement sur les protéines kinases durant cette étude.

I.5) Difficultés actuelles liées à l'identification et à la validation de nouveaux biomarqueurs

Avant de passer à la présentation des concepts fondamentaux relatifs à la détection de protéines en spectrométrie de masse, il est exposé ci-dessous les limitations des approches actuelles concernant la découverte et à la validation de biomarqueurs dans le domaine de la santé.

Une des raisons principales de la difficulté à valider de nouveaux biomarqueurs vient du coût de développement global de la méthodologie. De l'étape de développement à celle de validation, le coût est estimé, en 2007, à 1059 millions d'euros. C'est un chiffre qui n'a fait que grimper puisqu'il n'était que de 138 millions en 1975 et de 802 millions en 2001[65] (Figure 3). Les principales raisons de cette augmentation sont expliquées par une forte augmentation des coûts de recherche et développement qui seraient principalement liés à la complexité technique et scientifique nécessaire croissante (il est par exemple difficile de réunir des cohortes d'échantillons de tailles suffisantes pour procéder à l'étape de validation), aux exigences règlementaires de plus en plus élevées lors des tests cliniques et aux faibles taux de réussite lors de ces études. En effet, la

probabilité de succès de valider officiellement un biomarqueur est de seulement 0,2 % : sur 5000 entités biologiques identifiées et suspectées, seulement 1 se verrait formellement confirmée. Ces 2 observations corrélées peuvent expliquer la frilosité des laboratoires pharmaceutiques à se lancer dans de tels projets.

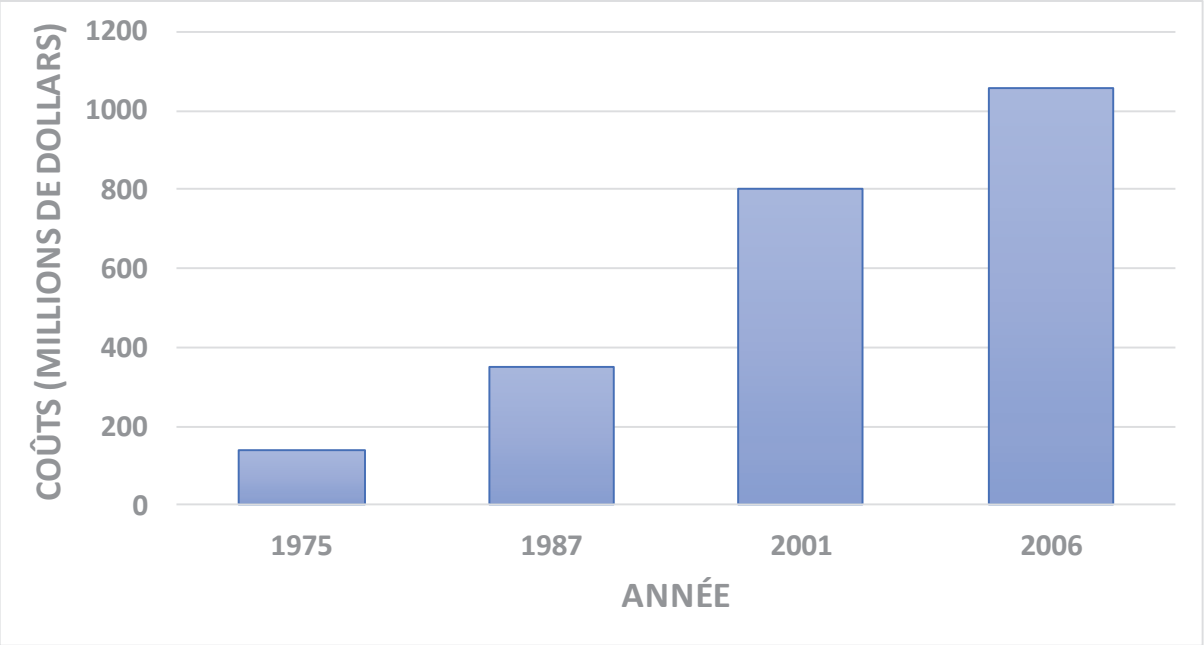


Figure 3 : Evolution du coût financier (en millions de dollars) pour identifier et valider un biomarqueur par un laboratoire pharmaceutique à travers 4 époques différentes[65]

Ces premiers problèmes pourraient être liés au fait que pour valider un biomarqueur une étude statistique mettant en avant des résultats similaires observés sur un nombre significatif de patients est nécessaire. Or, il est très rare qu'un évènement biologique soit observé chez la totalité des patients de l'étude en raison de la variabilité inter-individuelle. En effet, deux individus sains peuvent présenter des concentrations assez variables pour une même protéine brouillant les pistes et compliquant la tâche.

Dans le même temps il faut aussi tenir compte de toutes les incertitudes biologiques liées aux manipulations expérimentales :

- Tout d'abord la qualité des échantillons et leur altération en termes de prélèvement, stockage et de protocole appliqué. La non-uniformisation de ces étapes au niveau des hôpitaux, laboratoires de recherches ou pharmacologiques peut engendrer de la variabilité et nuire à ces étapes d'identification et de validation.

- Ensuite, la grande complexité intrinsèque des fluides biologiques humains. Au niveau de

l'organisme humain, les 20 000 gènes présents vont coder pour plus de 1 000 000 de protéines[5] toutes impliquées dans une multitude de processus reliés particulièrement les uns aux autres. Qui plus est, le rôle d'une protéine peut varier en fonction de sa localisation. Par exemple, la protéine IFI16 stimulera une réponse immunitaire ou sera un détecteur à divers virus en fonction de sa localisation dans la cellule (noyau, cytoplasme). Cette localisation est modulée par une PTM[66]. Ainsi, beaucoup de paramètres sont à prendre en considérations compliquant alors l'approche.

- A cette diversité, il faut finalement ajouter la difficulté liée à la gamme dynamique au niveau des concentrations auxquelles sont exprimées les différentes protéines dans l'organisme. Pour donner un exemple, dans le sérum, l'albumine se situe à une concentration de 1.10^{-3} M, en revanche, la protéine là moins abondante, elle, se situerait aux alentours de 1.10^{-12} M : la gamme dynamique est donc de l'ordre de 1.10^{-9} [66]. Or, en spectrométrie de masse, la gamme dynamique de détection est de l'ordre de 4 ordres de magnitude[67,68] ce qui ne couvre absolument pas celle du sérum humain. Une grosse part du contenu des échantillons est ainsi non analysable expliquant en partie la difficulté à détecter des biomarqueurs faiblement abondants.

II) Principes de l'analyse de protéines en spectrométrie de masse

L'identification des biomarqueurs protéines kinases putatifs sera expérimentée via le couplage chromatographie en phase liquide et spectrométrie de masse (LC/MS/MS). L'application de cette méthodologie à la protéomique s'accomplit principalement selon une approche dite bottom-up après digestion enzymatique. Cette approche est aujourd'hui la plus populaire et sûrement la plus utilisée pour sonder des matrices biologiques[69] (Celle-ci a permis le suivi d'un grand nombre de protéines dans des fluides biologiques pouvant aller au-delà de 4000 protéines en une seule analyse[70]) : elle sera présentée ci-dessous. Une autre voie, appelée top-down, se basant sur l'analyse intacte de protéines par spectrométrie de masse[71] est également accessible. Cependant, en raison des difficultés lors des étapes de fractionnement nécessaires mais encore non optimisées pour séparer efficacement les protéines[72] et à l'observation, pour des protéines à poids moléculaire plus élevé (supérieur à 30 kDa), d'un effet de « dilution du signal » (car la protéine est observée sous de nombreux états de charges différents), cette méthodologie ne sera pas gardée. Il est néanmoins important de préciser que la voie top-down est, à contrario, très efficace pour l'étude structurale de protéines avec des couvertures de séquences exhaustives pouvant aller jusqu'à 100 % permettant ainsi, de caractériser des isoformes résultant de l'épissage alternatif et des PTM[73].

II.1) Préparation d'échantillons lors d'une analyse bottom-up

L'approche bottom-up ne repose pas sur une analyse directe de protéines intactes. Une étape de digestion enzymatique suivie d'une étape de purification des protéines à sonder est à effectuer

avant la séparation et l'analyse de celles-ci sur le couplage LC/MS/MS.

Avant d'aborder l'étape de digestion il est important de rappeler que les protéines sont des biopolymères naturels constitués d'un enchainement linéaire d'acides aminés. Les acides aminés ont tous en commun la formule chimique H_2N-C(R)-COOH et possèdent deux groupements fonctionnels, un groupe carboxyle (COOH) et un groupe amine (NH_2), permettant le couplage peptidique entre les différents monomères (Figure 4). Il existe 21 acides aminés naturels constituants les protéines chez l'homme (Annexe 1). Ils se différencient entre eux par leurs chaines latérales R. C'est cette chaine latérale qui va définir les propriétés de l'acide aminé. Selon leur point isoélectrique, on peut distinguer les acides aminés basiques (Lysine, Arginine, Histidine) des acides aminés acides (acides Aspartique et Glutamique). Les acides aminés basiques ont tendance à capter un proton et sont donc naturellement positivement chargés tandis que les acides aminés acides portent les charges négatives par déprotonation. La charge portée par la protéine dépendra du nombre de résidus acides et basiques qu'elle contient dans sa séquence ainsi que du pH de la solution.

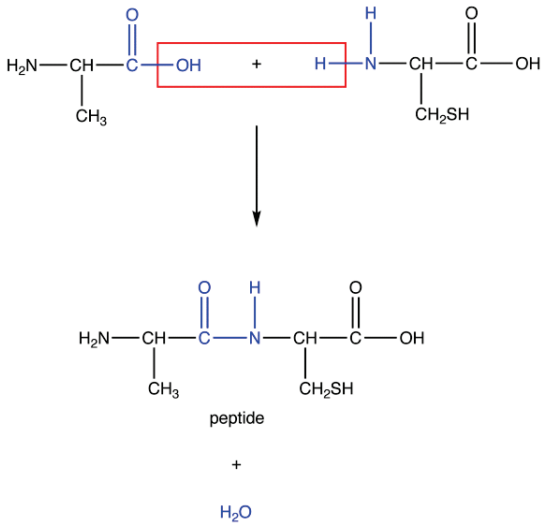


Figure 4 : Liaison peptidique entre 2 acides aminés

Lors d'une étape de digestion enzymatique les protéines sont dénaturées en milieu salin (en vue de les déplier en rompant leur pont disulfure et rendant ainsi accessible les sites de coupures) puis sont découpées par des protéases (enzymes) au niveau d'acides aminés particuliers. Le plus souvent, c'est la trypsine qui est utilisée, découpant les protéines au niveau des acides aminés lysines (K) et arginines (R)[72]. En résulte des polymères d'acides aminés de longueur de séquence moindre appelé peptides. Ceux-ci sont rapporteurs des protéines dont ils sont issus. Cette étape de digestion en vue d'obtenir des peptides est principalement réalisée en raison de la « dilution du signal », de la difficulté de séparation des protéines dans des mélanges complexes comme mentionné plus haut et de la facilité et de l'efficacité à analyser des peptides via un couplage

LC/MS/MS[5]. Une fois digérée, la matrice de peptides résultante sera purifiée et concentrée par extraction en phase solide (SPE) en vue de s'affranchir de composés non désirables en spectrométrie de masse (sels) et de gagner en sensibilité.

Des protocoles de digestions et de purification seront détaillés dans les différentes parties expérimentales présentées dans ce rapport.

Avant de passer en détail au couplage LC/MS/MS, il est important de préciser que cette voie de préparation d'échantillons n'est pas exempte de limitations. Tout d'abord, les enchainements d'acides aminés sont parfois similaires sur certaines portions de plusieurs protéines générant des peptides identiques pour des protéines différentes[74] : une dichotomie est donc appliquée entre des peptides dits protéotypiques c'est à dire unique à la protéine dont ils sont issus par rapport à toutes les protéines connues dans l'organisme étudié et ceux non protéotypiques c'est à dire pouvant être issus de plusieurs protéines au sein de cet organisme. Cela signifie qu'il faut détecter des peptides protéotypiques afin de valider l'identification d'une protéine unique. Il est d'ailleurs possible qu'aucun des peptides protéotypiques attendus, pour une protéine donnée, n'ait une bonne réponse à l'ionisation. Cela implique de multiplier le nombre de peptides suivis par protéines pour corréler leur comportement et s'assurer des bonnes identifications : ce n'est pas toujours possible[75]. Il faut aussi noter que lors de l'analyse de peptides par LC/MS/MS la totalité de la séquence des protéines n'est pas visualisée. Des pans de séquence entiers ne sont pas observés et par là-même de potentiels PTM. Une multitude des formes présentes pour une protéine peut donc être assimilée à un nombre moindre [74]. Le risque est alors de perdre de l'information biologique et de ne pas ou mal comprendre les mécanismes en jeu.

Une fois l'étape de préparation d'échantillons réalisée les peptides sont séparés au cours du temps à l'aide d'une colonne chromatographique le plus souvent en phase liquide (HPLC).

II.2) La séparation chromatographique en phase liquide (HPLC)

Lors de cette étape de séparation, les peptides sont entraînés par un liquide (phase mobile) à travers un solide divisé (phase stationnaire) placé dans un tube (colonne chromatographique). La séparation s'opère suivant les interactions chimiques ou physiques des analytes avec la phase mobile et la phase stationnaire. Suivant les types de phase stationnaire et mobile utilisés, différents modèles de séparation sont atteignables. Par exemple, en chromatographie liquide dite « échangeuse d'ions » les phases stationnaires arborent une résine insoluble munie de groupes fonctionnels capable de se dissocier (SO_3H ou $N(R)_3$)) et il devient possible de séparer les composés en fonction de leur charge. La chromatographie d'exclusion stérique, quant à elle, permet de séparer les composants selon leur dimension moléculaire via l'utilisation d'une taille de pore particulière.

Dans la majorité des cas, lors de la séparation de peptides avant une analyse en spectrométrie de

masse, c'est une séparation dite « en phase inverse » qui est utilisée. Celles-ci se basent sur la séparation des composés en fonction de leur hydrophobicité. La phase stationnaire inverse repose sur une composition de chaînes alkyles allant de 8 à 18 atomes de carbones (dites C_8 ou C_{18} respectivement) donnant le caractère apolaire et hydrophobe à la phase stationnaire. Les peptides sont ensuite séparés en fonction de leur polarité : les plus hydrophobes étant les plus retenus.

A chaque instant, lors de la séparation, tous les composés élués de la colonne chromatographique seront ionisés à l'aide d'une source d'ionisation et seront transférés au sein du spectromètre de masse sous forme d'ions en phase gazeuse où ils pourront être analysés pour identification à l'aide d'un processus de fragmentation[76]. Cette étape est décrite dans la partie suivante. Il est présenté ci-dessous un schéma récapitulatif de l'approche bottom-up décrite précédemment (Figure 5).

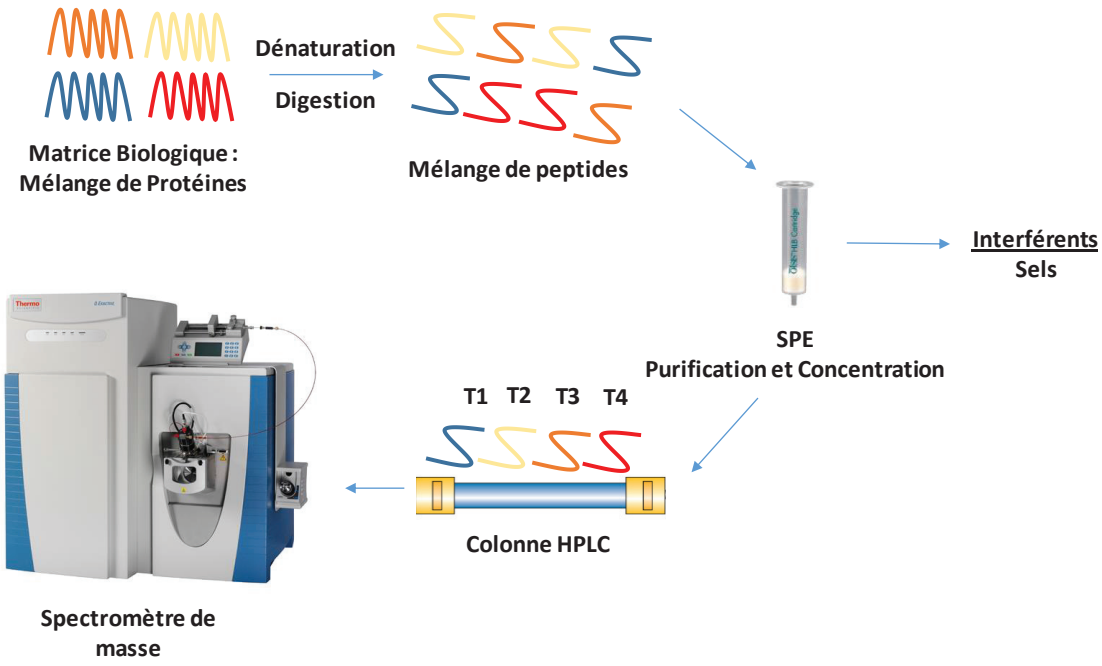


Figure 5 : Schéma récapitulatif de l'approche bottom-up pour l'analyse de protéines en LC/MS/MS

II.3) L'analyse de peptides par spectrométrie de masse

En fonction du spectromètre de masse utilisé différents modes sont accessibles. Ces modes n'offrent pas les mêmes avantages et limitations en fonction de ce que l'on souhaite réaliser : identification, quantification etc. Il faut donc choisir le mode adéquat ainsi que l'appareillage correspondant en vue d'identifier et quantifier les biomarqueurs protéines kinases putatifs au sein de matrices biologiques humaines.

Pour ce faire cette partie sera divisée en 2 sous parties : la première abordant les modes plutôt

relatifs à la quantification et la seconde sera orientée sur ceux en lien avec l'identification.

II.3.a) Méthodes de quantification (ciblées) de peptides par spectrométrie de masse

II.3.a.1) Selected/Multiple Reaction Monitoring (SRM/MRM)

Le mode « Selected Reaction Monitoring » SRM ou « Multiple Reaction Monitoring » MRM est la mode de référence pour quantifier des peptides de manière reproductible et précise au sein de matrices biologiques en spectrométrie de masse[77]. En effet, il a déjà été montré que cette méthodologie permettait de quantifier des dizaines de biomarqueurs putatifs de faible abondance dans le sang[78]. Qui plus est, c'est une méthodologie qui s'est avérée reproductible lors de transferts entre différents laboratoires[79,80] tout en se révélant facile à développer[81]. Les analyses en mode SRM sont effectuées principalement sur des triples quadripôles ou d'appareil dit QTRAP. Ceux-ci se composent de 3 analyseurs quadripôlaires (Q) alignés les uns derrière les autres ou de deux quadripôles suivis d'une trappe linéaire (TRAP) respectivement :



Le mode SRM est un mode ciblé (« targeted ») dans le sens où pour pouvoir effectuer une analyse il faut déjà savoir ce que l'on recherche au sein de l'échantillon analysé. Il faut donc au préalable de l'analyse établir la liste des protéines d'intérêt et les peptides protéotypiques ou non leur correspondant avant de les identifier et de les quantifier dans un quelconque échantillon[82]. Cette information peut être acquise de manière computationnelle, via de l'information extraite par des tests immuno- enzymatiques ou encore par d'autres approches en spectrométrie de masse. A travers ce mode, tous les peptides de la liste seront sondés un par un successivement en chaque point du chromatogramme.

Chaque peptide est sondé à l'aide de plusieurs transitions dites SRM : une transition correspond à deux sélections en masse : la première est celle d'un ion dit précurseur qui correspond au rapport masse sur charge (*m/z*) du peptide sous forme d'ion en phase gazeuse. Cette sélection sera effectuée au sein du quadripôle Q1 et ne laissera passer que les ions correspondants à la *m/z* du peptide d'intérêt (Figure 6 ions verts). Ensuite, les ions transmis au sein du Q2 subiront une étape de fragmentation, le plus souvent par dissociation par collision induite (CID) (les mécanismes de fragmentation de la CID seront décrits dans le chapitre I.III.2). Une seconde étape de sélection de *m/z* est effectuée en Q3 où il ne sera transmis qu'un seul ion fragment caractéristique de l'ion précurseur sélectionné en Q1 (Figure 6).

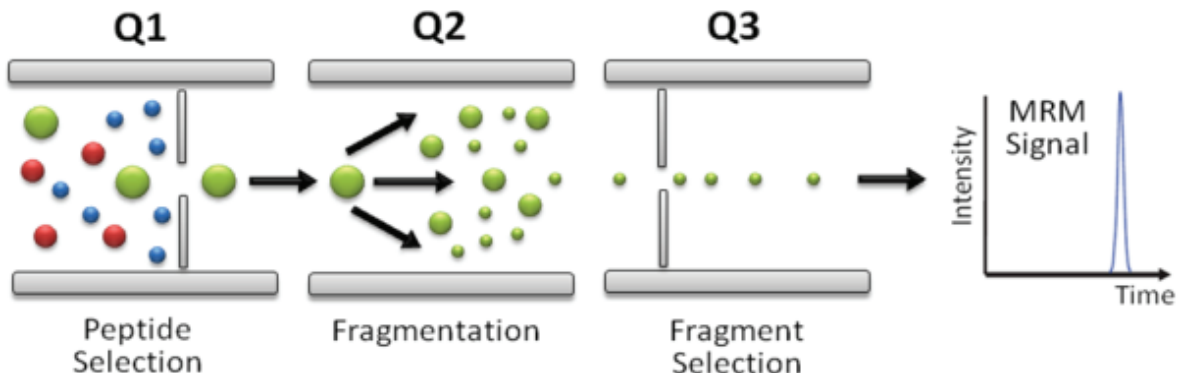


Figure 6 : Fonctionnement du mode SRM sur un triple quadrupole[83]

Lors d'une analyse plusieurs transitions seront suivies par peptides (même *m/z* en Q1 mais différentes *m/z* en Q3) ainsi que pour différents peptides (différentes *m/z* en Q1 et différentes/même *m/z* en Q3). L'étape de quantification (relative ou absolue) se fera par le biais de l'intensité de la transition SRM obtenue.

En chaque point du chromatogramme, toutes les transitions correspondant à chaque peptide seront sondées successivement : de la première à la dernière transition. Une fois la dernière sondée le cycle recommence pour un nouveau point. Le nombre de transitions maximales par analyse dépend de la vitesse de l'analyseur. En SRM, au laboratoire, nous sommes typiquement aux alentours de 120/130 transitions par analyse sur les derniers appareils type 6500 QTRAP.

En général 3 à 5 transitions spécifiques sont suffisantes pour identifier et quantifier des peptides[84] soit une quarantaine de peptides par analyse SRM au laboratoire. Pour maximiser la capacité de multiplexage en SRM, le scheduled SRM a été développé. Celui-ci ne repose plus sur le suivi des transitions SRM tout au long de la séparation mais simplement sur une portion de celle-ci par rapport à un temps de rétention attendu (RT) obtenu via des informations acquises expérimentalement par le passé. Cette approche permet de multiplier le nombre de peptides suivis par analyse : le nombre de protéines identifiables et quantifiables a été estimé à environ 200 protéines par analyse pour 400 peptides suivis soit 2 peptides par protéines au minimum[85] soit 1200 transitions. Cette capacité a été récemment augmentée via le développement d'une technique particulière au laboratoire : la méthode Scout-SRM. Celle-ci a permis le suivi de plus de 445 protéines via 782 peptides suivis en SRM[86] laissant entrevoir de nouvelles possibilités.

La méthode SRM est utilisée pour différents types de dosage dans les fluides biologiques tels que les drogues[87], les métabolites[88] mais aussi les peptides[89–91].

En vue de l'analyse de protéines kinases humaines, le mode SRM peut être une alternative viable de par sa gamme dynamique atteignable (5 ordres de magnitudes[75]) et du pré-ciblage des kinases. Couplé à des techniques d'enrichissement, des kinases ont déjà été identifiées en SRM récemment menant à la quantification de 83 d'entre elles[75]. Ce mode ne sera cependant pas

retenu en raison d'une volonté de se reposer sur une approche exclusivement non ciblée, sans a priori et du fait que le multiplexage en SRM ne semble pas suffisant pour couvrir la totalité du kinome en 1 analyse (518 protéines kinases humaines contre 200 analysables au maximum en scheduled SRM).

II.3.a.2) Parallel Reaction Monitoring (PRM)

Un autre mode « Targeted » similaire à la SRM intitulé « Parallel Reaction Monitoring » (PRM) a été mis en place en 2012[92] en vue de quantifier des peptides et protéines dans des matrices complexes sur des analyseurs à haute résolution types QTOF ou Qexactive (principe décrit en détail dans le chapitre II.I.5). Ces analyseurs sont composés respectivement d'un analyseur quadripôlaire suivi d'un second ou d'une cellule dite de dissociation induite par collision à haute énergie (HCD) suivi d'un analyseur temps de vol (TOF) ou d'un analyseur orbitrap :

Source d'ionisation → Q1 → Q2 (CID) Cellule HCD (HCD) → TOF Orbitrap

Le mode PRM est similaire dans le procédé à la SRM si ce n'est la manière d'identifier les peptides liés à l'analyseur en Q3. Contrairement à la SRM, en PRM il n'y aura aucune sélection des fragments. Tous les fragments du peptide sélectionné en Q1 et fragmenté en Q2 par CID sur les QTOF ou par HCD (la HCD est un processus de fragmentation similaire à la CID comme nous le verrons dans le chapitre I.I.5.c) sur les Qexactive seront transmis à l'analyseur haute résolution en vue d'obtenir un spectre de fragmentation comprenant tous ces fragments (Figure 7).

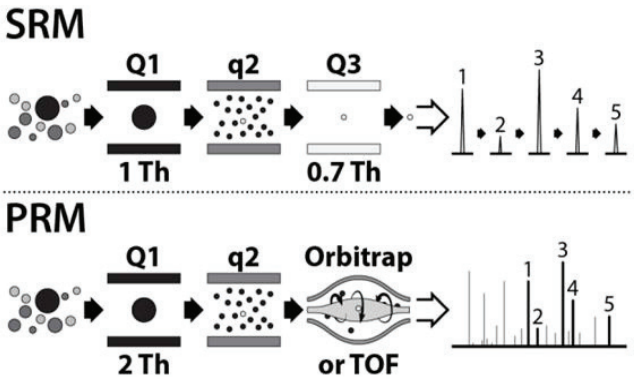


Figure 7 : Comparaison du fonctionnement du mode SRM et du mode PRM[85]

L'étape de quantification (relative ou absolue) se fera par le biais de l'intensité de 1 ou plusieurs des ions fragments issus du spectre de fragmentation de chaque peptide. La spécificité du mode PRM est issu de la haute résolution des analyseurs utilisés à contrario d'une seconde sélection en ion fragment.

Il a été démontré qu'en termes de reproductibilité et de précision la PRM était similaire à la SRM[85]. Cependant, des limites de quantification (LOQ) et des gammes dynamiques plus importantes ont été acquises en PRM[92,93]. Il a aussi été suggéré que via la multitude de fragments acquis lors de l'analyse lié à l'utilisation d'analyseurs hautes résolution, plus d'informations de meilleures qualités sont disponibles pour confirmer la bonne identification des peptides rendant l'approche plus spécifique que la MRM[94]. Cependant, ces avantages sont au détriment d'un temps d'analyse plus long en PRM qu'en MRM en raison de la vitesse plus lente des analyseurs haute résolution[95,96]. Pour répondre à cela, la scheduled PRM a aussi été développé similairement à la scheduled MRM[97].

II.3.b) Méthodes d'identification de peptides (sans a priori) par spectrométrie de masse

II.3.b.1) Data Dependent Acquisition (DDA)

Pour l'identification « à l'aveugle », la méthodologie bottom-up employée en protéomique est la « Data-Dependent Acquisition » (DDA) qui a été mise en place en 2003[98]. L'identification « à l'aveugle » signifie que pour cette analyse aucune liste de protéines n'a été établie au préalable. Ici nous sommes dans l'inconnu et l'on cherche à identifier (et quantifier) le maximum des composés présents au sein d'un échantillon pour ensuite établir un profil. Cette approche intitulée « shotgun proteomics »[70] est principalement appliquée sur des appareils types QTOF et Qexactive :



Pour chaque point du chromatogramme obtenu, l'analyse DDA se divise en 2 parties : tout d'abord un full MS (FMS) est réalisé c'est-à-dire que le quadripôle va filtrer les ions sur une plage de *m/z* étendue en vue d'obtenir toute la distribution en ions précurseurs (et donc en peptides) éluant à cet instant du système chromatographique (Figure 8). Cela se traduit par l'obtention d'un scan où tous les *m/z* correspondants à tous les ions sont présents. Ensuite, un Top N est réalisé. C'est-à-dire que pour une valeur de N = 4, comme sur la Figure 8, les 4 ions précurseurs les plus intenses seront sélectionnés 1 par 1 et successivement fragmentés par CID/HCD au sein du Q2 (l'ion le plus intense sera sélectionné en 1er, le 2nd en deuxième et ainsi de suite). Tous les ions fragments générés dans le Q2 seront ensuite transmis à l'analyseur orbitrap ou TOF pour obtenir un spectre de fragmentation. Une fois les 4 spectres de fragmentation acquis, une exclusion dynamique est opérée. Les 4 ions sélectionnés la première fois seront exclus pendant quelques secondes en vue d'aller sonder plus profondément l'échantillon et les peptides moins intenses éluant à cet instant.

Ce cycle global est réitéré jusqu'à la fin de l'analyse. Le nombre d'ions précurseurs qu'il est possible de sélectionner (N) va dépendre de la vitesse de l'analyseur : plus il sera rapide plus il sera possible d'augmenter la valeur de N et donc d'augmenter l'information collectée.

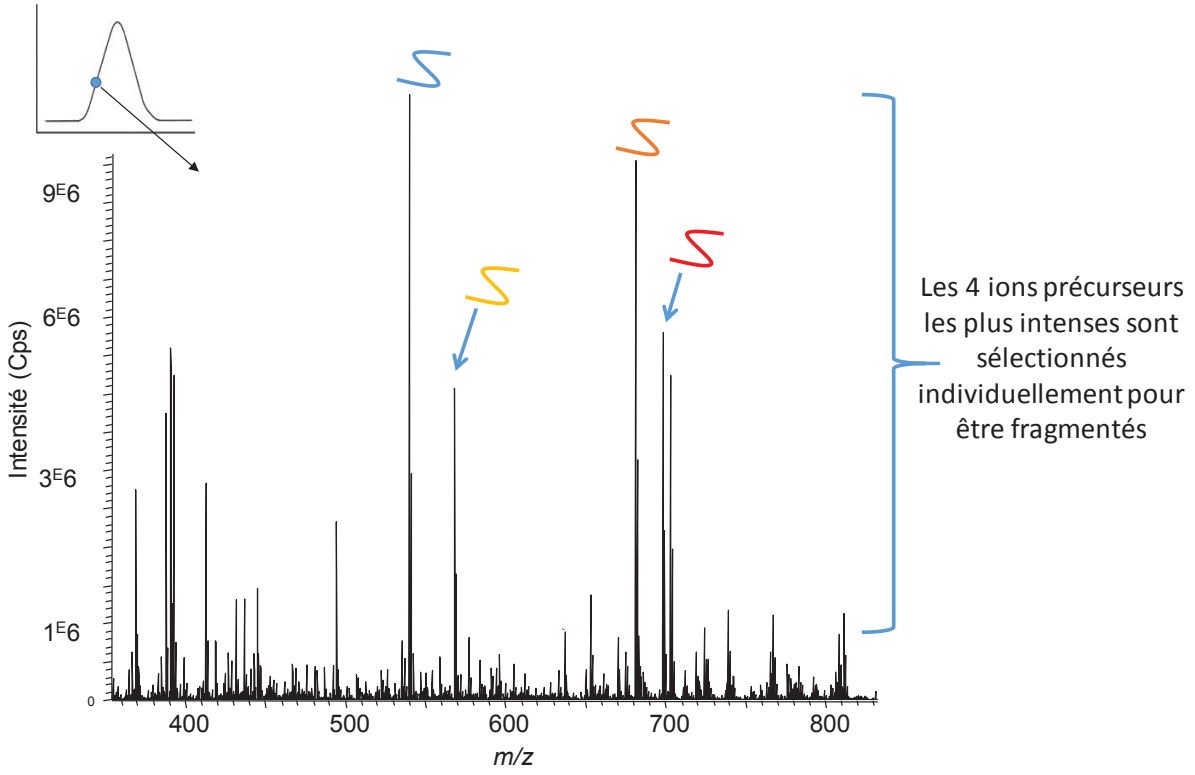


Figure 8 : Fonctionnement du mode DDA : 1 scan FMS est acquis suivi d'une sélection individuelle successive de N ions précurseurs en vue d'une étape de fragmentation par CID/HCD

Cette étape de fragmentation est effectuée car les spectres de fragmentation obtenus vont permettre d'identifier les peptides fragmentés. En effet, chaque peptide va générer un spectre de fragmentation qui lui est propre ce qui permettra de l'identifier. Pour ce faire, ceux-ci sont comparés à des spectres fragmentation théorique stockés dans des bases de données[97] (Figure 9) car il est possible de prédire la fragmentation des peptides par CID en fonction de leur séquence.

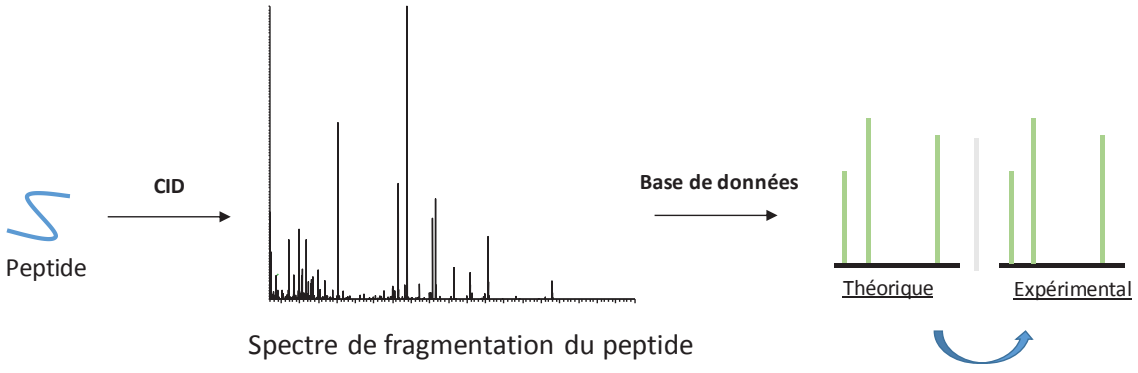


Figure 9 : Procédé utilisé pour identifier des peptides en mode DDA

A l'aide d'algorithme, il sera établi des corrélations entre les spectres expérimentaux et ceux théoriques. Voici une liste non exhaustive des logiciels traitant ce type de données : SEQUEST[99], Mascot[100], Maxquant[101]. L'identification peut aussi se faire à l'aide de bibliothèques spectrales[97].

Les analyses DDA ont permis la couverture de plusieurs milliers de protéines au sein d'échantillons biologiques : plus de 4000 protéines ont été identifiées au sein d'extrait cellulaire humain[70] et plus de 3900 protéines dans des extraits de levures[102], par exemple. En effet, beaucoup d'appareils peuvent actuellement acquérir 10 spectres MS^2 en approximativement 2 s[77] ce qui laisse entrevoir une capacité de multiplexage très importante en vue d'identifier des biomarqueurs ou de suivre une multitude de processus biologiques en simultané. La DDA s'est d'ailleurs révélée efficace pour la quantification relative de peptides à travers différents échantillons[103,104] et a été utilisée avec succès dans diverses études : la caractérisation d'une partie des réseaux d'activité de la protéine EGFR[105], la caractérisation du protéome sur différents organes de souris[106], les interactions protéines-protéines[107] ou encore en virologie[108].

Cependant, cette technique fait face à certaines limitations : tout d'abord, malgré la rapidité de l'analyse en DDA, tous les peptides présents au sein d'un spectre FMS ne pourront être sélectionnés en vue d'être fragmentés et identifiés. Qui plus est, la sélection commençant par les ions les plus intenses en évoluant de manière décroissante, ce sont les peptides les moins abondants qui seront omis[109]. Or, il semble que les biomarqueurs putatifs les plus informatifs soient de faibles abondances[85] ce qui nuit fortement à son potentiel dans cette optique. En outre, il faut préciser que la sélection en précurseur est effectuée de manière stochastique induisant une mauvaise répétabilité et/ou reproductibilité[77]. En effet, d'une part si beaucoup de peptides avec des intensités proches co-éluent, alors, en raison de la variabilité intrinsèque du système, ce ne seront pas forcément les mêmes qui seront sélectionnés à chaque analyse au même instant. Le recouvrement de similarité à l'identification d'une analyse à l'autre a été établi à 60 %[110]. D'autre part, les peptides ne sont échantillonnés qu'une fois ou deux lors de l'analyse en raison de l'exclusion dynamique et ne sont ainsi pas forcément sélectionnés au niveau de l'apex. Cela peut induire des CV insuffisants ou des spectres de fragmentation de mauvaise qualité[77].

Malgré le fort multiplexage offert par cette technique, il semble que le sous-échantillonnage, inhérent, soit un facteur limitant pour notre approche. En effet, 261 kinases humaines ont ainsi pu être identifiées au sein d'extrait cellulaire[59] et 212 quantifiés dans 6 extraits cellulaires cancéreux différents[57] à l'aide de la DDA en utilisant diverses techniques d'enrichissement. Il semble possible que là aussi les limites de la méthodologie en rapport à ce projet aient déjà été atteintes.

II.3.b.2) Data Independent Acquisition (DIA)

Une autre méthodologie bottom-up a été développée en 2003 en vue de détecter des peptides en spectrométrie de masse : la « Data Independent Acquisition » DIA[111]. Celle-ci est de plus en plus utilisée au niveau de la recherche clinique et fondamentale[112]. Son développement s'est

orienté dans le but de combiner les avantages offerts par les méthodes de types shotgun (haut multiplexage) et ciblé (bonne reproductibilité et précision)[112,113]. Comme en DDA, la DIA est principalement appliquée sur des appareils de types Qexactive et QTOF :



Son fonctionnement est le suivant :

Pour chaque point du chromatogramme obtenu, l'analyse DIA peut se diviser en 2 parties : un FMS est réalisé suivi d'une étape de fragmentation (toujours en CID/HCD) comme en mode DDA. Lors de cette seconde étape, c'est la plage de *m/z* transmise qui est déterminante. Ici, contrairement à la SRM et à la DDA, on ne sélectionne pas les ions précurseurs un par un. Le spectre de masse est découpé en fenêtres de *m/z* plus ou moins larges de X Th (selon les choix de l'analyste) (Figure 10) que l'analyseur quadripôlaire va filtrer successivement en vue de faire (co)-fragmenter tous les ions contenus dans chacune et d'obtenir des spectres de (co)-fragmentation.

Pour prendre l'exemple ci-dessous, la gamme de *m/z* analysée va de 350 à 850 Th. Le spectre est découpé en fenêtre de 25 Th. Pour un point du chromatogramme : le FMS est acquis entre 350 et 850 Th puis tous les ions (et donc peptides) contenus dans la fenêtre allant de 350 à 375 Th sont sélectionnés ensemble et vont aller (co)-fragmenter ensemble dans le Q2 puis seront transmis au sein de l'analyseur pour obtenir un spectre de (co)-fragmentation. Dans le spectre de (co)-fragmentation les ions fragments de tous les peptides contenus dans la plage de *m/z* filtrée seront mélangés ensemble. Ensuite, il en sera fait de même pour les ions contenus dans la plage 375 - 400 Th, pour celle de 425 - 450 Th et ainsi de suite jusqu'à 825 – 850 Th. Arrivé à ce stade, pour un point du chromatogramme 1 FMS et 20 spectres de (co)-fragmentation auront été obtenus. L'opération est ensuite réitérée tout au long de l'analyse.

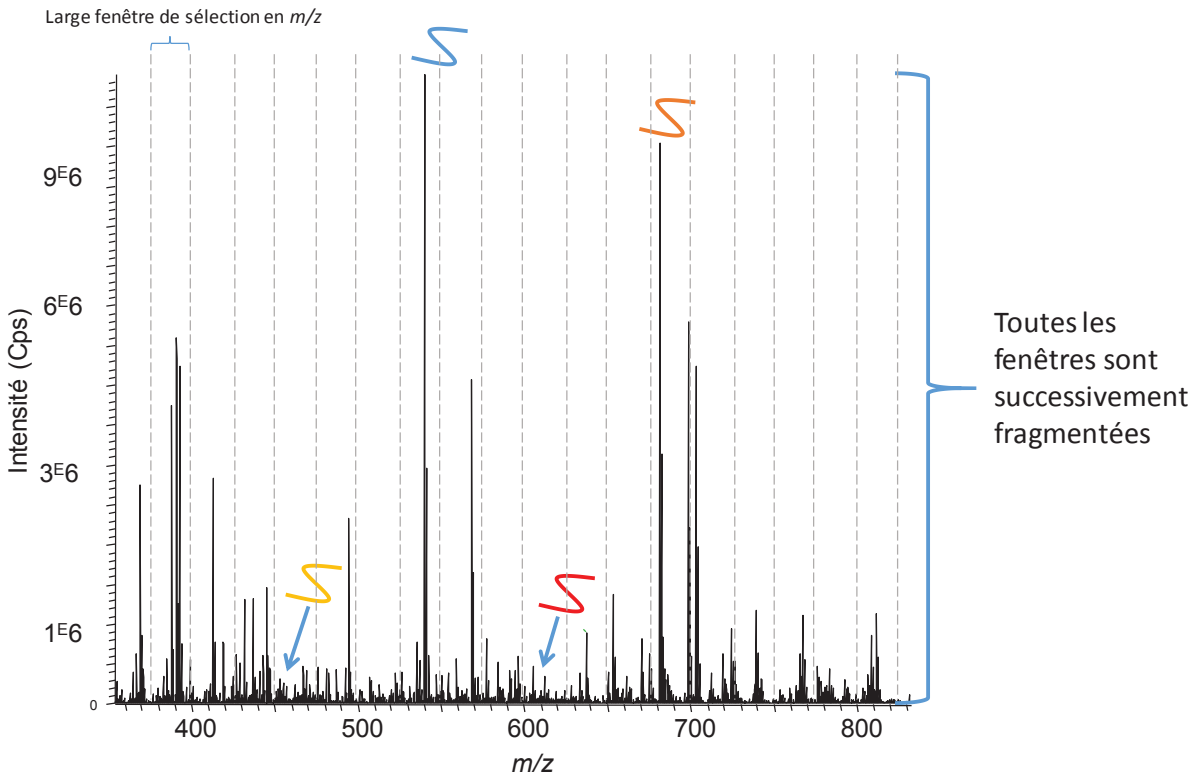


Figure 10 : Fonctionnement du mode DIA : 1 scan FMS est acquis suivi d'une fragmentation de la totalité de la plage de *m/z* par l'intermédiaire de larges fenêtres successives de plusieurs Th

L'étape d'identification se base sur les spectres de (co)-fragmentation obtenus et peut se faire via des bases de données théoriques comme en DDA (DIA-Umpire[114] ou DeMux[115] par exemple). Cependant, ces algorithmes ne se sont pas révélés efficaces pour traiter ces spectres de (co)-fragmentation au niveau des composés de faibles abondances[116]. Pour remédier à cela, une approche a été mise en place par Gillet *et al* en 2012[117]. Celle-ci se base sur l'utilisation de librairie spectrale contenant des spectres de fragmentation de peptides expérimentalement acquis par le passé. Ces spectres ont été obtenus soit par l'utilisation de peptides synthétiques soit par identification direct au sein de fluide biologiques via d'autres modes (DDA par exemple). Toutes les informations extraites telles que le *m/z* des fragments, l'intensité relative des fragments entre eux, seront compilées pour chaque peptide dans la bibliothèque. Puis chaque spectre de (co)-fragmentation sera évalué en regard des spectres de la bibliothèque pour identifier les peptides analysés[116]. Différents logiciels permettent cette approche tels que OpenSWATH[118], Spectronaut[119] ou encore Skyline[120].

Différentes manières d'effectuer de la DIA ont été développées au cours du temps. La première reposait sur de la fragmentation en source couplée à un TOF[111]. Le concept de MS^E a ensuite été proposé en 2005 sur des appareils QTOF par Waters reposant sur des alternances haute/basse en énergie de collision (CE) sur des plages de masses conséquentes (300-2000 Th) permettant

d'alterner l'acquisition d'un FMS et d'un spectre de (co)-fragmentation pour la plage entière. Cela s'est révélé efficace pour quantifier des protéines de manières relative et absolue[121–123]. Un concept similaire a été développé en 2010 sous le terme « All Ion Fragmentation » (AIF) pour les appareils Qexactive[124]. En 2004, une autre voie a été empruntée basée non pas sur les alternances en CE mais sur le découpage de la plage de masse comme explicité au-dessus en fonction de la vitesse de l'analyseur[109]. Cela a été effectué dans un piège à ion linéaire (LIT) avec des fenêtres de 10 Th. Ce concept a été ensuite approfondi menant à la « Fourier Transform All Reaction Monitoring » (FT-ARM) usant de 100 Th sur des couplages LIT-Orbitrap[125] ou encore la méthode « Sequential Window Acquisition of all Theoretical Spectra » (SWATH) développée en 2012 sur un Triple TOF 5600 avec des fenêtres de 25 Th et une haute résolution (précision en masse inférieure à 10 ppm). En 2013, la « Multiplexed data independent acquisition » (MSX) a été introduite sur un appareil Qexactive[126]. Dans cette approche, le spectre FMS est découpé en une centaine de fenêtres de 4 Th qui sont combinées aléatoirement par 5 en vue d'augmenter la sélectivité en précurseur. Cela a pour but de minimiser la (co)-fragmentation d'ions précurseurs pouvant générer des fragments similaires (peptide sans et avec PTM par exemple). Des améliorations de ces concepts ont été développées par la suite telle que la p-SMART[127].

Au niveau des applications en protéomique, la DIA peut être utilisée pour identifier et quantifier des peptides. Couplée à des librairies spectrales, elle a par le passé montré son potentiel quant à sa capacité de multiplexage et de répétabilité en quantifiant par exemple 1447 peptides (342 protéines) avec des coefficients de variation (CV) inférieurs à 25 % à travers 232 échantillons de plasma humains[128]. Une comparaison avec la DDA a pu aussi être effectuée lors d'une étude basée sur l'interaction entre des protéines. 5089 peptides ont pu être identifiés contre 2741 en DDA en triplicats. En DIA, 82,1 % des peptides identifiés avaient des CV < 20 % contre 74,5 % en DDA montrant alors l'apport en multiplexage et en répétabilité offert par la DIA[129].

L'approche DIA répond en effet à certaines limites inhérentes à la DDA. Tout d'abord, l'affranchissement de la sélection en précurseur. Ici, tous les ions potentiellement détectables par l'appareillage seront fragmentés et donc par là même identifiables[130]. La gamme dynamique de l'appareillage est exploitée à son maximum : tout ce qui peut être détecté le sera. Ensuite, de par son acquisition non stochastique, les problèmes liés à effectuer la fragmentation à l'apex du pic chromatographique ne sont plus et la répétabilité obtenue est meilleure que celle en DDA[109]. Il a même été établi que la reproductibilité dans l'identification et la précision dans la quantification des peptides pouvaient être comparable à la MRM mais que la méthode était moins robuste en raison des interférences générées par la (co)-fragmentation[131,132]. Il a d'ailleurs été montré que la méthode SWATH permettait de quantifier des peptides trypsiques avec une gamme dynamique de 4 ordres de magnitudes[117].

A contrario, de cette approche nait d'autres limitations : tout d'abord la taille de la libraire spectrale : on ne pourra identifier que les peptides présents au sein de la librairie. Cela implique

d'enrichir au maximum la libraire spectrale en fonction des processus biologiques que l'on veut suivre. Ce n'est pas toujours évident[133] même s'il est envisagé la construction d'une bibliothèque spectrale couvrant le protéome humain entier[112]. Le regard n'est donc pas encore exhaustif mais orienté. Ensuite la complexité des spectres de (co)-fragmentation rendant délicate l'extraction et l'interprétation des résultats[77] (Figure 11) :

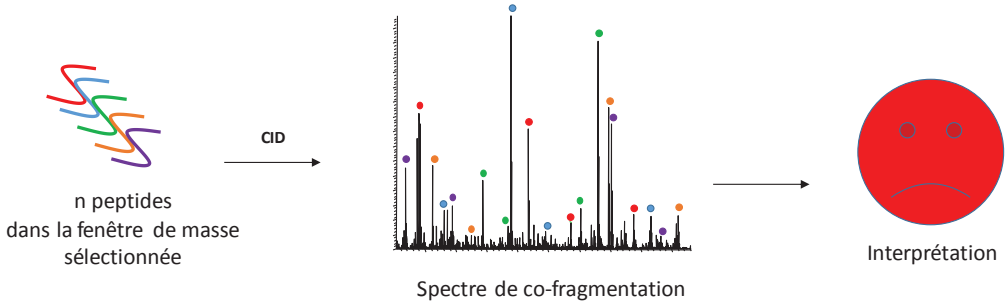


Figure 11 : Complexité des spectres de (co)-fragmentation obtenu en DIA

En effet, tous les peptides présents dans chaque fenêtre vont fragmenter ensemble pour obtenir un spectre de (co)-fragmentation où tous leurs fragments seront mélangés. La réattribution devient délicate car des peptides via leurs séquences peuvent générer les mêmes fragments et pose la question de leur réattribution et en quelle proportion[130]. La question du logiciel ou de la manière de traiter les données est aussi limitante en DIA car aucun consensus n'a encore été adopté quant au traitement adéquat[134]. Cette non efficience dans l'identification est critique pour développer cette approche dans des projets cliniques divers dont l'identification de biomarqueurs[135].

La manière d'y remédier simplement serait de réduire la taille des fenêtres sondées dans les approches bottom-up. Cependant, les analyseurs n'ont pas les vitesses requises à l'heure actuelle pour effectuer cette réduction. Il faudrait à contrario diminuer la taille de la fenêtre FMS dès le départ ce qui implique une perte d'information et, sachant que les plages de masse couverte en DIA sont de 400 Th classiquement contre 1600 Th en DDA, cela ne semble pas souhaitable[77]. De ce fait, outre l'attente liée aux développements technologiques et instrumentaux, d'autres approches peuvent être intentées en vue de réduire la complexité des échantillons sondés telle que l'utilisation des méthodologies de fractionnement ou d'enrichissement. Par exemple, l'utilisation de résines spécifiques aux glycopeptides a permis un meilleur suivi des glycoprotéines lors d'une étude de biomarqueurs putatifs ciblant le cancer de la prostate[136] ou encore l'obtention de LOQ de l'ordre de 5 ng/mL et un nombre de protéine unique couvert important grâce à l'utilisation de kit de déplétion[128]. Dans le même temps, une autre approche se basant l'utilisation d'une méthode d'activation des ions plus spécifique lors du processus de fragmentation peut être proposée. En effet, l'activation par CID n'est pas spécifique : tous les ions entrant dans la cellule de collision vont fragmenter. Or, si une méthode plus spécifique est utilisée, seule une population

d'ions fragmentera, permettant de simplifier les spectres de (co)-fragmentation et de facilité l' interprétation[9].

Cette dernière approche a été suggérée par l'équipe dans laquelle j'effectue ma thèse et sera celle qui sera retenue pour la mise en place d'une méthodologie d'identification des biomarqueurs protéines kinases putatifs au sein de fluides biologiques. Cette voie semble pouvoir utiliser tous les avantages offerts par la DIA (multiplexage, reproductible et précise) tout en ayant le potentiel pour s'affranchir de ses limitations (complexité des spectres de (co)-fragmentation). Qui plus est, aucune approche n'a encore été tentée dans cette direction, relativement à la simplification des spectres de (co)-fragmentation obtenue en DIA, ce qui semble être le bon moment pour la développer. Pour terminer, il ne semble qu'aucune étude n'a fait mention d'une tentative d'une couverture du kinome dans des extraits cellulaires humains en utilisant la DIA (seulement une étude quant au suivi de la phosphorylation de diverses protéines[137] alors que les approches DDA et SRM semblent avoir atteint leur maximum à ce niveau en cet instant).

Dans la partie suivante, il sera présenté les principales méthodes d'activation utilisées en spectrométrie de masse en vue de choisir la plus adaptée à notre approche pour identifier des biomarqueurs protéines kinases putatifs en mode DIA avec une bilbiothèque de spectres qui leur sera spécifique.

III) Différentes méthodes d'activation des ions en phase gazeuse pour l'analyse de peptides et de protéines en spectrométrie de masse

En vue de simplifier les spectres de (co)-fragmentation obtenus en DIA et de tenter de mettre en place une méthodologie performante pour l'identification de nouveaux biomarqueurs, nous nous sommes penchés sur les différents méthodes d'activation des ions en phase gazeuse. En effet, nous avons constaté que l'identification se fait grâce à l'étape de fragmentation et qu'une spécificité de fragmentation peut être atteinte selon la méthode d'activation utilisée.

A l'heure actuelle, différentes méthodes de fragmentation spécifiques ou non sont accessibles, chacune générant des modèles de fragmentation caractéristiques et donc une information particulière sur les composés étudiés. Ici, il sera présenté la fragmentation des peptides par spectrométrie de masse ainsi que les principales méthodologies (avantages, limitation, portée etc..) appliquées pour dissocier ces peptides.

III.1) Nomenclature des ions fragments résultant de la fragmentation des peptides par spectrométrie de masse

En ce qui concerne la fragmentation des peptides par spectrométrie de masse, une nomenclature universelle a été proposée par Roepstorff et Fohlman en 1984[138], nomenclature qui sera modifiée

en 1988 par Biemann[139]. Celle-ci est présentée ci-dessous (Figure 12) :

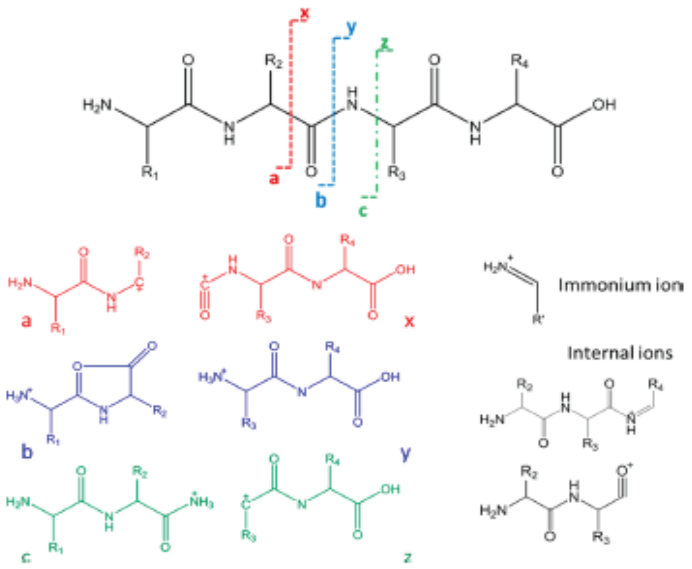


Figure 12 : Nomenclature des ions fragments issus de peptides[69]

La séquence d'un peptide est donnée de l'extrémité N-terminale (fonction amine libre) à l'extrémité C-terminale (fonction acide carboxylique libre). La fragmentation du peptide peut intervenir soit au niveau du squelette peptidique soit au niveau des chaines latérales des acides aminés composant le peptide. La nomenclature des fragments générés, utilisant des lettres de l'alphabet, va dépendre de 3 paramètres : l'endroit de la rupture, la localisation de la charge et le nombre d'acides aminés restants. Tout d'abord, les fragments chargés qui contiennent la partie N-terminale sont intitulés : a, b et c tandis que ceux chargés qui contiennent la partie C-terminale sont nommés : x, y et z. Ensuite en fonction du lieu de la rupture, l'une des trois lettres leur est attribuée : a/x pour une rupture au niveau de la liaison C_α-C, b/y pour une rupture au niveau de la liaison C-N et c/z pour une rupture au niveau de la liaison N-C_α. Puis une numérotation leur est donnée en relation avec le nombre d'acides aminés restants dans le fragment : s'il reste deux acides aminés dans le fragment il est noté 2 si trois 3 etc...

A cela il faut ajouter des ions appelés internes qui sont issus soit de la fragmentation des premiers ions fragments générés (a, b, c, x, y, z) durant laquelle les parties C-terminales ou N-terminales sont perdues, soit d'une rupture multiple au niveau de la séquence peptidique initiale. Des ions immonium apparaissent également lorsqu'une fragmentation de type a est proche d'une fragmentation de type y. Ils proviennent d'acides aminés uniques et possèdent donc de faible masse (inférieure à 200 Th). Ils peuvent aider à connaitre la composition en acide aminé du peptide fragmenté. De plus, des ions fragments issus de la fragmentation secondaire d'une partie ou de la totalité des chaines latérales des ions a, y et z sont décrits :

- Des ions d formés par la perte d'une partie de la chaine latérale des ions a.

- Des ions v formés par la perte de la chaine des ions y.

- Des ions w formés par la perte de la chaine latérale des ions z.

A noter que la nomenclature proposée ici a été développée pour le mode positif. En mode négatif, il n'en existe pas, c'est donc celle du mode positif qui est actuellement utilisée[140,141].

III.2) La dissociation induite par collision (CID)

La méthode de fragmentation CID est aujourd'hui la méthode standard de fragmentation utilisée en spectrométrie de masse pour identifier des peptides lors des analyses bottom-up. C'est celle qui est classiquement utilisée en DIA comme nous avons pu le voir précédemment. Celle-ci a été implémentée dans la majorité des instruments disponibles sur le marché (QTRAP, triple quadripôle, Q-TOF). Ceci, principalement en raison de ses performances robustes, ainsi qu'à la bonne compréhension de ses mécanismes de dissociation sur les peptides permettant de remonter à leur structure et de les identifier[142,143]. Son temps d'activation faible (environ 30 ms ou moins[144]) générant des temps de cycle court et permettant d'analyser de manière exhaustive les fluides biologiques est aussi une des raisons quant à son choix[145].

Brièvement l'activation par CID peut être réalisée de deux manières :

- La première approche repose sur l'accélération des ions en phase gazeuse lors de leur entrée au sein d'une cellule de collision. Un gaz inerte est présent au sein de celle-ci (hélium, diazote, argon). Les ions vont alors entrer en collision avec celui-ci engendrant une augmentation de leur énergie interne (par le biais d'un processus de conversion de leur énergie cinétique en énergie interne) qui va mener à leur dissociation[146] (Figure 13) :

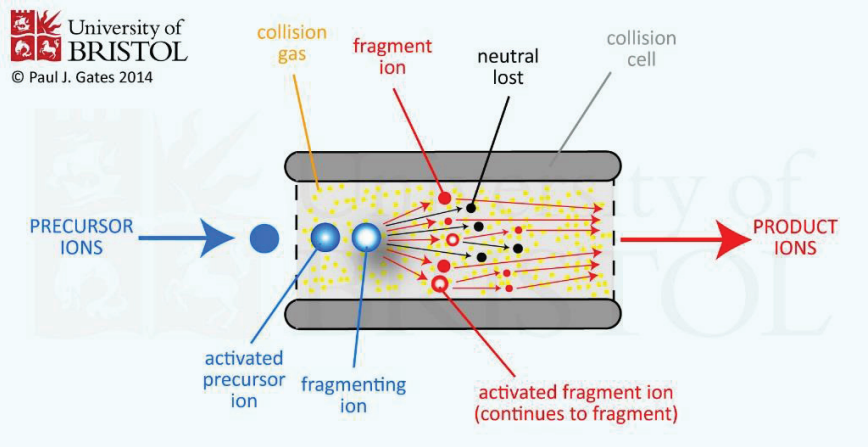


Figure 13 : Activation par CID des ions en phase gazeuse[147]

- La seconde approche a pour but d'utiliser un piège à ions rempli de gaz inerte où les ions sont excités par radiofréquences. Suite à cette excitation, on assiste à de nombreuses collisions avec le gaz inerte augmentant l'énergie des ions ce qui mène également à leur dissociation.

Les mécanismes de la fragmentation des peptides par CID ont été étudiés et une explication a été proposée à travers le modèle du proton mobile[148,149]. Brièvement, celle-ci stipule que les protons se situent sur les différents sites basiques présents sur le peptide : l'extrémité N-terminale et les chaines latérales des lysines, arginines et histidines. Suite à l'activation collisionnel, ceux-ci vont migrer le long du squelette peptidique engendrant des fragmentations par le biais de mécanismes particuliers tels que la protonation des groupements amides induisant une fragmentation à leur niveau pour produire des ions b et y (une rupture au niveau des liaisons C-N de la séquence peptidique). Ce sont ces ions qui sont majoritairement en CID sur des peptides.

Il est important de noter que lors de ce processus, l'ion doit entrer en collision avec de multiples atomes de gaz neutre avant qu'il ne puisse se fragmenter impliquant une limitation quant à l'énergie cinétique maximale qu'il est possible de convertir pour que l'ion se dissocie : cette technique favorise la fragmentation à des canaux de basse énergie. Cela a pour effet d'engendrer des ruptures aux niveaux des liaisons les plus labiles des peptides causant des pertes de H_2O, NH_3 peu informatives au détriment de fragments de la chaine peptidique permettant l'identification des peptides étudiés[142,143]. En outre, les PTM ne peuvent être gardées et, dans ce cas-ci, ce sont des informations pertinentes sur la régulation des protéines qui sont perdues. D'autre part, avec la technique CID, un « low-mass cut-off » est observé dans les trappes ioniques en raison des radiofréquences utilisées au cours du piégeage des ions lors de l'activation du processus de fragmentation[69]. La troncation des faibles *m/z* empêche l'observation des ions immoniums ainsi que des premiers ions b et y pouvant informer sur la composition des peptides en N-terminale et C-terminale. Cela a été corrigé via le développement des cellules HCD[150] qui fonctionnent aussi à de faible temps d'activation (environ 5 ms[151]).

Mais plus important pour nous encore, cette méthode de fragmentation n'est pas spécifique comme précisé auparavant, tous les ions entrant dans la cellule de collision étant fragmentés.

III.3) L'activation électronique

Les principales méthodes d'activation électronique utilisées pour dissocier les peptides sont intitulées « Electron Capture Dissociation » (ECD) et « Electron Transfer Dissociation » (ETD). D'autres méthodologies ont été mises en place ensuite (Electron Detachment Dissociation (EDD)[152], Negative Electron Transfer Dissociation (NETD)[153] etc…) mais seules ces 2 principales seront présentées.

- L'ECD repose sur l'irradiation des ions au sein d'un spectromètre de masse par un faisceau d'électrons de faible énergie (1-10 eV). En effet dans cette gamme d'eV, les peptides multichargés

ont une section efficace d'absorption plus importante pour des électrons de haute énergie[154]. Durant ce processus, les électrons sont capturés par les ions via un processus exothermique engendrant leur fragmentation[154–156] :

$$(M+nH)^{n+} + e^{-}\ (1\text{-}10eV) \xrightarrow{\text{Capture } e^-} (M+nH)^{n-1+\bullet} \xrightarrow[\text{de liaison}]{\text{Rupture}} \text{Fragments}$$

Cette méthode de fragmentation est souvent réalisée au sein d'appareil de « Fourier Transform Ion Cyclotron Resonance » (FTICR) où les ions chargés positivement et les électrons peuvent être piégés en même temps. Si cette application s'est révélée efficace pour les FTICR celle-ci n'a pu être réalisée sur les instruments comportant une trappe à ions en raison d'une incompatibilité entre les radiofréquences appliquées et les conditions nécessaires pour piéger ces électrons faibles énergie[69]. Une alternative a donc été mise en place :

- L'ETD a été développé en vue d'être implémenté sur les analyseurs comportant une trappe ionique[153,157] (Figure 14) :

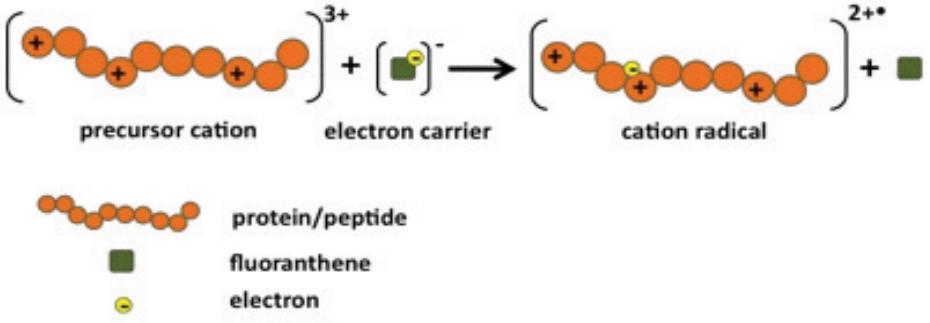


Figure 14 : Principe de l'activation ETD[158]

Ici, un ion chargé positivement est envoyé au sein d'un piège à ions où un anion est présent (typiquement du fluoranthène ou de l'anthracène). Lorsque les deux composés se trouvent à une distance suffisante, il y a un transfert d'un électron de l'anion vers l'ion chargé positivement pour induire sa fragmentation. Celle-ci va dépendre essentiellement de la structure de l'anion qui va transmettre l'électron ainsi que du temps de réaction entre les deux ions dans le piège.

Le processus de fragmentation engendrée via l'activation électronique a déjà été étudié[159,160]. Brièvement, la capture d'un électron engendre la formation d'une espèce radicalaire et une redistribution d'énergie au niveau vibrationnelle provoque une rupture au niveau de la liaison N-C_α du squelette peptidique. Les fragments produits sont principalement des ions fragments c et $z^{\cdot}$ (Figure 15). Ces méthodes d'activation électronique se révèlent donc réellement complémentaires de la CID (fragments b et y) pour l'étude de la structure des peptides analysés. En plus de cela, il faut préciser que l'ETD/ECD ne génère pas la perte des groupements les plus labiles type PTM. Elles peuvent donc contribuer à la compréhension de la régulation des protéines

et des processus en cours au sein de divers systèmes biologiques.

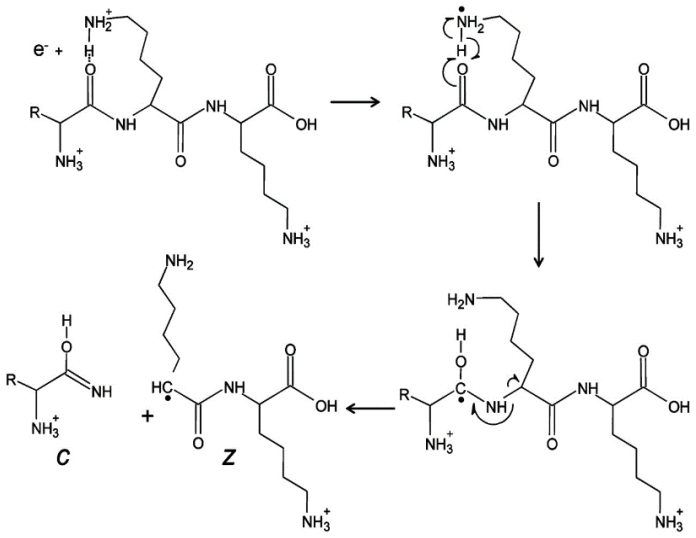


Figure 15 : Ions principalement produits lors de l'activation de peptides par ETD/ECD[157]

D'un point de vue plus expérimental, il faut préciser que la fragmentation issue de ces deux méthodologies dépend beaucoup de l'état de charge de l'ion précurseur. Il s'avère que les ions possédant des états de charges élevés (4^+, 5^+, 6^+) seront mieux fragmentés que ceux à des états de charges inférieurs tendant plutôt à former des radicaux (ou des neutres pour les monochargés) qu'à se dissocier. Cela a été démontré en ECD où la section efficace d'absorption augmente relativement à la charge de l'ion précurseur[154]. Pour remédier à cela, différentes voies ont été empruntées : certaines ont couplées à ces méthodes de la CID ou de la photodissociation laser sous le terme activated ion ECD/ETD (a – ECD/ETD) en vue d'apporter de l'énergie supplémentaire et de favoriser la fragmentation des ions[161,162]. D'autres approches se sont basées sur l'optimisation de la composition des solvants et des paramètres de sources, d'autres sur l'ajout de cations métalliques ou la modification du peptide avec un tag synthétique chargé[163–165].

En ECD/ETD il est aussi important d'aborder les notions de rendement de fragmentation et de temps d'activation. Tout d'abord le rendement de fragmentation (η) peut être exprimé selon cette équation[155] (Equation 1):

$$\eta = \frac{\textit{Abondance totale des ions fragments}}{\textit{Abondance totale des ions précurseurs}} \qquad \textbf{Equation 1}$$

Ce rendement a été estimé en moyenne à 20 % pour les cations dischargés mais à plus de 80 % pour les peptides hautement chargés en ECD avec un temps d'activation estimé à 50-70 ms même s'il peut baisser aux alentours de 20 ms dans le cas de protéines chargées[166–168]. Il en va de même pour l'ETD avec des valeurs légèrement supérieures en ce qui concerne le temps d'activation (50-200 ms pour les peptides et 40 ms pour les protéines)[169–171]. Ceci est illustré par les spectres de

fragmentation ECD et ETD du peptide NEREEHAMR (Figure 16) doublement et triplement protoné respectivement :

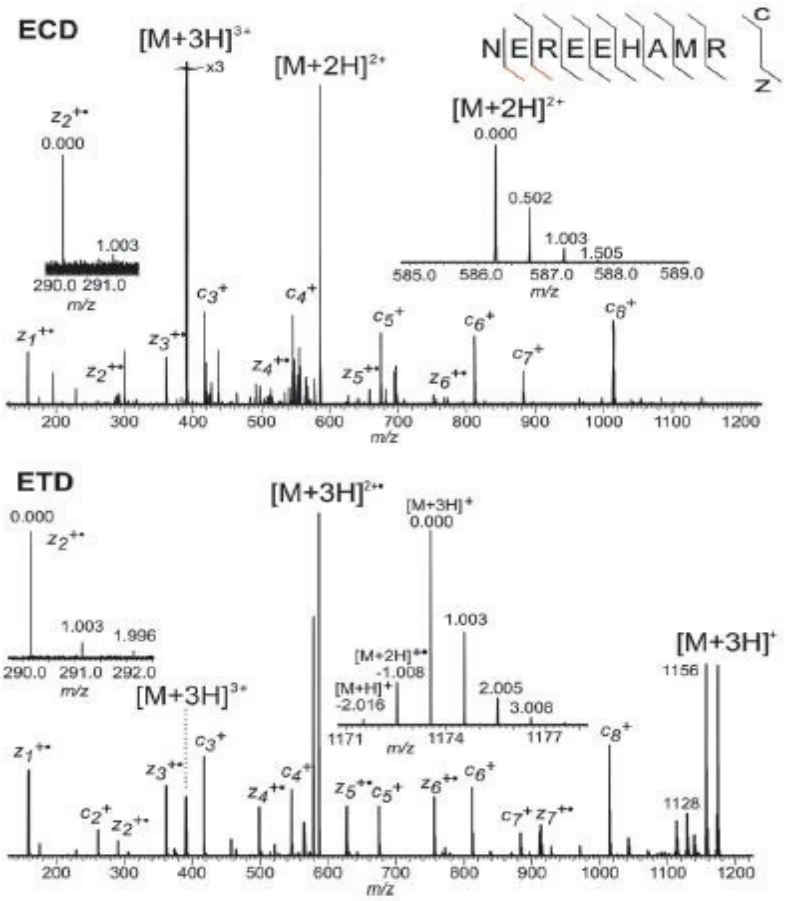


Figure 16 : Spectres de fragmentation du peptide NEREEHAMR dichargé en ECD et trichargé en ETD[155]

Comparativement à la CID/HCD où la fragmentation s'effectue en quelques millisecondes avec un rendement de fragmentation bien meilleur, ce sont là deux désavantages majeurs[145]. Il faut rappeler ici que les peptides sont identifiés via leurs spectres de fragmentation.

Ces méthodes d'activation nécessitent donc un long temps d'activation pour être efficaces et fonctionnent plus efficacement sur des peptides fortement chargés ou des protéines.

Au niveau de notre approche, les peptides étudiés seront en général chargés 2 ou 3 fois via les digestions trypsiques. Le temps d'activation devra être peu élevé afin d'avoir un temps de cycle court en vue d'obtenir le plus d'informations possibles sur nos échantillons.

Les activations types ETD/ECD ne semblent donc pas être adaptées pour des approches bottom-up mais plus pour des approches top-down ou structurale[150].

III.4) La photoactivation

La photoactivation repose sur l'absorption des ions en phase gazeuse lors de leur exposition à des rayonnements lumineux à différentes longueurs d'ondes entrainant la fragmentation de ces ions. Dans cette partie, les différentes méthodologies de "photodissociation"[10,13] seront abordées respectivement en fonction de la longueur d'onde du rayonnement utilisé car elle est un corolaire du nombre de photons qu'il sera nécessaire au composé d'absorber pour fragmenter. En découle le temps d'activation et surtout les canaux de fragmentation accessibles intimement reliés aux fragments générés.

III.4.a) Dissociation induite par photons infrarouges (IRMPD)

La première approche consiste en l'irradiation des ions en phase gazeuse à des longueurs d'ondes comprises entre 0,7 µm et 1 mm soit des rayonnements infrarouges à faible énergie « Infrared Multiple Photon Dissociation » (IRMPD). Le rayonnement IRMPD qui a été le plus utilisé pour l'étude des molécules biologiques est issu d'un laser à CO_2 émettant à 10,6 µm. Cela résulte du fait que les protéines ont une section efficace d'absorption importante à cette longueur d'onde, soit une forte probabilité d'absorber les photons émis. En effet, les peptides présentent un moment dipolaire permanent et leurs modes de vibration sont compatibles avec l'IR (liaison P-O, S-O et C-O). Cependant, du fait de la faible énergie par photons (0,12 eV) cette technique requiert l'accumulation d'un nombre important de photons avant de dissocier les peptides car l'énergie de dissociation de l'édifice moléculaire selon la voie la plus basse est typiquement de 1 à 2 eV. On parle alors de processus multiphotonique. L'IRMPD est ainsi caractérisé par l'accès à des canaux de fragmentation de basse énergie et un long temps d'activation[172]. En découle une fragmentation similaire à celle résultante de la CID et à la rupture des liaisons les plus labiles (fragments b et y). Néanmoins, il est important de noter que plusieurs modes vibrationnels absorbent et se dissocient très bien à cette longueur d'onde : celui de la liaison P-O[69] ainsi que les liaisons S-O et C-O[173]. L'IRMPD a été souvent implanté sur les FTICR en raison de la très faible pression de travail et à l'absence de refroidissement collisionnel rencontré dans d'autres instruments tels que les LIT. Ce refroidissement collisionnel repose sur l'introduction d'He à une pression de 1 milli Torr en vue de diminuer l'énergie cinétique des ions pour mieux piéger les ions et donc améliorer la résolution et la sensibilité de l'analyse[174]. Cependant, il a été montré que cette désexcitation entre en compétition avec l'activation des modes vibrationnels des peptides en IRMPD et donc rend délicates leurs fragmentations[175]. Malgré cela, l'IRMPD a quand même été implémenté sur des LIT : des lasers plus puissants (plus de photons par secondes)[176], un temps de piégeage allongé et une optimisation de la température lors de ce piégeage ont permis de surmonter la compétition du refroidissement collisionnel[177].

Malgré la génération de fragments similaire et un temps d'activation plus long que la CID, cette méthode d'activation a pu se révéler complémentaire. En effet, l'IRMPD ne génère pas de « low-mass cut-off » (l'IRMPD est indépendant des radiofréquences appliquées pour le piégeage des

ions dans un LIT) ce qui permet d'avoir accès à des fragments de basse masse et d'enrichir le panel d'information. Cela peut être bénéfique lors de l'analyse de composés de faible masse tels que les drogues[178]. De plus, via un temps d'activation allongé, l'IRMPD permet l'accès à des fragments de secondes générations : comme en CID, l'IRMPD génère la fragmentation des groupements les plus labiles et donc des pertes d'H_2O et de NH_3 peu informatifs. Or, grâce à cette seconde fragmentation, ceux-ci peuvent être à nouveau dissociés et compléter l'information obtenue[179].

En raison de son efficacité d'activation au niveau de liaison P-O beaucoup d'oligonucléotides[180] ou d'interaction ADN-ligand[181] ont été étudiées par cette technique. Il en va de même pour les peptides grâce à la richesse des fragments plus importants générés[176,182]. Cependant, cette méthode d'activation ne propose pas de spécificité quelconque au niveau de la fragmentation.

III.4.b) Photodissociation par laser Ultraviolet (UVPD)

Des photons de plus hautes énergies (ultraviolet (UV)) ont aussi été utilisés dans le but de faire photodissocier des peptides. Ceux-ci menant à des canaux de fragmentations plus élevés permettent d'obtenir plus rapidement des spectres de fragmentation plus riches. La méthodologie s'intitule UVPD. Voici une liste recoupant les principales longueurs d'ondes accessibles dans l'UV via les lasers disponibles sur le marché et utilisés pour l'analyse de peptides :

- F2 laser excimer: 157 nm – 7,9 eV/photon[183]

- ArF laser excimer: 193 nm – 6,4 eV/photon[184]

- Nd:YAG laser : 213 nm – 5,8 eV/photon[185]

- Nd:YAG laser : 266 nm – 4,7 eV/photon[186]

- XeF laser excimer: 351 nm – 3,5 eV/photon[187]

Concernant les lasers se situant à 157 et 193 nm, un à deux photons suffisent à faire dissocier un peptide et offrent une plus grande diversité de fragments qu'en CID. Cela peut-être constater sur la Figure 17 où il a été comparé la fragmentation du peptide DLYANTVLSGGTTMYPGIADR sous forme dichargé en UPVD (a) à 193 nm et en CID (b). Une meilleure identification en UPVD, de par la richesse du spectre obtenu, en a résulté (plus le Xcorr de SEQUEST est haut, meilleure est la confiance quant à l'identification) (Figure 17) :

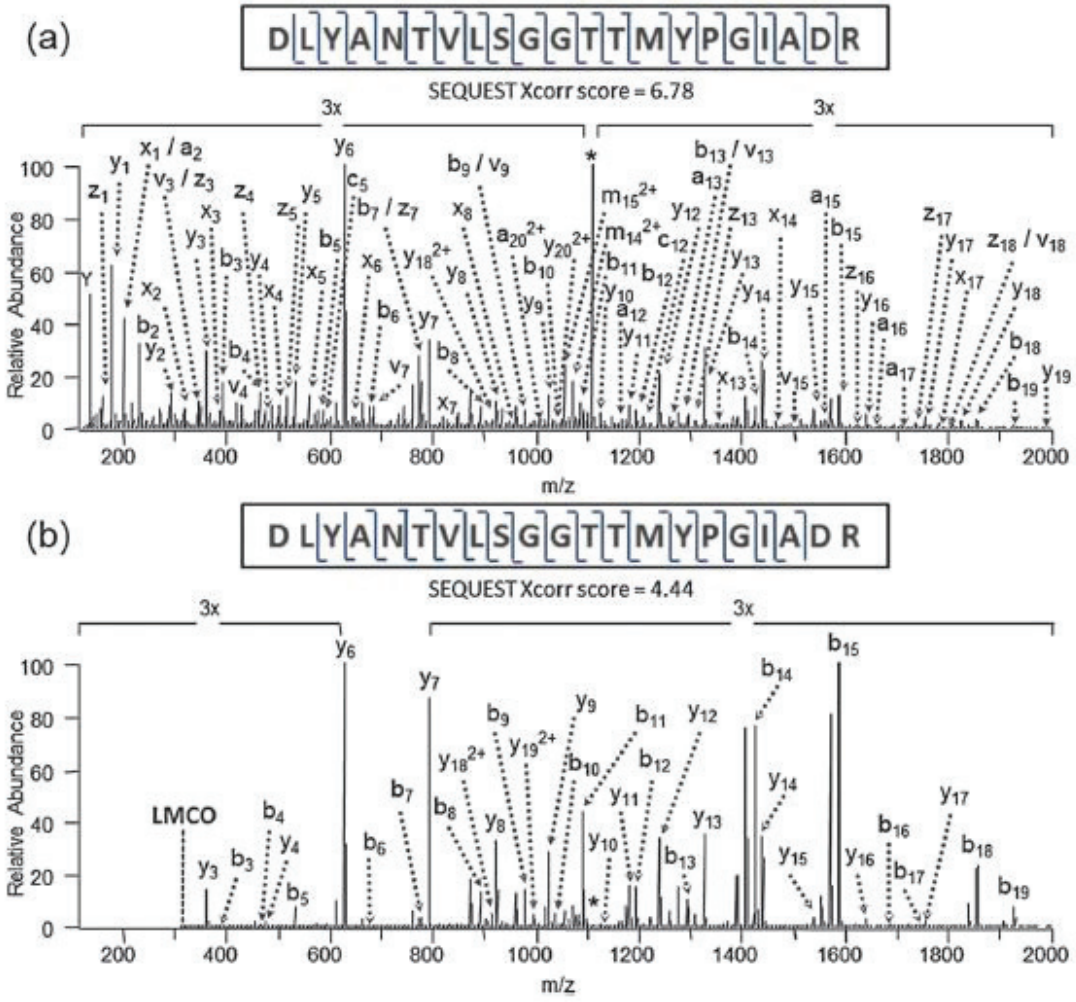


Figure 17 : Spectres de fragmentation du peptide DLYANTVLSGGTTMYPGIADR sous forme dichargé après activation en UPVD à 193 nm (a) et en CID (b)[184]

En effet, il est important de rappeler ici que la dissociation induite par des photons permet d'activer les états électroniques excités des molécules, orbitales moléculaires. Pour que les molécules soient excitées électroniquement, elles doivent absorber à la longueur d'onde d'émission du laser.

Pour le cas simple d'une molécule diatomique on peut représenter l'absorption d'un photon d'énergie hν par un état électronique fondamental ψ_0 d'énergie E_0 jusqu'à un état excité ψ_1 d'énergie E_1. Avec $E_1 - E_0 = h\nu$ (Figure 18).

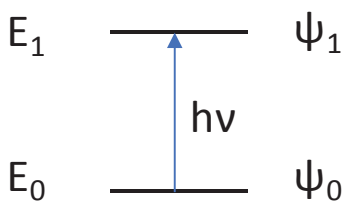


Figure 18 : Transition d'un état stable E_0 à un état excité E_1

Les orbitales moléculaires se différencient en trois catégories :

· Orbitales liantes : si l'on considère l'orbitale initialement vide et que l'on y dépose un électron puis deux électrons, la liaison entre les deux atomes se renforce et devient plus difficile à rompre, l'orbitale est dite constructive.

· Orbitales anti-liantes : si l'on considère l'orbitale initialement vide et que l'on y dépose un électron puis deux électrons, la liaison entre les deux atomes s'affaiblit et devient plus facile à rompre, l'orbital est dite destructive.

· Orbitales non-liantes : si l'on considère l'orbitale initialement vide et que l'on y dépose un électron puis deux électrons, la liaison entre les deux atomes n'est ni affaiblie ou renforcée.

La transition d'un état fondamental vers un état excité peut être décrite par l'excitation d'une orbitale liante, dite occupée, vers une orbitale non-liante, dite virtuelle. Dans le modèle simple d'une molécule diatomique, si l'énergie apportée par le photon est telle qu'elle permet de passer d'un état fondamental ψ_0 a un état excité non lié ψ_1, la molécule va se dissocier.

Comme nous l'avons décrit précédemment, plusieurs types de lasers sont utilisés pour l'UVPD. Aux longueurs d'onde inferieures à 260 nm, la plupart des électrons des liaisons p des peptides peuvent être excités et, en mode positif, de nombreuses fragmentations du squelette peptidique sont observées mais également des ions fragments de type v et w, non observées en CID. Les lasers émettant dans le VUV (150 - 220 nm) permettent d'exciter, entre autres, toutes les liaisons amides du squelette peptidique sans aucune dérivation [188].Or celles-ci sont présentes tout au long du squelette peptidique. Elles vont absorber, fragmenter à leur niveau et/ou redistribuer l'énergie aux alentours engendrant la production de tous les types de fragments (a, b, c, x, y, z) comme on peut le voir sur le spectre ci-dessus (Figure 17 a). La compréhension des mécanismes de fragmentation à 157 nm fut l'objet d'études menées en particulier par l'équipe de Reilly's[183,189]. Ces études proposent une rupture homolytique au niveau de la liaison C-C_α engendrant la production des fragments a et x. Il s'en suit des processus secondaires tels que la migration de protons qui engendreront cette diversité de fragments. Il a même été possible de générer des fragments v et w

qui ont permis de discerner la présence de leucine ou d'isoleucine dans des séquences peptidiques[190,191]. La conservation des PTM a aussi été observée[191]. Il a été montré que les mécanismes mis en jeu à 193 nm étaient les mêmes qu'à 157 nm[192–195]. Corrélé à cette richesse, la rapidité de l'étape de fragmentation permet l'application de cette méthodologie dans les approches bottom-up en protéomique. En effet, il a été montré qu'en usant d'une impulsion laser de 5 ns, 20 à 98 % de rendement de photodissociation a été obtenu sur des peptides tout en précisant qu'une séquence contenant des acides aminés aromatiques induisait de meilleur rendement[184]. Cela est tout à fait compétitif avec la CID et bien au-delà de ce qu'il est possible d'obtenir en IRMPD. Dans le même temps, une autre méthodologie intitulée « activated EPD » (a-EPD) a également été décrite, où en mode négatif, les édifices moléculaires étaient irradiés à 220 nm, menant à un photodétachement d'électrons suivi d'une activation par CID qui aboutissait à des couvertures de séquences importants[196]. Cependant, si pour une approche bottom-up, de manière générale, cette méthodologie UPVD est performante, il en va beaucoup moins pour le mode DIA. Ces longueurs d'ondes, générant des spectres de fragmentation encore plus riches, l'interprétation des spectres de (co)-fragmentation ne pourra être qu'encore plus problématique.

A 266 nm seuls les acides aminés aromatiques peuvent absorber. Cependant, l'histidine et la tyrosine ont un rendement de photodissociation très faible et, le tryptophane fluorescant, il ne redistribue pas l'énergie absorbée à l'édifice peptidique pour le faire fragmenter. L'intérêt principal, ici, résulte de la spécificité de fragmentation des ponts disulfures et des liaisons C-I à cette longueur d'onde. Si des peptides contiennent des ponts disulfures et qu'ils sont soumis à ce rayonnement en phase gazeuse, alors une rupture homolytique aura lieu au niveau de ces ponts tout en laissant les autres liaisons intactes[197]. Il en va de même pour les liaisons C-I. Julian *et al* ont donc développé un procédé reposant sur l'iodination des peptides au niveau des tyrosines permettant la production de radicaux à 266 nm. Ces radicaux pourront ensuite être dissociés via une activation par CID ou IRMPD. Cette manipulation a été effectuée au sein d'un FTICR ou d'une LIT, spectromètre où les ions peuvent être piégés longtemps pour subir les 2 étapes d'activation[198]. Les principaux fragments observés correspondent à des fragments a, b, c, z, des ruptures au niveau des chaines latérales ainsi que des pertes de CO_2[198,199]. Grâce à cette technique, des applications au niveau protéomique ont pu être décrites telles que l'identification des sites de phosphorylation[200] ou la différenciation des acides aminés D ou L[201]. De par la nécessité de la fragmentation en 2 étapes (long temps d'activation) et d'une spécificité de fragmentation qui ne semble ne pas répondre aux problèmes de réduction de la complexité des spectres de (co)-fragmentation, cette irradiation à 266 nm ne sera pas retenue.

Une autre démarche a été développée via l'utilisation de laser émettant à 351 nm. A cette longueur d'onde, les peptides n'absorbent pas : il faut donc les dériver avec un groupement chromophore. Cela a été réalisé par Brodbelt *et al* avec la dérivation des acides aminés histidines et tyrosines à l'aide d'un chromophore nitroazobenzène qui augmente la section efficace

d'absorption à 351 nm. Le rendement de dérivation est estimé à plus de 90 % et les fragments générés sont similaires à ceux obtenus en CID (b, y et pertes de NH_3 et de H_20) pour un temps d'activation correspondant à 20 ms[202]. Le taux de photo-fragmentation est lui aussi très satisfaisant comme on peut l'observer sur le spectre ci-dessous (Figure 19 b). Il convient aussi de préciser que seuls les peptides contenant des acides aminés histidines ou tyrosines photo-fragmentent. Les peptides qui ne contiennent pas le chromophore n'absorbent pas et ne sont donc pas photo-fragmentés (Figure 19 a).

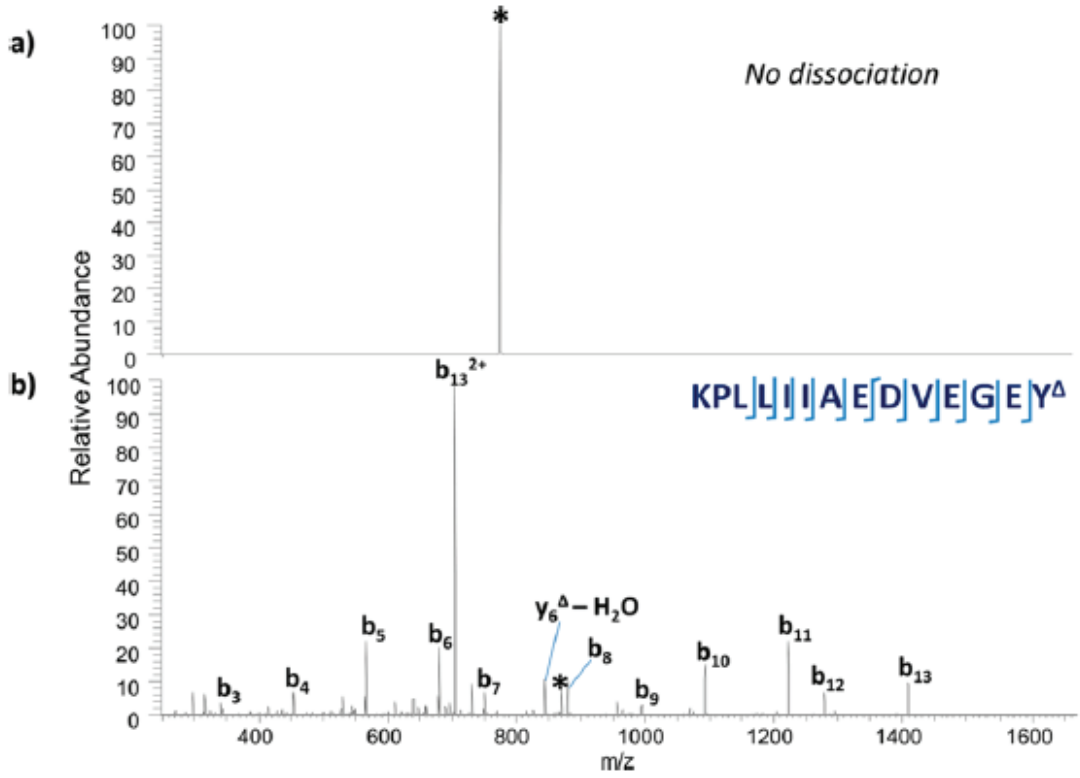


Figure 19 : Spectre de photo-fragmentation après irradiation à 351 nm d'un peptide non dérivé avec le chromophore nitroazobenzène (a) et du peptide KPLLIIAEDVEGEY dérivé avec ce chromophore sur l'acide aminé Y (b)[202]

La même approche a été développée pour la dérivation des acides aminés cystéine avec le chromophore Alexa Fluor réagissant spécifiquement avec les thiols. Cette méthode a été appliquée pour le suivi des peptides à cystéine issue des régions spécifiques d'anticorps et a permis d'aider à la caractérisation des régions CDR d'anticorps[203]. Cette spécificité de dérivation/fragmentation des peptides contenant des acides aminés cystéine se révèle très intéressante pour notre approche DIA. En effet, l'acide aminé cystéine est un acide aminé rare : il est présent à une fréquence de 2 % dans le protéome. Cependant, 89 % des protéines humaines contiendraient un acide aminé cystéine et plus de 60 % en contiendraient 2[15]. Ce qui signifie qu'en ciblant spécifiquement les peptides à cystéine, nous pourrions couvrir potentiellement tous le protéome humain à raison d'un peptide au minimum par protéine tout en diminuant

drastiquement la complexité des spectres DIA générés en raison de la rareté de ces peptides. Cela implique bien évidemment le fait de supposer leur détectabilité en spectrométrie de masse. Cependant, comme le montre la figure ci-dessous, le taux de photo-fragmentation de l'Alexa Fluor à 351 nm ne semble pas être optimal malgré la richesse des fragments produits posant peut-être la limite de cette approche (Figure 20).

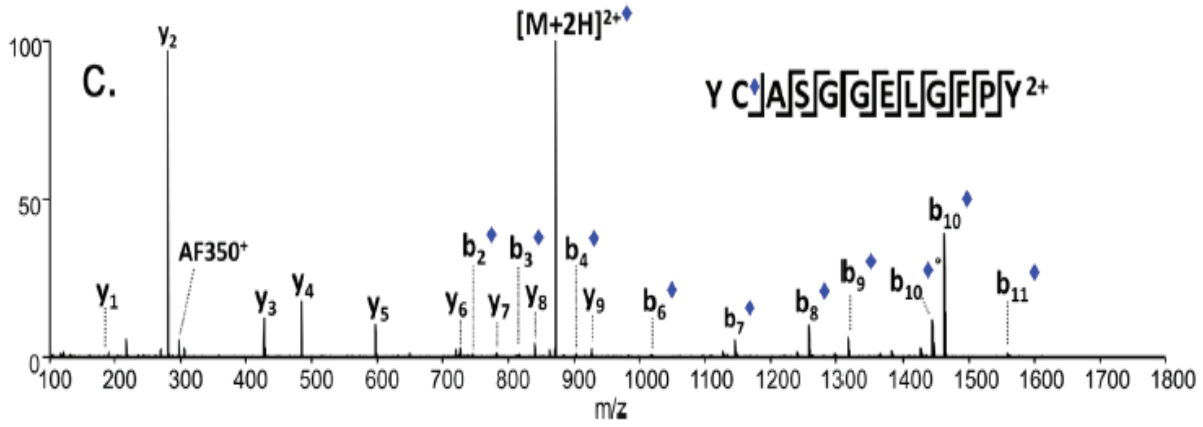


Figure 20 : Spectre de photo-fragmentation après irradiation à 351 nm du peptide à cystéine YCASGGELGFPY dérivé avec le chromophore Alexa Fluor sur l'acide aminé cystéine[203]

Il nous faut trouver maintenant une manière de détecter ces peptides à cystéine spécifiquement tout en obtenant un rendement de dérivation et un taux de photo-fragmentation le plus important possible.

III.4.c) Dissociation induite par laser dans le visible (LID)

Dans cette optique, l'équipe de Jérôme Lemoine a développé une méthodologie intitulée photo-SRM[14]. Celle-ci repose sur une activation laser dans le visible à 473 (ou 532) nm étant donné que les peptides n'absorbent pas naturellement à ces longueurs d'ondes. En revanche, une fois dérivés avec un chromophore qui, lui, absorbe à ces longueurs d'onde, ils peuvent photo-fragmentés. Cette méthodologie a été appliquée pour l'analyse de peptides à cystéine dans un mélange de plasma humain dilué dans du plasma de rat : les peptides à cystéine étaient dérivés avec un groupement chromophore DABCYL (quencher) puis analysés par photo-SRM au sein d'un spectromètre de masse triple quadripôles : la dissociation par CID étant remplacée par une irradiation laser à 473 nm.

Comparativement à la SRM, cette méthode a permis d'atteindre, dans le cas de la C4b binding protéine, une LOQ 10 fois inférieure à celle obtenue en SRM du fait de la réduction de l'effet d'isobarie beaucoup plus important en SRM. En effet, des composés détectés dans les blancs interféraient sur la mesure dans les basses concentrations (Figure 21). En Photo-SRM à 0 µg/ml il n'y a pas d'interférence (Figure 21 a) contrairement en SRM (Figure 21 c) ce qui permet de diminuer la LOQ a 6,75 µg/mL pour la Photo-SRM au lieu de 67,5 µg/mL en SRM.

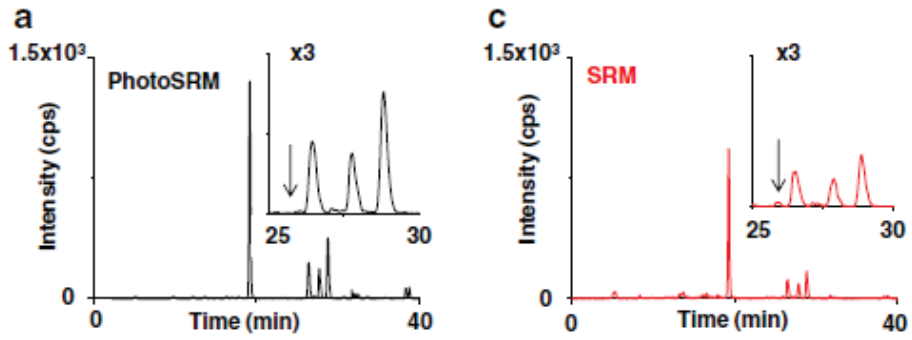


Figure 21 : Absence d'interférence dans le transition SRM suivi du peptide rapporteur de la C4b binding protéine en photo-SRM (a) alors que présente en SRM classique (c)[14]

La LID à 473 nm ciblant spécifiquement des peptides à cystéine a donc déjà permis de répondre à des problèmes d'interférences liées à la complexité des fluides biologiques. A la suite de ces résultats, la méthodologie a été implémentée sur un Qexactive [9]. Cette étude a montré que cette méthodologie était efficace pour détecter spécifiquement des peptides à cystéine dérivés lors d'approche bottom-up dans des matrices complexes. Le potentiel quant à la réduction des spectres de (co)-fragmentation en DIA est donc réel. Qui plus est, une attention a été portée sur le taux de photo-fragmentation et le temps d'activation (Figure 22) :

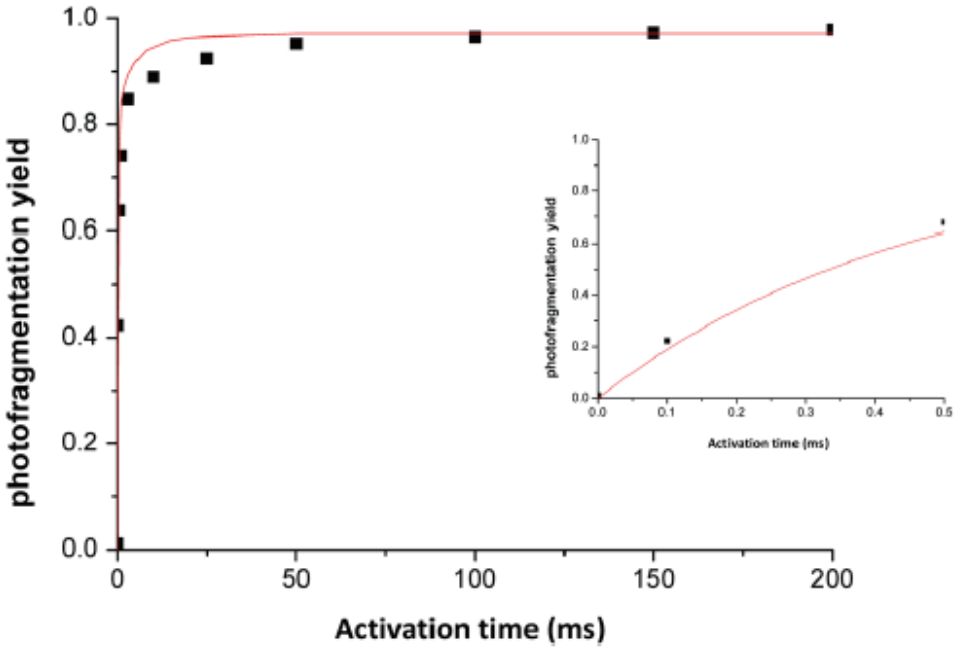


Figure 22 : Evolution du taux de photo-fragmentation du peptide à cystéine IVGGWECEK (2+) après irradiation à 473 nm en fonction du temps d'activation[9]

On peut y remarquer qu'un taux de photo-fragmentation de plus de 90 % est atteint en 25 ms ce qui est suffisamment court pour rendre cette irradiation effective avec un couplage chromatographique et donc en bottom-up[9].

De toutes les méthodes d'activation approchées celle-ci semble être l'une des plus prometteuses, elle sera retenue pour être utilisée avec le mode en DIA en remplacement de la CID.

Conclusion

Nous avons pu voir lors de ce chapitre que la spectrométrie de masse est, de par sa rapidité et son multiplexage massif, une alternative efficace pour identifier et valider des biomarqueurs protéines putatifs au sein de matrices biologiques malgré une sensibilité moindre comparativement aux tests immunologiques.

Nous avons également pu constater que, étant impliquées dans la régulation d'une multitude de processus en rapport avec l'homéostasie du corps humains, les protéines kinases humaines étaient des biomarqueurs protéines putatifs pertinents.

A travers l'introduction au différents modes accessibles en spectrométrie de masse en vue d'identifier et quantifier des biomarqueurs protéines au sein de matrices biologiques, nous avons pu prendre conscience que le mode DIA, exploitant au maximum la gamme dynamique des appareillages haute résolution tout en permettant le suivi de plusieurs milliers de composés par analyse de manière précise et répétable, était un mode intéressant selon ce but. Seulement, celui-ci fait face à une limite considérable : la difficulté à identifier des composés en raison de l'obtention de spectres de (co)-fragmentation trop complexes. Il est donc nécessaire de pallier à ce problème pour pouvoir profiter de tous les avantages offerts par cette technique. Nous avons montré qu'en se tournant vers l'utilisation d'une méthode d'activation des ions plus spécifique lors du processus de fragmentation, il est possible de lever ce problème.

Dans cette optique, nous nous sommes tournés vers la méthode d'activation LID à 473 nm qui permet le suivi exclusif des peptides à cystéine. Ce procédé implique de dériver spécifiquement les peptides avec le chromophore DABCYL C2-maléimide sur leur acide aminé cystéine et de les faire photo-fragmenter à l'aide d'un laser émettant à la longueur d'onde d'absorption du chromophore (473 nm). Ainsi, seul les peptides dérivés avec le chromophore seront photo-fragmentés. Ce qui nous a poussé à nous tourner vers cette méthode d'activation résulte de son ciblage des peptides à cystéine. En effet, l'acide aminé cystéine est un acide aminé rare (présent dans 89 % des protéines humaines à 1 raison d'un acide aminé par protéine) de ce fait, en suivant spécifiquement les peptides à cystéine nous pouvons d'une part, être exhaustif en couvrant potentiellement tous le protéome humain à raison d'un peptide au minimum par protéine, et, d'autre part, diminuer drastiquement la complexité des spectres de (co)-fragmentation générés en DIA en raison de la rareté de ces peptides.

Dans le chapitre suivant, nous présenterons le développement et la mise en application de la DIA-LID à 473 nm sur un appareil Qexactive avec une bibliothèque de spectres en vue d'identifier et de quantifier spécifiquement des protéines kinases humaines au sein de matrices biologiques.

Chapitre II : Développement de la méthodologie DIA-LID à 473 nm au sein d'un spectromètre de masse haute résolution pour l'identification et la quantification de protéines kinases humaines

Introduction

Dans ce chapitre il sera premièrement introduit le matériels et méthodes utilisés pour l'établissement de la bibliothèque de spectres et la mise en place de la méthodologie DIA-LID à 473 nm sur le Qexactive : il sera présenté tous les produits utilisés, l'appareillage LC/MS/MS dont le couplage Qexactive-laser ainsi que le protocole en détail.

Un point sera ensuite effectué sur la méthode d'activation LID permettant la photo-fragmentation de peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm : il sera présenté toutes les informations relatives à la dérivation, la stabilité et à la fragmentation laser des peptides à cystéine dérivés.

Troisièmement, nous nous attarderons sur la construction de la bibliothèque de spectres : il sera présenté les informations relatives à la sélection des peptides à cystéine composant cette bibliothèque ainsi que les manipulations expérimentales et computationnelles nécessaires à sa construction.

Quatrièmement, il sera démontré l'efficacité de la méthodologie DIA-LID à 473 nm pour identifier et quantifier des peptides à cystéine dérivés au sein de matrices biologiques : cette opération sera effectuée en dopant des échantillons d'E.Coli avec des peptides à cystéine synthétiques issus de la bilbiothèque de spectres.

Ensuite, la méthodologie DIA-LID à 473 nm sera appliquée pour quantifier des peptides à cystéine dérivés endogènes au sein d'un extrait cellulaire mammaire cancéreux humain.

Enfin, il sera présenté des perspectives en vue de perfectionner la méthodologie DIA-LID à 473 nm pour identifier et quantifier des peptides à cystéine dérivés au sein de matrices biologiques : il sera présenté 2 approches. La première abordera une optimisation du protocole utilisé dans le matériel et méthode, la seconde se penchera sur l'utilisation d'un nouveau chromophore.

I) Matériels et Méthodes

Cette section introduit le matériels et méthodes utilisé tout au long de ce chapitre (quelques précisions spécifiques à chaque partie pourront être néanmoins reprécisées le cas échéant). Il sera d'abord présenté les échantillons qui ont été sélectionnés en vue des différentes expérimentations. Un point sera ensuite fait sur le chromophore retenu pour dériver spécifiquement les peptides à cystéine. Le protocole standard de dérivation et de préparation des échantillons avant l'analyse LC/MS/MS sera également explicité. Pour terminer, le système chromatographique suivi du montage Qexactive-laser ainsi que leurs paramétrages respectifs seront décrits.

I.1) Echantillons

- Pour la réalisation de la bibliothèque de spectres spécifique des protéines kinases humaines, 446 peptides à cystéine correspondant à 446 protéines kinases humaines ont été commandés et synthétisés chez Thermo Scientific (Peptide grade 2). Ils ont été reçus déjà solubilisés dans un milieu H_2O/Acétonitrile (ACN) (50/50) + 0,5 % acide formique (FA) à différentes concentrations selon les peptides. En vue de commencer à enrichir la bibliothèque de spectres avec d'autres composés 63 peptides à cystéine supplémentaires issus de 33 protéines humaines ont également été commandés.

- Pour étudier la stabilité des composés dérivés avec le DABCYL C2-maléimide l'hormone peptidique Vasopressine et le neuropeptide Ocytocine seront utilisés (Proteogenix – 96 % et Sigma Aldrich – 97 % respectivement)

- 3 peptides à cystéine (VMCDSNDLPK, DICNDVLSLLEK et YMPQNPCIIATK) utilisés pour démontrer l'efficacité de la méthodologie DIA-LID pour quantifier des peptides à cystéine dans des matrices biologiques de manière répétable ont été commandés chez JPT à une pureté de 70 %.

- Les échantillons d'Escherichia Coli (E.Coli) employés pour démontrer l'efficacité de la méthodologie DIA-LID pour identifier et quantifier des peptides à cystéine dérivés au sein de matrices biologiques ont été préparés au laboratoire via une culture en milieu LB de la souche BL21(DE3).

- Pour l'analyse des protéines kinases humaines endogènes au sein de fluides biologiques, nous nous sommes tournés vers un extrait cellulaire mammaire cancéreux. Un extrait de la lignée HS578T nous a été fourni par le centre de recherche en cancérologie affilié au centre hospitalier Léon Berard à Lyon.

- Pour expérimenter l'étape de dérivation des peptides à cystéine après l'étape de digestion 3 protéines conservées sous forme de poudre seront utilisées (Sigma Aldrich – 97 % pour les 3) : Albumine humaine, Lysozyme de poulet et Serotransferrine humaine.

En vue de mesurer les concentrations protéiques des matrices biologiques (E.Coli et extrait cellulaire humain) deux tests de Bradford ont été effectués. Ce protocole ne sera pas présenté ici. Les concentrations obtenues sont pour E.Coli de 1177 µg/mL et de 340 µg/mL pour l'extrait cellulaire humain.

I.2) Choix du chromophore

Une fois les échantillons sélectionnés, il a fallu opter pour le chromophore adéquat en vue de dériver spécifiquement les peptides à cystéine sur l'acide aminé cystéine. A partir des travaux précédemment réalisés au laboratoire[204] le chromophore DABCYL C2-maléimide a été sélectionné (Figure 23).

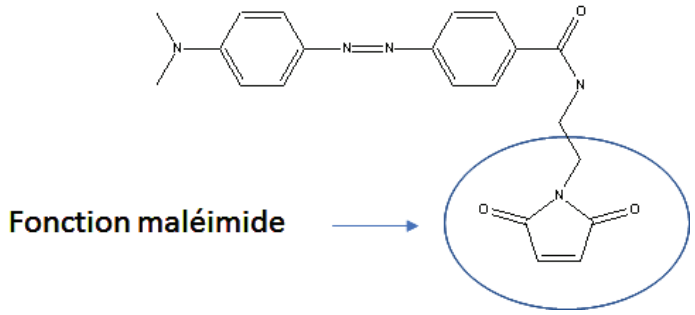


Figure 23 : Structure chimique du DABCYL C2-maléimide

Le choix de se tourner vers le DABCYL C2-maléimide (masse molaire (Mw) = 391 g/mol) vient premièrement du fait de sa propriété quencher. C'est-à-dire que suite à une excitation laser ce chromophore ne va pas fluorescer mais va absorber l'énergie et la redistribuer à l'édifice moléculaire auquel il est relié générant ainsi sa photo-fragmentation. Greffé sur des peptides, il permet l'obtention de leurs spectres de photo-fragmentation nécessaire à leurs identifications[9]. Ensuite, ce chromophore possède une fonction maléimide qui dérive spécifiquement le groupement thiol (SH), groupement qui est aussi spécifique de l'acide aminé cystéine, parmi les autres acides aminés composant les protéines. Ce chromophore permet donc de cibler spécifiquement les peptides à cystéine, parmi tous les peptides issus d'une digestion trypsique. Enfin, pour terminer, ce chromophore absorbent à 473 nm (longueur d'onde émise par le laser disponible au laboratoire) comme nous pouvons le voir sur son spectre d'absorption en solution (Figure 24). Son λ_{max} se situe certes aux alentours de 430 nm, mais à 473 nm il possède encore une absorption importante aux alentours de 0,7.

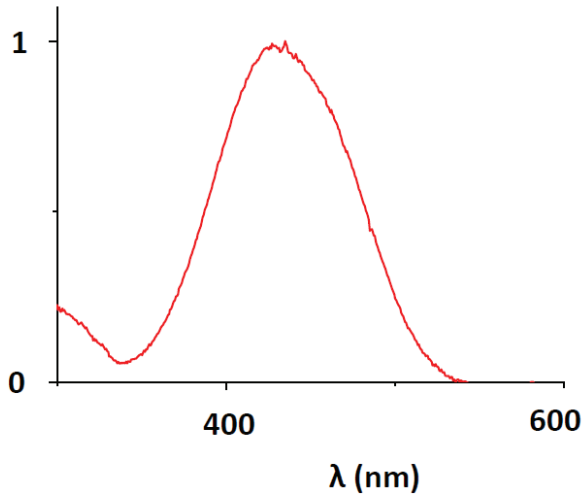


Figure 24 : Spectre d'absorption du DABCYL C2-maléimide en solution

Son spectre d'action (taux de photodissociation d'un ion d'intérêt corrigé du flux de photons en fonction de la longueur d'onde du laser) en phase gazeuse a été étudié lorsque greffé sur des peptides révélant une bonne absorbance à 473 nm[204].

Il est à noter que peu de chromophore quencher, dérivable sur des cystéines avec un taux de photo-fragmentation important sont disponible commercialement. Le chromophore QSY7 en fait partie mais il s'est révélé moins efficace que le DABCYL C2-maléimide lors d'une étude effectuée au laboratoire par le passé (moins bon taux de photo-fragmentation)[204]. Un nouveau chromophore en phase de test sera présenté dans la partie VI.2 de ce chapitre.

I.3) Préparation d'échantillons

Un protocole de dérivation et de préparation d'échantillons avant analyse LC/MS/MS a été établi pour les échantillons suivants :

- Les 446 peptides à cystéine correspondant à 446 protéines kinases humaines ont été solubilisés dans un milieu H_2O/ ACN (50/50) + 0,5 % FA par pool de 10 et de 9 à une concentration théorique de 400 µg/mL par peptide à cystéine. Ils ont été ensuite regroupés en 1 seul pool pour l'analyse DDA-LID à 473 nm servant à la construction de la bibliothèque (1 µg de chaque peptide à cystéine).

- Les 63 peptides à cystéine correspondant à 33 protéines humaines ont été solubilisés dans un milieu H_2O/ ACN (50/50) + 0,5 % FA par pool de 10 et de 3 à une concentration théorique de 400 µg/mL par peptide à cystéine. Ils ont été ensuite regroupés en 1 seul pool pour l'analyse DDA-LID à 473 nm servant à la construction de la bibliothèque (1 µg de chaque peptide à cystéine).

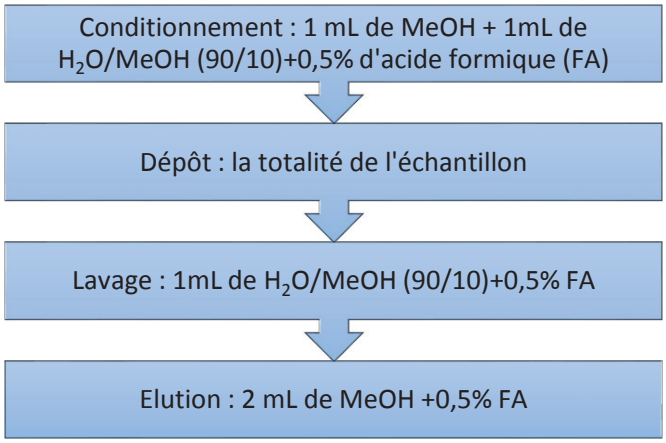
- 2 µg de l'hormone peptidique Vasopressine et du neuropeptide Ocytocine en matrice aqueuse

- 100 µg d'E.Coli dopé avec des peptides à cystéine issus de la bibliothèque de spectres. Les quantités ajoutées pour chaque peptide à cystéine seront explicitées dans les parties IV.1 et IV.2 de ce chapitre.

- 100 µg d'E.Coli non dopé

- 255 µg d'extrait cellulaire mammaire cancéreux HS578T non dopé

Tout d'abord les échantillons sont dérivés avec le DABCYL C2-maléimide dans un mélange méthanol (MeOH)/acétate d'ammonium (AA) 60 mM (50/50). Le choix de ce mélange résulte de la nécessité de maintenir le pH aux alentours de 6,8 pour la dérivation et de la faible solubilité du DABCYL C2-maléimide en milieu aqueux. Lors de cette étape, 3 équivalents molaires de TCEP sont ajoutés pour réduire les ponts disulfures et 5 équivalents de chromophore pour dériver les cystéines. Les 2 réactifs sont ajoutés au même moment et la mixture est gardée dans l'obscurité pendant 4h. Le noir est une nécessité car le DABCYL C2-maléimide se dégrade à la lumière. Lors de l'analyse de l'extrait cellulaire et des échantillons de E.Coli une étape de digestion trypisque a ensuite été réalisée à 37°C pendant 1 nuit avec un ratio massique de 1 pour 20. Les échantillons sont ensuite purifiés par SPE-HLB dont voici les différentes étapes.



Les composés élués dans 2 mL de MeOH sont alors évaporés sous azote puis resolubilisés dans un volume de reprise (V.rep) de 100 µL ou 85 µL de H_2O/MeOH (95/5) + 0,5 % FA afin de reconcentrer les échantillons avant l'injection.

Le choix de se tourner vers un solvant de reprise H_2O/MeOH (95/5) + 0,5 % FA après l'évaporation résulte d'une étude menée afin de préconcentrer nos échantillons au maximum en tête de colonne via l'utilisation d'une composition en solvant de reprise moins éluante que la phase mobile utilisée. Dans cette optique différentes compositions en solvant ont été testées. Les gradients choisis débutant à 15 % en ACN (comme nous le verrons par la suite), il a été testé 3 volumes de reprises différents sur un mélange de 10 peptides à cystéine dérivés (Tableau 2). Il est

à préciser que dorénavant tous les peptides à cystéine dérivés avec le DABCYL C2-maléimide seront annotés avec un C souligné (C) :

Peptide	Intensité (95/5+ 0,5%FA)	Intensité (90/10 +0,5 %FA)	Intensité (85/15 + 0,5%FA)
CHAPENVELALR	3,12E+05	2,58E+05	2,51E+05
CLAALASLR	2,74E+05	2,44E+05	2,21E+05
VHSENNACINFK	3,90E+05	1,99E+05	2,61E+05
YSFGATCVK	3,40E+05	2,76E+05	2,46E+05
ICQAHFAPISPGFSR	2,44E+05	2,03E+05	1,97E+05
CPAAGNPTPTIR	7,59E+05	6,72E+05	6,38E+05
NCLVGENNVLK	7,00E+05	6,39E+05	5,66E+05
SQCLSTLVRPVFGELK	3,49E+05	3,16E+05	3,08E+05
GECVITPSTDVK	3,04E+05	2,54E+05	2,12E+05
AACLLDGVPVALK	2,15E+05	1,69E+05	1,54E+05

Tableau 2 : Variation de la réponse en intensité de 10 ions précurseurs tri ou quadrichargés issus de 10 peptides à cystéine dérivés en fonction de la composition du solvant de reprise post évaporation

Les 3 échantillons ont été analysés successivement dans cet ordre de solvant de reprise : 95/5, 90/10 et 85/15 tout en étant séparés par des blancs de H_2O/MeOH (95/5) + 0,5 % FA. Le signal présenté dans le tableau est celui de l'ion précurseur le plus intense obtenu pour chaque peptide après une analyse en FMS. Comme nous pouvons le voir, l'utilisation d'un solvant de reprise à 95/5 + 0,5 % FA tend à augmenter le signal obtenu. Qui plus est, dû au fait que cet échantillon a été injecté en premier et à l'absence de carry over au niveau des blancs, cet effet résulte bien de la préconcentration en tête de colonne.

Lors de l'injection sur le système chromatographique il a été injecté sur la colonne :

- 200 ng pour chaque peptide à cystéine (correspondant à 1 des 446 protéines kinases humaines) analysé soit 89 µg pour les 446 (V.rep = 100 µL)

- 400 ng pour l'hormone peptidique Vasopressine et le neuropeptide Ocytocine soit 800 ng pour les 2 (V.rep = 100 µL)

- 200 ng pour chaque peptide à cystéine (correspondant à 1 des 33 protéines humaines) analysé soit 13 µg pour les 63 (V.rep = 100 µL)

- 24 µg d'E.Coli dopé ou non (V.rep = 85 µL)

- 60 µg d'extrait cellulaire mammaire cancéreux HS578T (V.rep = 85 µL)

I.4) Partie chromatographie

Une fois les différents échantillons préparés, les peptides ont été séparés à l'aide d'un système chromatographique en amont du couplage Qexactive-laser. Le système utilisé se décompose en 2 parties :

- Une pompe Surveyor MS 68009 de chez Thermo Scientific. Celle-ci imposera un débit de 300 µL/min ainsi qu'une composition en phase mobile faite de 2 solvants (un solvant A (H_20 + 0,5 % FA) et un solvant B (ACN + 0,5 % FA) pour toutes les analyses présentées par la suite. La pression maximum tolérée par cette machine est de 400 bars.

- Un passeur automatique de chez CTC Analytics qui prélève un volume défini d'échantillon contenu dans un flacon. Le volume, dit d'injection, utilisé pour toutes les manipulations présentées par la suite est de 20 µL.

La colonne chromatographique qui a été utilisée pour toutes les analyses est une colonne chromatographique C_{18} Xbridge de chez Waters avec les caractéristiques suivantes (Tableau 3) :

Diamètre de particule	3,5 µm
Diamètre interne	2,1 mm
Taille des pores	300 A
Longueur	10 cm

Tableau 3 : Caractéristique de la colonne chromatographique utilisée pendant tout ce chapitre

Deux gradients chromatographiques ont été utilisés durant les diverses expérimentations menées au laboratoire. Ils sont présentés ci-dessous et affiliés à un numéro qui sera réutilisé au long du chapitre en vue de préciser quel est le gradient qui a été utilisé pour telle analyse.

Gradient 1 (Tableau 4) :

Temps (min)	% Solvant B (ACN + 0,5 % FA)
0	15
1,5	15
23	75
23,1	95
26	95
26,1	15
30	15

Tableau 4 : Variation de la composition en solvant B de la phase mobile en fonction du temps lors du gradient 1

Gradient 2 (Tableau 5) :

Temps (min)	% Solvant B (ACN + 0,5 % FA)
0	15
4	15
74	50
74,1	95
79	95
79,1	15
85	15

Tableau 5 : Variation de la composition en solvant B de la phase mobile en fonction du temps lors du gradient 2

I.5) Spectrométrie de masse

Aux vues du matériel disponible au laboratoire, l'approche DIA-LID à 473 nm a dû être effectuée sur un spectromètre de masse Qexactive. Ci-dessous est présenté le fonctionnement de cet instrument ainsi que de la mise en place du dispositif permettant d'irradier avec un laser à 473 nm au sein de la cellule de collision en remplacement de la méthode d'activation par HCD. Un point sera également fait sur les différents paramètres utilisés lors de toutes les analyses.

Le Qexactive est un analyseur en masse haute résolution confectionné par Thermo Scientific. Ci-dessous est présenté un schéma exhaustif du Qexactive (Figure 25) :

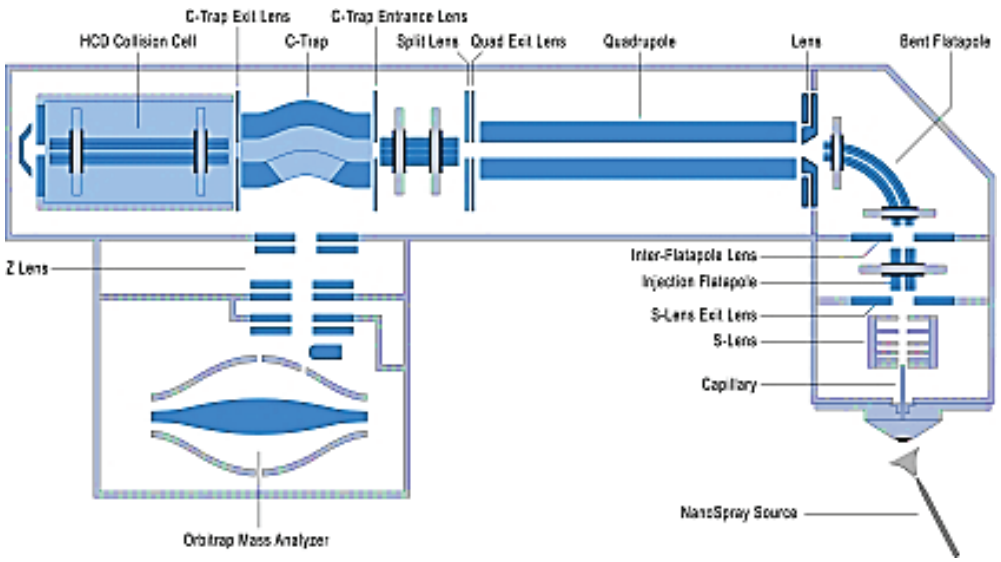


Figure 25 : Schéma détaillé d'un Qexactive[205]

La présentation de l'appareil s'effectuera selon le chemin parcouru par l'ion de son entrée dans le spectromètre jusqu'à la détection.

I.5.a) La source d'ionisation à pression atmosphérique (API)

Le 1er élément composant le Qexactive est une source API. Cette source se situe à l'interface entre la partie chromatographique et le spectromètre de masse. Elle permet la formation d'ions en phase gazeuse à partir d'ions en phase liquide élués en amont par le système chromatographique. Cette source permet différents modes d'ionisations dont l'électrospray (ESI), l'ionisation chimique à pressions atmosphérique (APCI), la photo-ionisation à pression atmosphérique (APPI) ou encore l'ionisation nanospray (NSI). Durant cette thèse, il n'aura été utilisé que le mode ESI exclusivement. Il sera le seul à être détaillé par la suite.

La source ESI est une source d'ionisation douce qui va désolvater les ions à pression atmosphérique. Tout d'abord, un champ électrique intense est généré sur le liquide émergeant du système chromatographique à l'aide d'une forte différence de potentiel (entre 3 et 5 kVolts) appliqué entre une contre électrode et le capillaire amenant ce même liquide. Les charges seront ainsi mises en mouvement. Il est aussi possible d'apporter de la chaleur pour aider à la désolvatation. En fonction de la polarité choisie, une concentration des ions chargés positivement ou négativement s'effectuera à la surface du liquide et à la pointe du capillaire engendrant une accumulation de charges en surface jusqu'à un surplus. Dès lors, une optimisation de la forme du liquide devient nécessaire et se matérialisera par la formation d'un cône de Taylor et la production de gouttelettes (Figure 26).

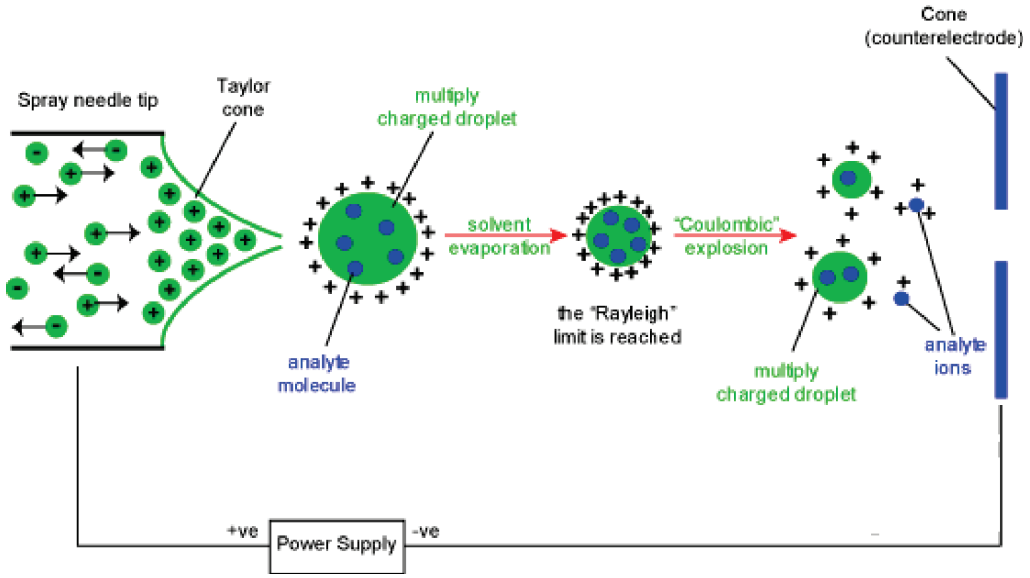


Figure 26 : Principe de fonctionnement d'une source ESI[206]

Une fois ces gouttelettes obtenues, elles diminueront en taille par évaporation : lorsque la répulsion électrostatique deviendra trop forte, atteignant la limite de Rayleigh les gouttelettes se dissocieront en microgouttelettes par des processus de fissions de Coulomb. Ce phénomène s'appliquera sur plusieurs générations de gouttelettes jusqu'à former des ions chargés porteurs

de plus ou moins de charges en fonction de leurs nombres de sites ionisables.

La source ESI va produire essentiellement des ions multichargés selon quatre procédés :

- L'ionisation par séparation de charges : les analytes existent en solution comme ions et se retrouvent directement ionisés en gouttes.

- La formation d'adduits : elle se produit avant la séparation de charge. En mode positif, les analytes s'adduisent avec du sodium ou d'autres cations présents dans les solvants utilisés. En mode négatifs peu d'adduits sont formés.

- L'ionisation par réaction en phase gazeuse : beaucoup de molécules chargées sont générées et des transferts de protons peuvent avoir lieux entre des composés déjà protonés en solution vers des composés de plus forte basicité en phase gazeuse. Des analytes neutres qui s'évaporent peuvent donc être chargés en phase gaz.

- L'ionisation par oxydation ou réduction électrochimique au niveau du capillaire métallique : il y a conversion d'un analyte non chargé en ion.

Cette étape se prolongera ensuite lors de leurs transferts dans la partie suivante du Qexactive : un tube en métal long d'une dizaine de centimètre : le capillaire de transfert. Les ions seront chauffés et traverseront un gradient de pression décroissant (de la pression atmosphérique jusqu'à 1,5 Torr (2 millibar)) favorisant leur désolvatation. Le modèle expliquant les mécanismes d'ionisation ESI est intitulé « évaporation d'ions » et a été décrit par Iribane et Thomson dans les années 1970[207,208].

Une fois le capillaire de transfert parcouru les ions sont ensuite transférés au sein de la S-lens. Celle-ci est composée d'électrodes en acier inoxydable espacées progressivement. En son sein, les ions seront focalisés à l'aide de tensions radiofréquences appliquées jusqu'aux optiques ioniques (Injection et Inter-flatapole Lens) où ils continueront leurs focalisations. Ils sont ensuite transférés au sein du Bent Flatapole où ils seront guidés selon un arc de 90°. Cette étape est réalisée en vue d'éliminer les composés neutres qui auraient pu être transmis jusqu'à ce point, dû à la courbure, ils ne pourront être guidés. Avant d'entrer dans le quadripôle, les ions sont refocalisés à l'aide du flatapole exit-lens et passent à travers la split lens. Celle-ci contrôle l'entrée des ions au sein du quadripôle à l'aide d'une tension. En fonction de la tension appliquée les ions sont ou ne sont pas autorisés à entrer.

I.5.b) Principe d'un analyseur quadripôlaire

Les ions pénètrent ensuite au sein d'un analyseur quadripôle représenté ci-dessous (Figure 27). Son principe a été décrit la première fois par Paul et Steinwegen en 1953[209] :

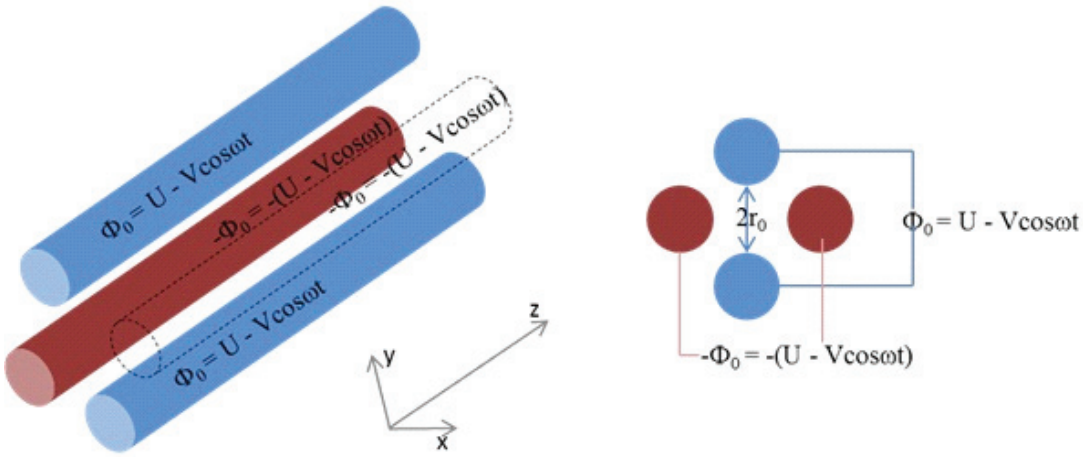


Figure 27 : Schéma d'un quadripôle avec les potentiels appliqués sur les 4 barreaux[210]

Le quadripôle comprend 4 barreaux hyperboliques parallèles ou circulaires, disposés selon les 4 points d'un losange, sur lesquels sont appliqués des potentiels positifs ou négatifs de manière alternative. Comme on peut le voir sur la figure ci-dessus (Figure 27), figé à un instant t les deux barreaux opposés (haut/bas) se voient apposer une tension similaire (positive) tandis que les deux autres (gauche/droite) se verront apposer la tension opposée (négative). A l'instant t+1 cela sera inversé et ainsi de suite.

Sous cet effet de tensions alternatives, les ions positifs vont adopter une trajectoire oscillante selon l'axe z. Les ions stables traverseront le quadripôle tandis que ceux instables iront percutés un des 4 barreaux.

Ci-dessous il est présenté une description théorique du fonctionnement d'un quadripôle :

Le potentiel ϕ_o (Equation 2) appliqué positivement ou négativement sur les 4 barreaux du quadripôle est composé d'une tension continue U et d'une tension alternative d'amplitude maximale V et de fréquence angulaire ω en radiant avec $\omega = 2\pi f$ ou f est la fréquence du champ alternatif.

$$\Phi_0 = U - V \cos\omega t \quad \textbf{Equation 2}$$

Lors de la migration des ions dans le quadripôle ceux-ci vont opter pour un mouvement particulier. Ce mouvement est la résultante des phénomènes d'accélération que vont subir les ions dû aux champs électriques appliqués. Suivant les axes x et y, deux équations ont été établies avec a pour l'accélération, m la masse de l'ion et -ze ou q la charge de l'ion en question (Equation 3 et 4).

$$Fx = ma = m\frac{d^{2x}}{dt^2} = -ze\frac{\delta\Phi}{\delta x} \quad \textbf{Equation 3}$$

$$Fy = ma = m\frac{d^{2y}}{dt^2} = -ze\frac{\delta\Phi}{\delta y} \quad \textbf{Equation 4}$$

S'en suit l'établissement de la fonction ϕ définit par rapport à ϕ_o et avec r_o le rayon du cercle inscrit entre les 4 barres du quadrupole (Equation 5 et 6) :

$$\Phi_{(x,y)} = \frac{\Phi_0 (x^2-y^2)}{r_0^2} \quad \textbf{Equation 5}$$

Soit :

$$\Phi_{(x,y)} = \frac{(U - V\cos\omega t)(x^2-y^2)}{r_0^2} \quad \textbf{Equation 6}$$

Celle-ci est dérivée et réarrangée avec les termes ci-dessous pour obtenir l'équation de Matthieu établie en 1866 en vue de la description de la propagation des ondes dans les membranes.

$$u = x \text{ ou } y,\ \varepsilon = \frac{\omega t}{2},\ a_u = \frac{8zeU}{mr_0^2\,\omega^2} \text{ et } q_u = \frac{4zeV}{mr_0^2\,\omega^2}$$

Avec :

$$a_u = a_x \text{ ou } -a_y \text{ et } q_u = q_x \text{ ou } -q_y$$

On obtient (Equation 7) :

$$\frac{d^2u}{d\varepsilon^2} + (a_u - 2\,q_u\cos 2\,\varepsilon)u = 0 \quad \textbf{Equation 7}$$

A partir de cette équation, il devient possible d'établir quels seront les ions qui pourront traverser ou non le quadripôle. Un diagramme de a_u en fonction de q_u peut-être tracé. Il permet de mettre en avant les zones syables pour un ion donné soit les valeurs des tensions U et V à appliquer pour que les valeurs de x et de y restent inférieures à r_o (Figure 28 a et b) soit les valeurs des tensions pour confiner l'ion au sein des 4 barreaux durant la totalité de son trajet au travers du quadripôle. En combinant les deux diagrammes il devient possible de définir intégralement cette zone (Figure 28 c).

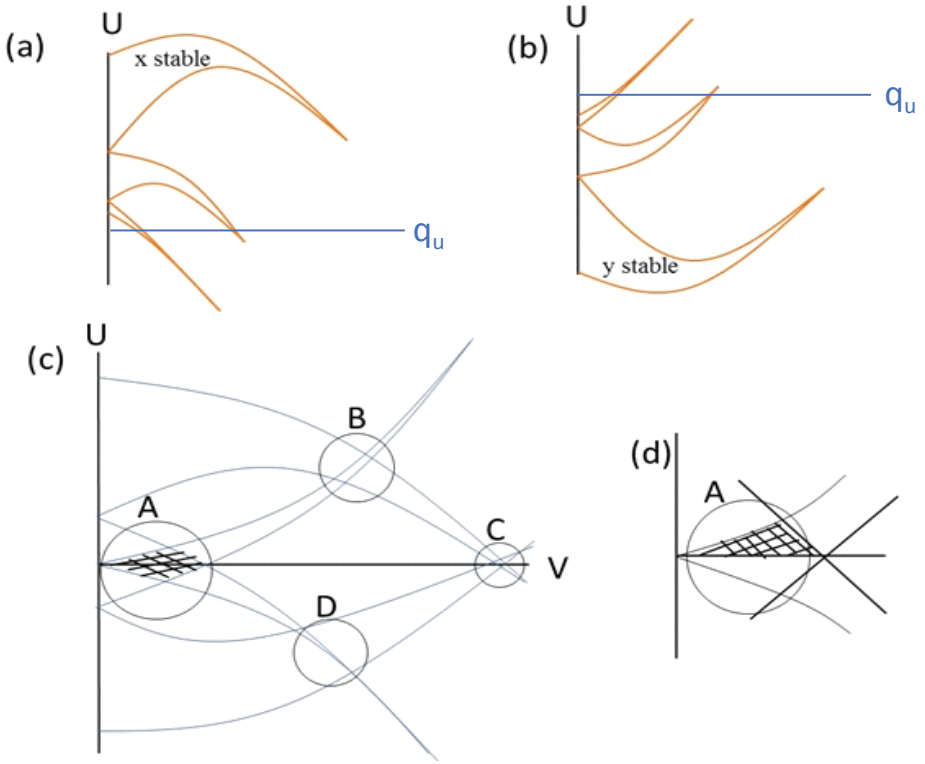


Figure 28 : Zones stables dans un quadripole d'un ion donné suivant l'axe x ou y (a et b). Les quatre zones de stabilités sont notées de A à D (c). La zone A est celle utilisée dans un quadripole (d). La partie hachurée correspond à un potentiel U positif[210]

Selon ce diagramme combiné, 4 zones stables émergent, les zones A, B, C et D. Pragmatiquement, c'est la zone A qui est la plus utilisée (Figure 28 d). En vue d'un balayage en masse dans un quadripôle, une attention particulière doit-être posée sur le rapport U/V à appliquer. Lors d'un balayage U et V vont varier simultanément et le rapport U/V sera gardé constant (Figure 29). Plus ce rapport sera élevé plus la sélectivité du filtre en masse sera importante signifiant une meilleure résolution. En d'autres termes, lorsque le coefficient directeur de la droite U/V augmente, la résolution de détection augmente. En contrepartie, la transmission des ions sera réduite signifiant une perte en sensibilité. Il est donc possible de jouer sur la sensibilité en augmentant ou en diminuant la résolution sur un analyseur type quadripôlaire.

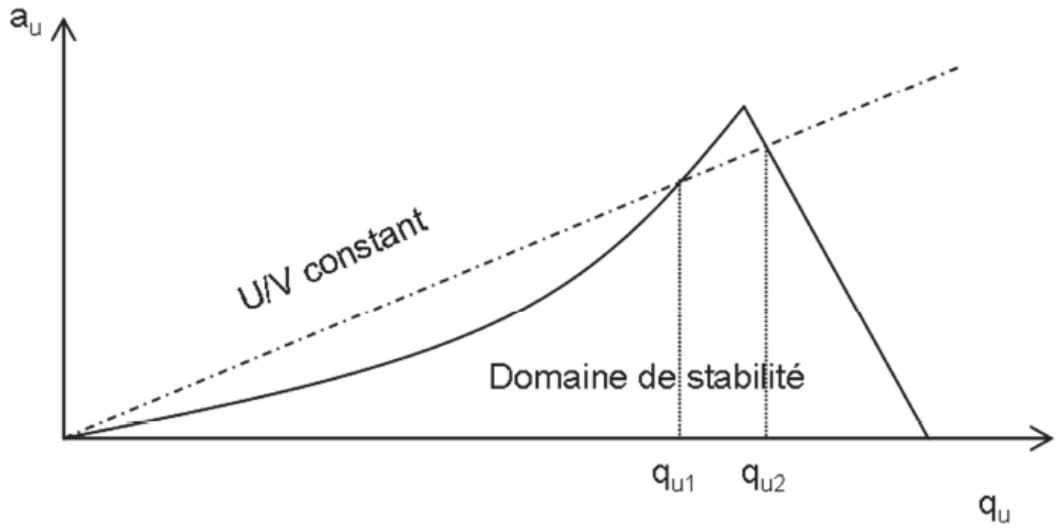


Figure 29 : Diagramme de stabilité en fonction de U et V donné pour un ion[204]

Le quadripôle, étant un analyseur séquentiel, permet l'analyse d'un seul ion. L'analyse des ions s'effectue masse après masse : une masse m_1 pendant une durée t_1, une masse m_2 pendant une durée t_2 et ainsi de suite. Les ions m_2 n'étant pas stables durant la durée t_1 ils sont perdus car non désirés. En corollaire, la vitesse de balayage de cet appareil est donc un paramètre critique puisque plus elle sera rapide plus d'ions pourront être analysés en un minimum de temps.

Les ions ayant été transmis (dû au paramétrage fixé pour le quadripôle) pourront continuer leurs routes au sein du Qexactive où ils seront transmis au sein de la C-Trap. Celle-ci est composée de 4 électrodes rectangulaires disposées en carré (1 électrode en haut, 1 autre en bas et 2 autres à gauche et à droite) confinant les ions arrivant par paquet au centre d'elles à l'aide de tensions radiofréquences. Les ions confinés peuvent se mouvoir librement tout au long de la trappe tout en subissant un effet de refroidissement collisionnel en rentrant en collision avec du gaz N_2. S'en suit un abaissement brutal des radiofréquences (< 200 ns) ne permettant plus le confinement. Immédiatement, des tensions continues sont appliquées pour envoyer les ions dans la cellule HCD en vue d'une étape de fragmentation ou directement au sein de l'orbitrap pour une analyse FMS. La cellule HCD sera d'abord présentée.

Il est à noter que plusieurs méthodes ont été proposées quant à l'injection des ions dans l'orbitrap : des lentilles électrostatiques[211] ou encore une injection axiale à l'aide d'un piège à ion linéaire[212]. La C-Trap par son injection radiale permet de minimiser l'énergie cinétique initiale des ions dans la direction z. C'est elle qui sera conservée.

I.5.c) La cellule HCD

La cellule HCD sera brièvement abordée ici. Les ions sont envoyés dans cette cellule pour être fragmentés par HCD. Succinctement la cellule HCD est un multipole fixé à l'intérieur d'un tube de

métal directement relié à la C-Trap. Les ions seront accélérés pour rentrer en collision avec le gaz neutre qu'elle contient (N_2). Une fois fragmentés ceux-ci seront ré-envoyés dans la C-trap pour être transférés à l'orbitrap en vue d'obtenir un spectre de fragmentation. Il est important de noter que l'avantage permis par la cellule HCD est l'affranchissement du « low-mass cutt-off » parfois limitant en CID.

I.5.d) L'orbitrap

Une fois que les ions sont expulsés de la C-Trap ils sont envoyés au sein de l'analyseur orbitrap par le biais de la Z-lens. L'orbitrap est un analyseur dans lequel les ions sont piégés électrostatiquement. Son principe a été présenté par Kindon en 1923[213]. Par la suite, dans les années 1990, Alexander Makarov a poursuivi le développement de ce concept en termes de géométrie et d'injection des ions menant au premier spectromètre de masse orbitrap commercial (Orbitrap XL, 2005) présenté par Thermo Scientific [211,214].

I.5.d.1) Principe de l'orbitrap

Voici un schéma de l'orbitrap fourni par Alexander Makarov (Figure 30) :

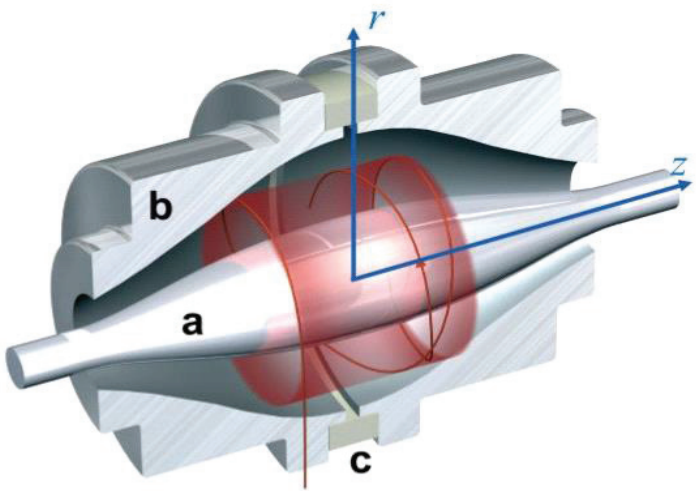


Figure 30 : Schéma d'une orbitrap[215]

L'orbitrap est composé de deux électrodes. L'une est centrale (Figure 30 a) en forme de fuseau et est placée coaxialement par rapport à une électrode creuse externe (Figure 30 b). La forme des électrodes découle de la forme des champs électriques. Les ions seront injectés tangentiellement à ce champ (en c suivant l'axe r). La stabilisation des ions est expliquée par l'équation de Matthieu suivante (Equation 8) :

$$U(r,z) = \frac{k}{2}\left(z^2 - \frac{r^2}{2}\right) + \frac{k}{2}(R_m)^2 \ln\left[\frac{r}{R_m}\right] + C \qquad \textbf{Equation 8}$$

r et z sont des coordonnées cylindriques, R_m : rayon caractéristique, R_1 : rayon d'électrode centrale, R_2 : rayon d'électrode creuse, C : constante, k : paramètre déterminé par la forme de l'électrode centrale et la tension appliquée.

Les ions injectés au sein de l'obitrap adopteront un mouvement en spirale (Figure 30 flèche rouge) autour de l'électrode centrale et seront séparés en fonction de leurs fréquences axiales d'oscillation. La détection s'effectuera à l'aide du courant induit par les oscillations mesurées de manière différentielle entre les électrodes. Cette fréquence suivant une transformation de Fourier permettra de remonter à la masse de ces mêmes composés (Equation 9) :

$$Z(t) = z_0 \cos(\omega t) + \left(\frac{2E_z}{k}\right)^{1/2} \sin(\omega t) \quad \textbf{Equation 9}$$

z_0 est l'amplitude axiale initiale. E_z : énergie cinétique initiale selon la direction z et (Equation 10) :

$$\omega = \left(\frac{kq}{m}\right)^{1/2} \quad \textbf{Equation 10}$$

Il est important de noter que selon la dernière équation présentée, la fréquence d'oscillation des ions est directement reliée à l'inverse du rapport masse sur charge ce qui permet de remonter aisément à celle-ci. Pour donner un ordre d'idée quant à la fréquence axiale des ions au sein d'un orbitrap, un ion de masse 542 Th aura une fréquence de 300 kHz soit environ 3 allers-retours le long de l'électrode centrale en 1 µs.

I.5.d.2) La résolution

L'orbitrap offre la possibilité d'analyse à très haute résolution. En point de comparaison, en vue d'estimer les capacités de résolution de l'orbitrap, nous pouvons prendre les travaux de Makarov effectués en 2006 où il a comparé les capacités de l'orbitrap et celle d'un FTICR[216]. Pour un ion à 400 *m/z*, lorsque l'orbitrap atteint une résolution de 60 000 le FTICR est à 100 000. Qui plus est, les ions peuvent être gardés plus longtemps au sein d'une cellule ICR engendrant la possibilité d'augmenter encore la résolution tandis que la cohérence est perdue plus rapidement dans l'orbitrap : augmenter le temps d'enregistrement ne mènera pas à une meilleure résolution. Cependant, la résolution de l'orbitrap est inversement proportionnelle à la racine carrée de la masse tandis que pour le FTICR elle est inversement proportionnelle à la masse (Equation 11) :

$$\frac{m}{\Delta m_{50\%}} = \frac{1}{2\Delta\omega_{50\%}} \left(\frac{kq}{m}\right)^{1/2} \quad \textbf{Equation 11}$$

Cela abouti au fait que pour des ions ayant un rapport masse sur charge supérieur à 1100 *m/z* et un même temps d'analyse, l'orbitrap offre une meilleure résolution.

Pour améliorer la résolution au sein d'un orbitrap on peut augmenter le champ électrique (Equation 12) :

$$\omega^2 = \frac{e}{m/z} \cdot \frac{2Ur}{R_m^2 \ln(\frac{R_2}{R_1}) - 0{,}5.[R_2^2 - R_1^2]}$$ **Equation 12**

Selon cette équation la fréquence des ions augmente proportionnellement avec la racine carrée de la tension appliquée. La vitesse d'analyse et la résolution peuvent ainsi être augmentées. Une autre méthode qui a permis d'augmenter la capacité de résolution de l'orbitrap a été de modifier la géométrie de la cellule (Figure 31). C'est ce qu'a réalisé Makarov en 2008[217] en appliquant un potentiel supérieur (5000 V contre 3500 V classiquement) :

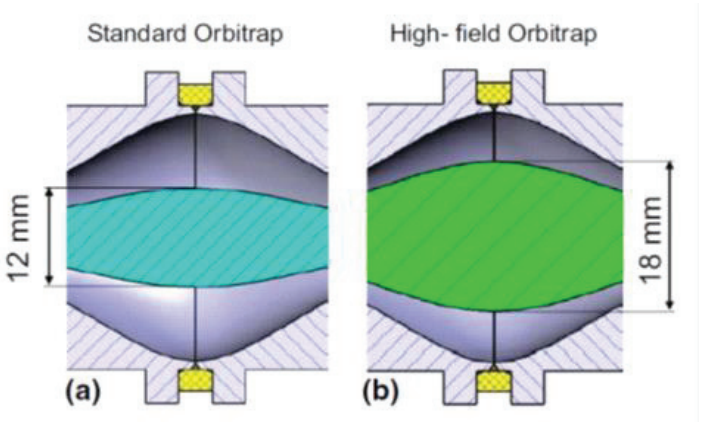


Figure 31 : Modification de la géométrie de la cellule entre une orbitrap standard et une orbitrap High-field[215]

Ce nouvel orbitrap a été introduit dans les modèles Qexactive à partir de l'Orbitrap Elite.

I.5.d.3) La précision en masse

La précision en masse va dépendre de la résolution qui sera choisie. Cependant, pour améliorer cette précision, le concept de « lock mass » a été introduit. Celui-ci repose sur la recherche d'une masse connue qui sera présente dans chaque spectre et qui servira à recalibrer les autres masses par rapport à la déviation subite par celle-ci même. Pour les analyses effectuées en protéomique l'ion 445,12005 est utilisé. Il s'agit d'un ion formé à partir des molécules de siloxanes présentes dans l'atmosphère. Il a été démontré que l'utilisation d'une lock masse permet d'augmenter la précision en analyse de 3,7 à 0,9 ppm[218] sans nuire à la vitesse d'analyse.

I.5.e) Conclusion sur le fonctionnement d'un Qexactive

Pour rappeler le fonctionnement du Qexactive de manière succincte : les ions en phase liquide provenant du système chromatographique sont d'abord ionisés à l'aide d'une source ESI puis sont transférés au sein d'un quadripôle qui laissera passer une certaine portion d'ions en fonction de la plage de *m/z* sélectionnée. Les ions transférés sont accumulés dans la C-Trap et sont transférés par paquets soit directement dans l'orbitrap pour obtenir un spectre FMS soit dans la cellule HCD pour être fragmentés. Une fois les ions fragments obtenus, ils sont retransférés dans la C-Trap pour être envoyés dans l'orbitrap en vue de l'obtention d'un spectre de fragmentation.

Ce qu'il est maintenant important de préciser c'est que le Qexactive travail en parallèle. Lorsque l'orbitrap est en train d'effectuer la mesure en fréquence d'oscillation pour un scan, les ions du scan suivant sont déjà en train d'être accumulés dans la C-Trap et/ou fragmentés, dans la cellule HCD. Cela permet de ne pas perdre de temps et de gagner en vitesse d'analyse.

I.5.f) Couplage du Qexactive avec un laser émettant à 473 nm

En vue de photo-fragmenter les peptides à cystéine dérivés avec le DABCYL C2-malémide, un laser émettant à 473 nm a été implémenté au sein du Qexactive disponible au laboratoire. Voici le montage (Figure 32) :

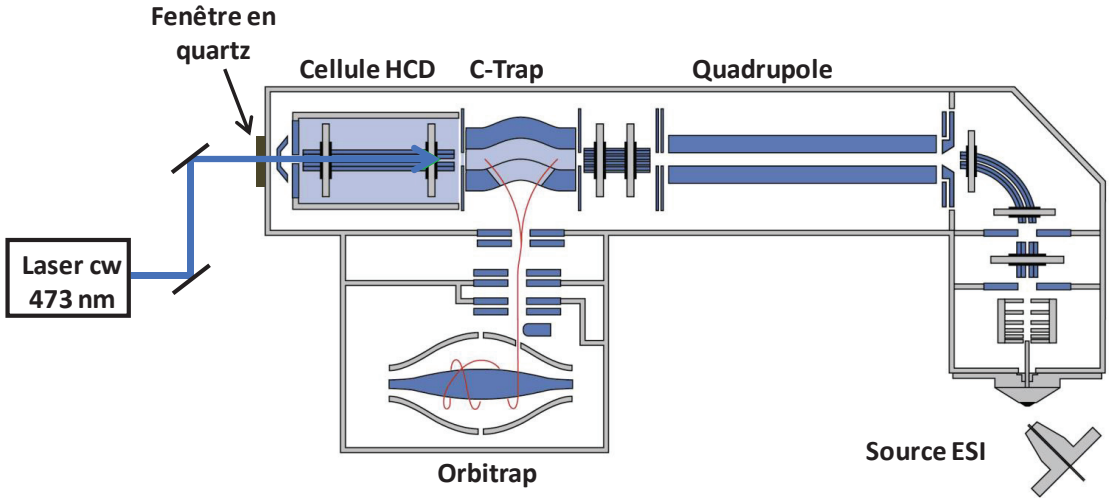


Figure 32 : Schéma d'un Qexactive couplé avec un laser irradiant à 473 nm dans la cellule HCD

La substitution de la fragmentation HCD par la méthode d'activation LID à 473 nm nécessite l'entrée du laser directement dans la cellule HCD, lieu initialement prévu pour la fragmentation. Pour ce faire, le Qexactive a tout d'abord été percé au niveau du côté gauche de la cellule HCD pour laisser pénétrer le laser. Celui-ci émettant dans le domaine du visible, une fenêtre en quartz a été disposée. En effet, ce matériau n'absorbant pas les rayonnements émis (dans ce domaine), il permet une bonne transmission. Il est important de préciser qu'au niveau d'entrée choisi pour le faisceau laser, un détecteur était présent. D'après le service de recherche et de développement

à Brême de Thermo Scientific, ce détecteur n'est utilisé que pour les maintenances lors des services après ventes et son perçage ou altération ne nuit aucunement aux manipulations effectuées sur la machine. Celui-ci a donc été percé. Ensuite, en vue de régler le trajet optique de manière optimale entre la sortie du laser et la cellule HCD, un jeu de miroir a été disposé. Dans le même temps, il a fallu s'assurer que la fragmentation, ayant lieu au sein de la cellule HCD, n'allait être induite qu'à la LID à 473 nm et non à une association LID/HCD. Idéalement, il faudrait que l'énergie d'activation par collision soit nulle. Or, le logiciel de base fourni par Thermo Scientific pour contrôler les paramètres du Qexactive, ne permet pas de fixer une valeur de « Normalized Collision Energy » (NCE) (utilisée pour fragmenter les composés au sein de la cellule HCD) inférieure à 10. Nous avons vu que pour cette valeur les peptides fragmentent encore. Une demande a donc été faite au constructeur pour obtenir plus de contrôle sur les valeurs de NCE. Cette demande a été acceptée nous donnant accès à la CE directe en eV. Nous avons pu déterminer qu'une valeur de CE de 2 eV permettait de transmettre les ions au sein de la cellule tout en n'induisant aucune fragmentation par HCD. En dessous de 2 eV, les ions ne sont pas transmis et le signal est perdu. Dans le même temps, nous avons pu avoir accès au temps d'activation. Celui-ci est habituellement fixé à une valeur de 3 ms en HCD. Dorénavant nous pouvons le modifier selon le besoin.

Durant toute ma thèse un seul laser a été utilisé : il s'agit d'un laser à diode MBL-FN-473 produit par la société CNI. Celui-ci émet un rayonnement continu à une longueur d'onde de 473 nm à +/- 1 nm avec une puissance pouvant aller de 0 à 500 mW et un diamètre de 1,5 mm ayant une divergence de 1 mrad. Le choix de se tourner vers un laser continu plutôt que pulsé (plus puissant) vient d'une part du prix d'achat plus important pour ces derniers et du fait de la difficulté à synchroniser de manière adéquate le pulse du laser et l'injection des ions dans la cellule HCD.

I.5.f) Paramétrage du spectromètre de masse

Toutes les manipulations effectuées par la suite auront été réalisées sur ce montage par le biais de ces 4 modes : FMS, PRM, DDA et DIA.

Le choix du paramétrage, pour ces 4 modes, a été établi en vue d'obtenir au minimum 8 points par pic chromatographique estimé à 20 s en moyenne. Dans cette optique, il a fallu régler différents paramètres de manière cohérente pour optimiser le temps de cycle :

Le temps d'injection (IT) est un paramètre clé à régler lors de l'optimisation du temps de cycle pour n'importe quel mode du spectromètre. Pour être expliciter, il doit être mis en relation avec un autre paramètre : le contrôle automatique de gain (AGC Target). L'AGC Target correspond à la quantité d'ions maximale qui sera collectée dans la C-Trap pour chaque scan du chromatogramme. L'ITmax (IT maximale fixée) correspond au temps maximal pour cette collecte. Lors de cette collecte, si l'AGC Target est atteint avant la fin de l'ITmax alors la collecte sera stoppée et l'ITexp (IT expérimental) sera plus court que l'ITmax. En revanche, si l'AGC Target n'est

pas atteint au bout de l'ITmax alors la collecte est aussi stoppée mais cette fois-ci la quantité d'ions accumulée sera moindre que celle fixée. Il est possible de savoir, pour chaque point, lequel des deux paramètres était limitant en aval de l'analyse mais en amont cela est impossible : cela dépend de la composition à l'élution en chaque instant du chromatogramme et ne peut donc être prédit. Le minimum de 8 points par pic est alors calculé pour un IT max atteint.

L'ITmax sera choisi en fonction du temps nécessaire à l'orbitrap pour analyser les ions à la résolution souhaitée. En effet, il est important de repréciser ici que le Qexactive travail en parallèle, c'est-à-dire que pendant que l'orbitrap analyse des ions pour un scan, les ions du scan suivant sont déjà en train d'être accumulés dans la C-Trap ce qui permet de gagner en vitesse d'analyse. Pour optimiser le temps de cycle il faut donc corréler l'ITmax au temps d'analyse de l'orbitrap. Celui-ci dépend de la résolution fixée : plus la résolution est grande, plus le temps d'analyse de l'orbitrap est important. Voici les temps d'analyse en fonction de la résolution fourni par le constructeur (Tableau 6) :

Résolution	Temps d'analyse (ms)
140 000	512
70 000	256
35 000	128
17 500	64

Tableau 6 : Temps d'analyse de l'orbitrap en fonction de la résolution

Voici le paramétrage respectif optimisé utilisé pour chacun des modes[70] (Tableau 7) :

FMS	
Résolution	140 000
AGC Target	1,0E+06
IT max (ms)	250
Plage de *m/z* (Th)	150-1200

PRM - LID	
Résolution	70 000
AGC Target	5,0E+05
IT max (ms)	215
Plage de *m/z* (Th)	4
Temps d'activation (ms)	25

	DDA			DIA	
FMS	Résolution	35 000	FMS	Résolution	35 000
	AGC Target	1,0E+06		AGC Target	1,0E+06
	IT max (ms)	95		IT max (ms)	95
	Plage de *m/z* (Th)	150-1200		Plage de *m/z* (Th)	400-848
LID	Résolution	35 000	LID	Résolution	35 000
	AGC Target	5,0E+05		AGC Target	5,0E+05
	IT max (ms)	95		IT max (ms)	95
	Top N	10		Nombre de fenêtre	14
	Plage de *m/z* (Th)	4		Plage de m/z (Th)	32
	Temps d'activation (ms)	25		Temps d'activation (ms)	25
HCD	Résolution	35 000	HCD	Résolution	35 000
	AGC Target	5,0E+05		AGC Target	5,0E+05
	IT max (ms)	95		IT max (ms)	95
	Top N	10		Nombre de fenêtre	14
	Plage de *m/z* (Th)	4		Plage de m/z (Th)	32
	Temps d'activation (ms)	3		Temps d'activation (ms)	3
	NCE	28		NCE	28

Tableau 7 : Paramétrage du Qexactive fixé en fonction du mode utilisé

Un minimum de 8 points par pic chromatographique, estimé à 20 s en moyenne, a été contrôlé et validé pour chaque mode et pour chaque analyse manuellement.

Le temps d'activation de la LID à 473 nm a été fixé à 25 ms pour toutes les analyses ce qui a mené au choix des 32 fenêtres sans recouvrement de 14 Th en DIA.

La plage de *m/z* restreinte en DIA de 400 à 848 Th résulte à la fois d'un temps cycle trop important pour pouvoir sonder une plus large gamme, et, d'autre part, du fait que la majorité des ions produits en ESI le sont dans cette fenêtre de *m/z*.

I.6) Logiciels

Lors de ce chapitre les données brutes auront été visualisées avec le logiciel Xcalibur 2.2.48 (Thermo Scientific). La construction de la bilbiothèque de spectres aura été effectuée à l'aide du logiciel Mascot 2.5.1 (Matrix Science) et du logiciel Proline 1.4. Toutes les données auront été traitées à l'aide du logiciel Skyline 3.7 et de Excel.

II) Présentation générale de la dérivation et de la photo-fragmentation LID à 473 nm des peptides à cystéine dérivés

Avant d'aborder la construction de la libraire spectrale, un récapitulatif quant à la dérivation et à la photo-fragmentation laser à 473 nm des peptides à cystéine dérivés avec le DABCYL C2-maléimide sera présenté dans cette section. Il sera introduit toutes les informations relatives à la dérivation, la stabilité et à la photo-fragmentation laser à 473 nm des peptides à cystéine dérivés avec le DABCYL C2-maléimide. Toutes les expérimentations ont été effectuées à l'aide des peptides à cystéine synthétiques commandés chez Thermo Scientific.

II.1) Dérivation des peptides à cystéine en solution

En ce qui concerne l'étape de dérivation, celle-ci avait déjà été mise en place et validée par le passé au laboratoire [9]. Tous les peptides à cystéine sont dérivés, suivant le protocole établi dans la section I.3 de ce chapitre, au niveau de leur acide aminé cystéine sur leur fonction thiol via le groupement maléimide présent sur le chromophore DABCYL C2-maléimide (Figure 33) :

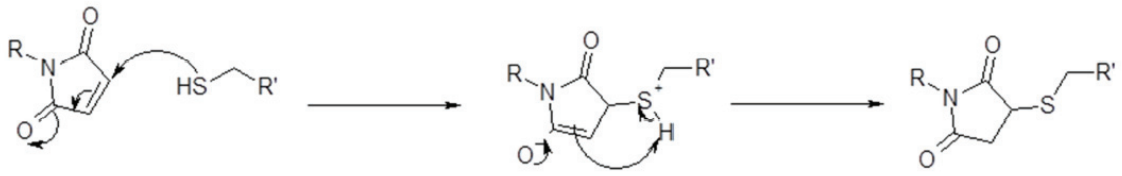


Figure 33 : Dérivation de la fonction thiol à l'aide d'un groupement maléimide

Les peptides à cystéine ainsi obtenus présentent un incrément de masse de 391,17 Th correspondant au chromophore DABCYL C2-maléimide (Figure 34) :

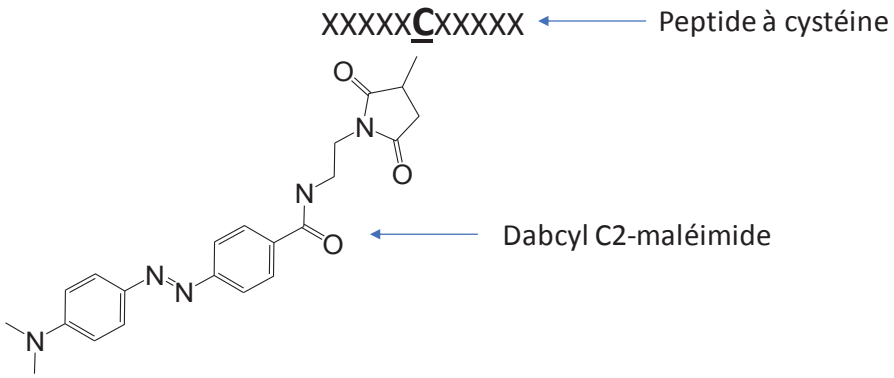


Figure 34 : Peptide à cystéine dérivé avec le chromophore DABCYL C2-maléimide sur l'acide aminé cystéine

Ci-dessous est présenté, à travers l'exemple de 2 peptides à cystéine dérivés (NILWSAENECFK, ACDALYPPGPPSNPK) l'efficacité de cette étape de dérivation mais aussi les questions résiduelles à son niveau.

Les 2 peptides à cystéine exposés ici ont été solubilisés dans une matrice aqueuse puis traités avec le protocole type de dérivation (échantillon A), séparés avec le gradient 1 puis analysés en mode FMS. Deux échantillons témoins ont également été réalisés, ceux-ci contenant les peptides purs purifiés mais non soumis au protocole de dérivation (échantillon B).

Voici les résultats présentés pour le peptide à cystéine NILWSAENECFK (Figure 35) :

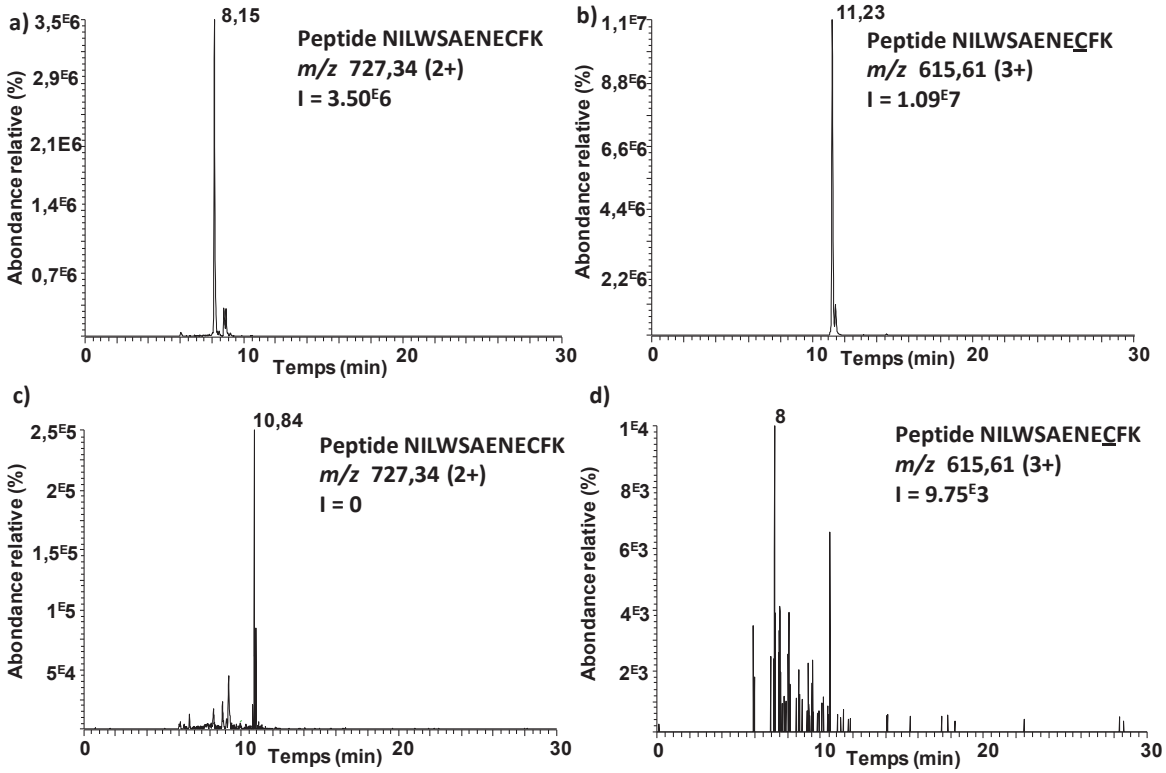


Figure 35 : XIC du peptide à cystéine NILWSAENECFK non dérivé dichargé dans l'échantillon non traité (B) (a) et traité avec le DABCYL C2-maléimide (A) (c), dérivé trichargé dans l'échantillon traité avec le DABCYL C2-maléimide (A) (b) et non traité (B) (d)

Nous pouvons voir que lorsque le peptide NILWSAENECFK n'est pas dérivé celui-ci est observé à 8,15 min sous forme doublement protoné ($M+2H^{2+}$) de *m/z* 727,34 comme le montre le chromatogramme extrait (XIC) de cet ion (Figure 35 a). Lorsque le peptide NILWSAENECFK est dérivé, en revanche, l'ion correspondant, triplement protoné, est bien observé à 11,23 min comme le montre le XIC de l'ion *m/z* 645,61 (Figure 35 b). On peut aussi voir que le peptide non dérivé n'est plus présent dans l'échantillon traité avec le protocole de dérivation avec une intensité à 0 pour le XIC de l'ion de *m/z* 727,34 correspondant à 8,15 min (Figure 35 c). Un ion

précurseur interférent est noté à 10,84 min mais il n'est pas relatif à NILWSAENECFK. Nous avons pu également vérifier que la forme dérivée n'est pas présente, dès le départ, dans l'échantillon non traité (Figure 35 d). La dérivation semble totale. Ne pouvant comparer les rendements d'ionisation des 2 formes du peptide (dérivée et non dérivée), une estimation du rendement de dérivation (R) peut être établie en comparant la réponse du peptide non dérivé dans les 2 types d'échantillons suivant cette équation (Equation 13) :

$$R = \frac{\text{I peptide non dérivé échantillon B} - \text{I peptide non dérivé échantillon A}}{\text{I peptide non dérivé échantillon B}} \text{ X 100} \qquad \textbf{Equation 13}$$

Ici il est de 100 %. Cependant, l'estimation de ce rendement de dérivation ne peut être pleinement affirmée. Un autre comportement porteur de questionnement a été observé pour le peptide à cystéine ACDALYPPGPPSNPK (Figure 36) :

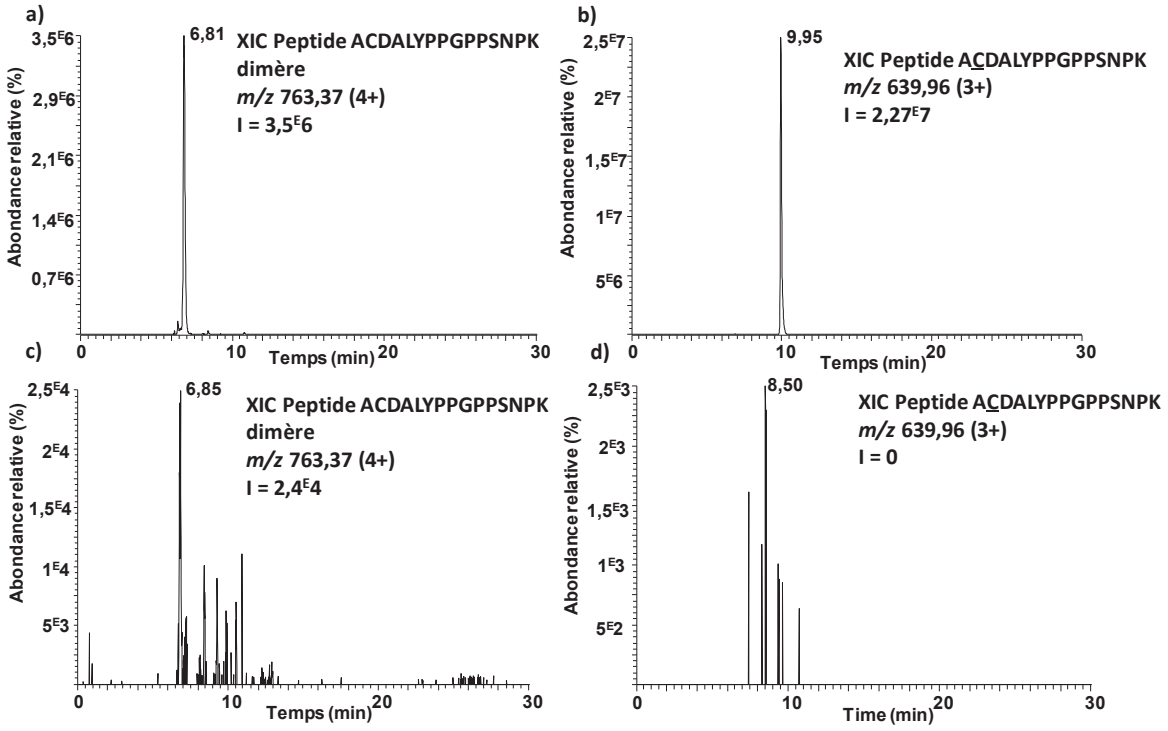


Figure 36 : XIC du peptide à cystéine ACDALYPPGPPSNPK non dérivé dimérisé dichargé dans l'échantillon non traité (B) (a) et traité avec le DABCYL C2-maléimide (A) (c), dérivé trichargé dans l'échantillon traité avec le DABCYL C2-maléimide (A) (b) et non traité (B) (d)

Ici, lorsque le peptide ACDALYPPGPPSNPK n'est pas dérivé il n'est pas présent sous forme de monomère mais sous forme de dimère. Ceci est dû à la reformation des ponts disulfures entre 2 acides aminés cystéines comme le montre le XIC de l'ion précurseur quadrichargé *m/z* 763,37 du dimère avec un pic à 6,81 min (Figure 36 a). Dans ce cas-ci, le dimère a quasiment disparu après dérivation : il est encore présent à 6,85 min mais à une intensité très faible de $2,4^{E}4$ (Figure 36 c). Lorsque le peptide A<u>C</u>DALYPPGPPSNPK est dérivé, le comportement attendu est observé : l'ion

triplement chargé *m/z* 639,96 est observé dans l'échantillon traité avec le protocole de dérivation avec un pic à 9,95 min (Figure 36 b) et absent dans l'échantillon non traité (Figure 36 d). Le rendement de dérivation est optimal (99,4%).

Cependant, il faut savoir que dans des milieux plus complexes, une multitude de composés possédant des fonctions thiols pourront se dimériser avec les peptides à cystéine ciblés nuisant ainsi à toute estimation potentielle quant au rendement de dérivation.

Dans le même temps, il est important de préciser que les peptides à cystéine commandés chez Thermo Scientific sont des peptides de basse qualité (peptide grade 2). Une multitude de composés peuvent être présents en plus des peptides attendus au sein des échantillons : les peptides à cystéine attendus tronqués, des acides aminés résiduels, des éléments inconnus etc…Ces composés aussi peuvent présenter un acide aminé cystéine ou une fonction thiol avec lequel le peptide non dérivé aurait pu former des dimères et ainsi biaiser l'estimation précédente puisque non incorporés à celle-ci. Le rendement de dérivation semble donc optimal mais ne peut être que spéculer à proprement parler.

Outre le rendement de dérivation, deux autres observations sont à mettre en avant après ces premières analyses. Tout d'abord, nous pouvons voir que le signal du peptide dérivé est plus important que celui du non dérivé (Figure 35 a et b ; Figure 36 a et b). Il semble que grâce à la dérivation un gain en sensibilité est obtenu grâce à l'amine tertiaire du chromophore qui est facilement protonable (en mettant de côté les observations faites relativement au phénomène de dimérisation). D'autre part, un recul en RT résultant de la dérivation est observé (Figure 35 a et b ; Figure 36 a et b). Les peptides à cystéine dérivés sont plus hydrophobes que les peptides à cystéine non dérivés de par l'hydrophobicité du chromophore.

Après ces premières expérimentations, nous pouvons affirmer que le protocole de dérivation fonctionne efficacement, les 2 peptides à cystéine ont pu être dérivés et détectés. La répétabilité de ce protocole a déjà été démontré par le passé [14].

II.2) Etude de stabilité des peptides à cystéine dérivés

Une fois les peptides à cystéine dérivés, il a fallu estimer leur stabilité en vue d'établir une durée limite de conservation ainsi que l'environnement (température ambiante, frigidaire et congélateur) le plus à même de les garder indemnes. Cette étape a pu être réalisée en amont de ma thèse durant mon stage de deuxième année de master. Cette étude s'était basée sur le suivi de deux hormones possédant des fonctions thiols (Figures 37 et 38) :

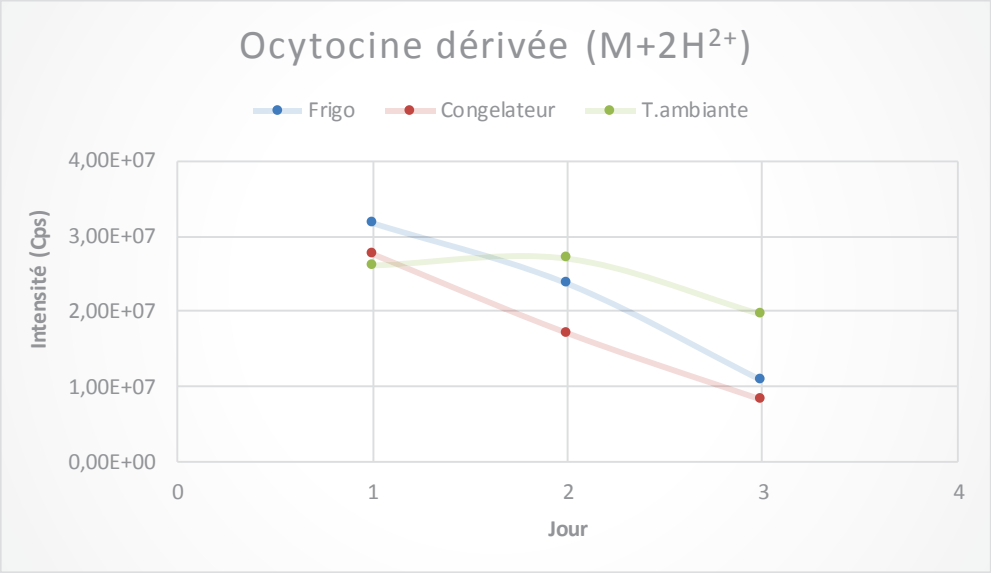


Figure 37 : Etude de stabilité de l'ocytocine dérivée dans 3 milieux de conservations différents durant 3 jours

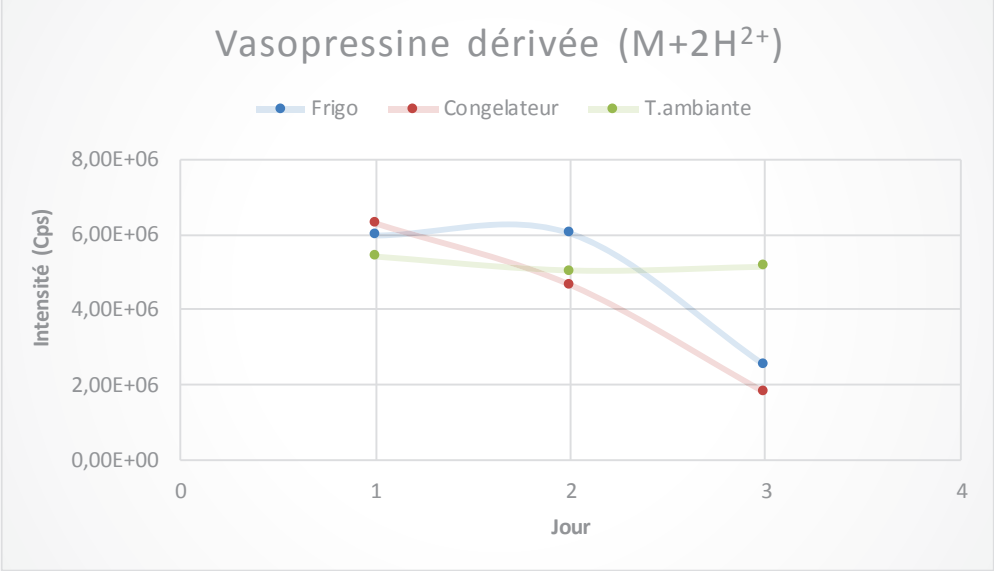


Figure 38 : Etude de stabilité de la vasopressine dérivée dans 3 milieux de conservations différents durant 3 jours

A travers les figures 37 et 38 il est présenté l'évolution de l'intensité du signal de l'ion dichargé correspondant à l'ocytocine et à la vasopressine durant 3 jours dans 3 milieux de conservation différents (température ambiante, frigidaire et congélateur).

Ce qui a pu être mis en avant durant ces 3 jours était que, la stabilité des composés dérivés dépendait à la fois du milieu et des composés en eux-mêmes. Nous pouvons voir par exemple que la vasopressine se conserve correctement à température ambiante a contrario de l'ocytocine (l'intensité du signal de l'ion dichargé correspondant à l'ocytocine a diminué en jour 3 contrairement à celui de la vasopressine) (Figures 37 et 38). Egalement, de façon assez

surprenante, on peut voir que la vasopressine est encore stable 3 jours après dérivation à température ambiante mais qu'elle ne l'est pas lorsqu'elle est placée au congélateur (Figure 38). Une variabilité intra et inter-composé en fonction des milieux a pu être observée. Ces remarques induisent que, pour toutes analyses comparatives, les échantillons, après dérivation, doivent être gardés dans le même milieu et surtout analysés après le même laps de temps de conservation.

Devant les résultats obtenus et cette variabilité intra et inter-composé le choix le plus simple a été fait : il a été remarqué qu'après 1 jour de conservation, pour les 3 milieux de conservation, l'intensité du signal des ions était équivalent pour les deux hormones (CV inférieurs à 10 et 7 % pour l'ocytocine et la vasopressine, respectivement) (Figure 37 et 38). Il a donc été décidé d'analyser tous les échantillons quels qu'ils soient dès le premier jour sans exception et qu'il serait conservé dans un même milieu : le congélateur.

II.3) Photo-fragmentation laser à 473 nm des peptides à cystéine dérivés avec le DABCYL-C2 maléimide

L'identification des peptides à cystéine dérivés avec le DABCYL-C2 maléimide se fait grâce à leur comportement de photo-fragmentation induite par laser à 473 nm. Une étude plus exhaustive, sur la photo-fragmentation des peptides à cystéine dérivés par irradiation laser à 473 nm sera présentée dans le chapitre III, ici il ne sera question que d'une présentation synthétique.

Il est important de préciser, qu'avant toute analyse par photo-fragmentation LID à 473 nm, un contrôle est effectué au niveau du couplage Qexactive-laser pour s'assurer du bon alignement de ce dernier au sein de la cellule HCD. Le recouvrement entre le faisceau de photons et le nuage d'ions en phase gazeuse se doit d'être optimal. Pour ce faire, le chromophore DABCYL-C2 maléimide a été solubilisé dans du MeOH à une concentration d'environ 50 µg/mL puis infusé seul à un débit de 10 µL/min. Il est ensuite analysé en mode PRM-LID à 473 nm avec le laser éteint ou allumé avec un temps d'activation de 3 ms (Figure 39) :

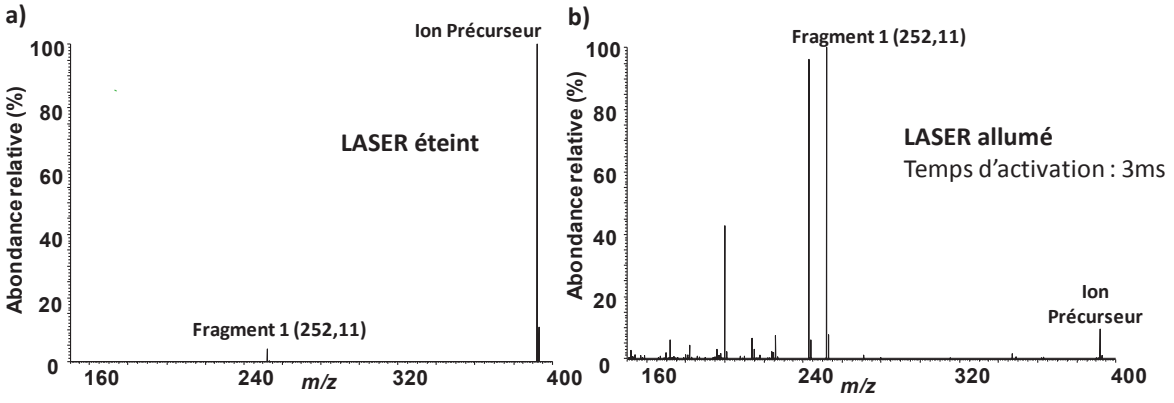


Figure 39 : Comportement de l'ion précurseur monochargé (*m/z* 392.17) du DABCYL C2-maléimide en PRM-LID à 473 nm lorsque le laser est éteint (a) ou allumé (temps d'activation 3 ms) (b)

Lorsque le laser n'est pas allumé il ne doit y avoir aucune fragmentation visible au sein du spectre obtenu en raison de l'absence du laser ainsi que de la faible CE appliquée (2 eV). Nous pouvons voir que pour une CE de 2 eV les ions sont bien transmis dans la cellule HCD comme assuré par Thermo Scientific mais aussi qu'une légère fragmentation résiduelle pour cet ion précurseur peu stable a quand même lieu (Figure 39 a). Un seul ion fragment est produit (fragment 1 *m/z* 252,11) avec une intensité relative de 3-4 % ce qui n'est pas problématique : la perte de signal est minime.

Lorsque le laser est allumé (Figure 39 b), un spectre de photo-fragmentation du DABCYL-C2 maléimide est obtenu, résultant de la photo-fragmentation de l'ion précurseur, avec 7 ions fragments dont 1 majoritaire à *m/z* 252,11. Ensuite, comme il le sera mentionné et explicité par la suite, nous pouvons voir que le taux de photo-fragmentation n'est pas de 100 %. Ici, relativement à l'ion fragment le plus intense, le signal de l'ion précurseur n'est plus que de 15 %.

Ce spectre de photo-fragmentation du chromophore DABCYL-C2 maléimide à 3 ms fait office de référence (réglage optimal du laser). Avant chaque analyse un spectre identique devra être obtenu après réglage du laser.

Il est important de préciser ici que l'ion fragment 1 *m/z* 252,11 cité précédemment est un ion dit rapporteur (Figure 39 b). Cet ion est un fragment caractéristique du chromophore lorsque soumis à l'irradiation laser à 473 nm seul ou lié à un édifice moléculaire (Figure 40) :

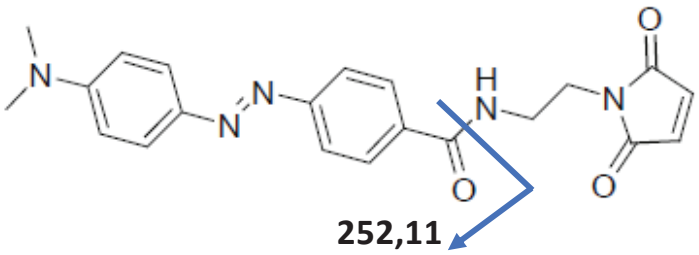


Figure 40 : Structure du DABCYL C-2 maléimide avec localisation du site de fragmentation générant l'ion rapporteur *m/z* 252,11

Le chromophore seul et tous les composés dérivés avec le DABCYL-C2 maléimide (dont les peptides à cystéine) vont produire cet ion rapporteur qui agira en tant que marqueur sur un spectre de fragmentation. En effet, dès que cet ion sera présent sur un spectre de fragmentation, alors le composé fragmenté aura précédemment été dérivé avec le DABCYL-C2 maléimide. Cet ion est, sur la majorité des spectres, le plus intense qui est généré.

En ce qui concerne la photo-fragmentation laser des peptides à cystéine dérivés avec le DABCYL-C2 maléimide à 473 nm, le spectre de photo-fragmentation du peptide à cystéine dérivé YCSVTTQADGR est exposé ci-dessous. Celui-ci a été solubilisé dans une matrice aqueuse, dérivé à l'aide du protocole type, séparé avec le gradient 1 puis analysé en mode PRM-LID à 473 nm.

La majorité des spectres de photo-fragmentation LID à 473 nm pour des peptides à cystéine dérivés avec le DABCYL-C2 maléimide se présentent de la sorte (Figure 41) :

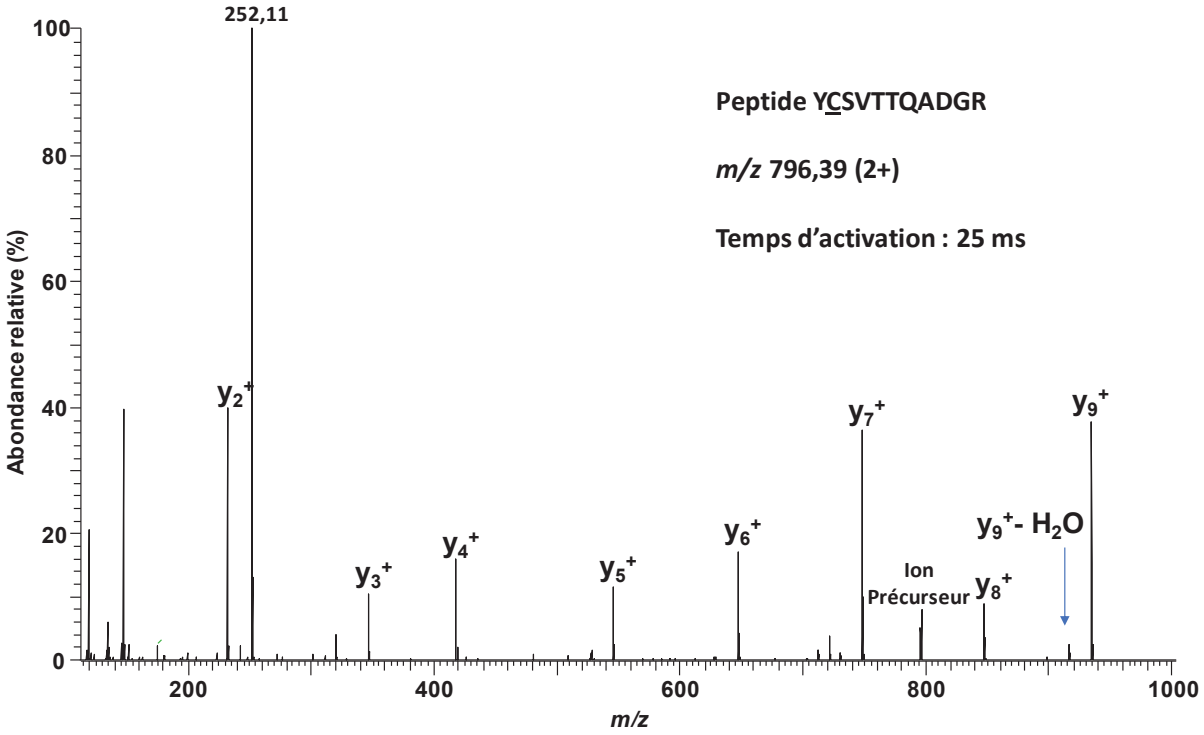


Figure 41 : Spectre de photo-fragmentation de l'ion dichargé *m/z* 796,39 du peptide à cystéine dérivé YCSVTTQADGR obtenu en PRM LID à 473 nm avec un temps d'activation de 25 ms

Plusieurs observations peuvent être faites sur ce spectre. Tout d'abord, ce sont principalement des fragments y qui sont générés. On y retrouve aussi des pertes d'eau. Les fragments produits sont similaires à ceux obtenus en CID (b et y majoritairement) comme nous le verrons dans le chapitre III. Nous pouvons également observer l'ion rapporteur *m/z* 252,11, qui est l'ion fragment le plus intense généré sur ce spectre. Ensuite, comme vu pour le DABCYL-C2 maléimide auparavant, le taux de photo-fragmentation n'est pas total : l'ion précurseur est présent au sein de ce spectre. Si cela a un impact négatif en termes de sensibilité, cette information peut néanmoins servir à confirmer la bonne identification des peptides à cystéine suspectés lors d'analyse PRM ou DDA-LID à 473 nm.

II.4) Temps d'activation laser des peptides à cystéine dérivés avec le DABCYL-C2 maléimide à 473 nm

En vue d'obtenir des spectres de photo-fragmentation optimaux permettant l'identification des peptides à cystéine dérivés, le temps d'activation nécessaire à leur bonne photo-fragmentation a été étudié. Il y a quelques années, lors d'une précédente étude, un temps d'activation de 25 ms

permettait une photo-fragmentation optimale[9]. Cette conclusion a néanmoins été redémontrée à l'aide du peptide à cystéine dérivé LCDSGELVAIK qui a été irradié à quatre temps d'activation différents (100, 75, 50 et 25 ms). Ce peptide a été solubilisé dans une matrice aqueuse, dérivé à l'aide du protocole type, séparé avec le gradient 1 puis analysé en mode PRM-LID à 473 nm (Figure 42).

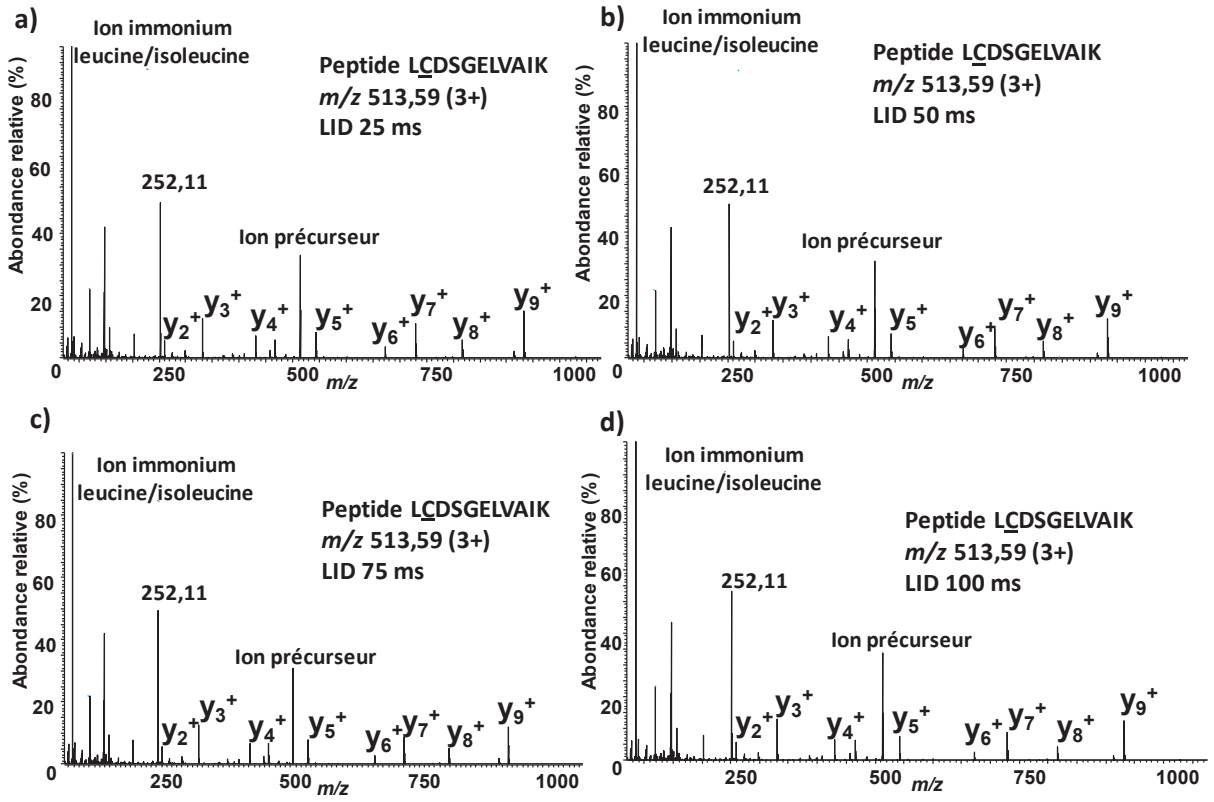


Figure 42 : Spectre de photo-fragmentation de l'ion trichargé *m/z* 513,59 du peptide à cystéine dérivé LCDSGEL-VAIK en PRM LID à 473 nm à 4 temps d'activation différents 25 ms (a) 50 ms (b) 75 ms (c) et 100 ms (d)

Il est ici bien observé que pour les 4 temps d'activation différents testés le même spectre de photo-fragmentation est obtenu. Les intensités relatives entre tous les ions fragments sont conservées quel que soit le temps d'irradiation utilisé. De ce fait, le temps d'activation de 25 ms a été conservé.

II.5) Taux de photo-fragmentation des peptides à cystéine dérivés avec le DABCYL-C2 maléimide à 473 nm

Comme vu précédemment le taux de photo-fragmentation n'est pas total. En vue d'éclaircir ce point, 2 aspects ont été abordés :

Tout d'abord, il a été calculé le taux de photo-fragmentation pour 10 peptides à cystéine dérivés afin d'établir une estimation d'un taux moyen. Les peptides à cystéine ont été solubilisés dans

une matrice aqueuse, dérivés avec le protocole type puis séparés à l'aide du gradient 1 et analysés en mode PRM-LID à 473 nm.

Ci-dessous est présenté, pour les 10 peptides, le taux de photo-fragmentation (Equation 14 et Tableau 8) :

$$\text{Taux de photo-fragmentation} = \frac{\text{E Ifrag}}{\text{E Ifrag} + \text{Iparent}} \text{ X 100} \quad \textbf{Equation 14}$$

Avec I_{frag} : l'intensité des ions fragments et I_{parent} : l'intensité de l'ion précurseur

Sequence	Taux de photo-fragmentation (%)
FSPEFVDFTAQCLR	98
GYPSEFATYLNFCR	99
GYPSEFSTYLNFCR	98
NDYTQGCEPGHR	92
VVECLDHAR	96
SLAYIHSQGVCHR	99
LCDSGELVAIK	95
ACDGRPYAGAVQK	93
CLDEGGYYISPR	98
AGLECSDIEK	99

Tableau 8 : Taux de photo-fragmentation obtenus pour 10 peptides à cystéine dérivés

Nous pouvons observer que le taux de photo-fragmentation n'est pas total pour les 10 peptides à cystéine dérivés mais qu'il est très élevé : 92 % au minimum. Nous pouvons donc conclure que le taux de photo-fragmentation atteint n'est jamais de 100 % mais qu'il est suffisamment haut pour que les populations d'ions précurseurs soient photo-fragmentées quasiment en totalité.

Ensuite, sur ces 10 spectres, les couvertures de séquence associées ont été calculées en vue de s'assurer que les spectres de photo-fragmentation étaient assez riche pour pouvoir identifier des peptides à cystéine par leurs intermédiaires. (Equation 15 et Tableau 9) :

$$\text{Couverture de séquence} = \frac{\textit{Nombre de fragments non redondants observés}}{\textit{Nombre d'acide aminé composant la séquence peptidique}} \text{ X 100} \quad \textbf{Equation 15}$$

Dans cette équation le terme non redondant signifie que tous les fragments de charge différentes situés à une même location (par exemple y_5^+ et y_5^{++}) pour un même peptide, valent pour 1.

Sequence	Couverture de séquence (%)
FSPEFVDFTAQCLR	71
GYPSEFATYLNFCR	100
GYPSEFSTYLNFCR	100
NDYTQGCEPGHR	50
VVECLDHAR	100
SLAYIHSQGVCHR	100
LCDSGELVAIK	73
ACDGRPYAGAVQK	54
CLDEGGYYISPR	83
AGLECSDIEK	60

Tableau 9 : Couvertures de séquence obtenues pour les 10 peptides à cystéine dérivés

Les informations extraites indiquent que la couverture de séquence moyenne obtenue est de 80 % avec un minimum à 50 % pour le peptide à cystéine dérivé NDYTQGCEPGHR et un maximum à 100 % pour 4 peptides à cystéine dérivés. Il est à noter que les peptides à cystéine dérivés NDYTQGCEPGHR et ACDGRPYAGAVQK, qui possèdent les couvertures de séquence les plus petites, sont aussi ceux qui ont un taux de photo-fragmentation le plus bas (92 et 93 %). Cependant, il ne semble pas qu'un taux de photo-fragmentation élevé assure une couverture de séquence importante comme le démontre les peptides à cystéine dérivés AGLECSDIEK ou FSPEFVDFTAQCLR avec des taux de photo-fragmentation supérieurs à 98 % et des couvertures de séquence inférieures à 72 %. Il semble donc que d'autres paramètres influent ici (ils seront abordés en partie dans le chapitre III). Néanmoins, avec une couverture de séquence moyenne supérieure à 80 %, il est démontré que la méthode d'activation LID à 473 nm engendre des spectres de photo-fragmentation suffisamment riches pour identifier des peptides à cystéine dérivés avec le DABCYL C2-maléimide.

Les taux de photo-fragmentation élevés obtenus et les couvertures de séquence offertes prouvent que la photo-fragmentation laser à 473 nm est bien une méthode d'activation efficace pour identifier des peptides à cystéine dérivés avec le DABCYL-C2 maléimide en spectrométrie de masse.

II.6) Comparaison entre la méthode d'activation LID à 473 nm et la fragmentation HCD sur des peptides à cystéine dérivés avec le DABCYL-C2 maléimide

Une fois le rendement de dérivation et l'efficacité de photo-fragmentation des peptides à cystéine dérivés avec le DABCYL-C2 maléimide à 473 nm évalués, une étude comparative a été menée entre la méthode d'activation HCD et LID à 473 nm afin de visualiser les apports/limites de la méthode par rapport à celle de référence.

Dans cette étude, les 10 mêmes peptides à cystéine que précédemment ont été solubilisés dans une matrice aqueuse, dérivés avec le protocole type puis séparés à l'aide du gradient 1 et analysés en mode PRM-LID à 473 nm et PRM-HCD. Les résultats obtenus pour les 10 peptides à cystéine dérivés peuvent être représentés par les peptides à cystéine dérivés CTEALYAPAEGDK et YCSVTTQADGR (Figure 43) :

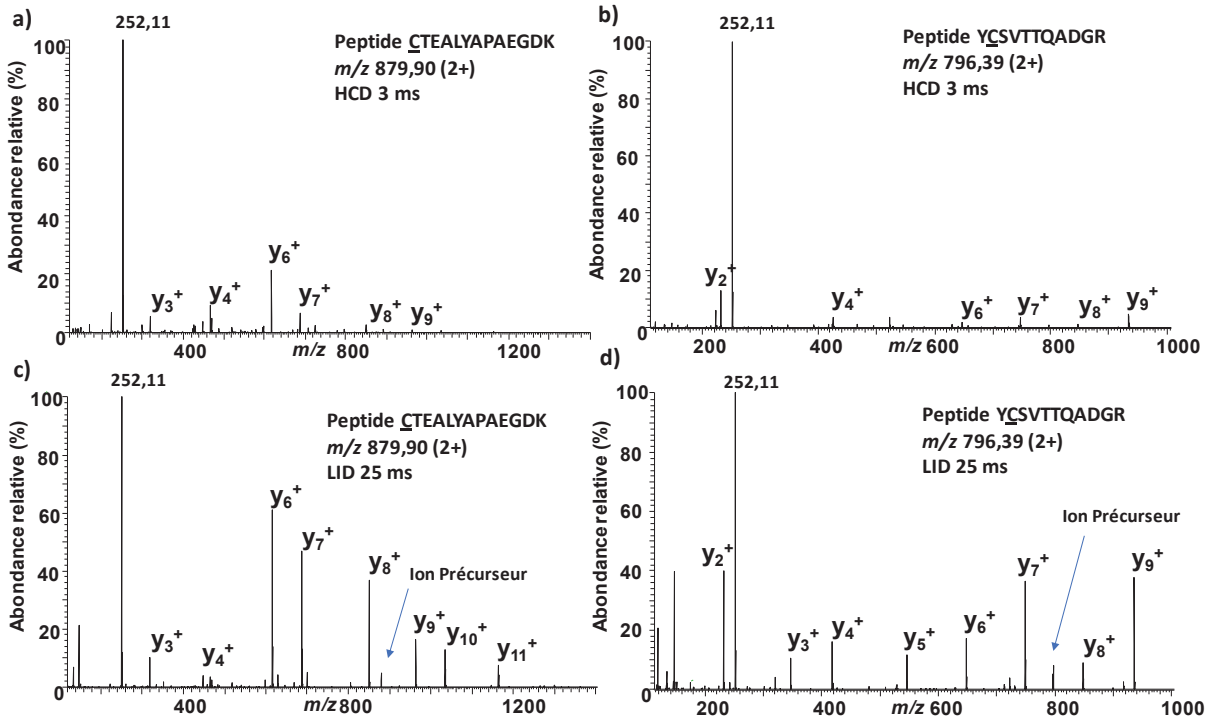


Figure 43 : Spectres de fragmentation des ions dichargés *m/z* 879,90 et 796,39 des peptides à cystéine dérivés CTEALYAPAEGDK et YCSVTTQADGR en PRM-HCD (a,b) et en PRM-LID à 473 nm (c,d)

Nous pouvons tout d'abord observer qu'en HCD ou en LID à 473 nm, se sont majoritairement les mêmes types d'ions (y) voir les mêmes fragments qui sont générés. Ensuite, en HCD, il n'y a aucun ion précurseur résiduel : la fragmentation est totale et donc la méthode d'activation permet théoriquement d'obtenir une meilleure sensibilité. Cependant, nous pouvons voir qu'en HCD les fragments b et y visualisés sont moins intenses qu'en LID à 473 nm. Ceci semble dû à une orientation particulière du processus de redistribution d'énergie lors de la fragmentation favorisant essentiellement la production de l'ion rapporteur. De plus, les intensités relatives entre les fragments produits par ces deux méthodes sont diverses. En effet, sur le spectre HCD du peptide CTEALYAPAEGDK (Figure 43 a) le fragment y_4^+ est plus intense que le fragment y_3^+ et le fragment y_7^+ or sur le spectre de photo-fragmentation le y_7^+ est de loin le plus intense des 3 suivi du fragment y_3^+ et du fragment y_4^+ (Figure 43 c).

Nous pouvons donc conclure à travers cette étude que la méthode d'activation LID à 473 nm permet l'obtention de fragments b et y plus intenses pour les peptides à cystéine dérivés avec le

DABCYL-C2 maléimide comparativement à la HCD, permettant plus de facilité et d'assurance lors de l'étape d'identification.

III) Construction d'une bibliothèque de spectres LID à 473 nm spécifiques de peptides à cystéine dérivés rapporteur de protéines kinases humaines

Dans cette section il sera présenté la manière par laquelle la bibliothèque de spectres LID à 473 nm de peptides à cystéine dérivés rapporteurs de protéines kinases humaines a été construite. Les parties expérimentales et computationnelles seront abordées. Un point sera également fait concernant la couverture de séquence moyenne offerte par celle-ci en vue de s'assurer que la bibliothèque construite permet d'identifier sans équivoque les peptides à cystéine dérivés qu'elle contient. Pour terminer, une comparaison de spectre de photo-fragmentation LID à 473 nm acquis en mode DIA-LID et DDA-LID à 473 nm sera également effectuée. La transférabilité des spectres de photo-fragmentation entre ces 2 modes sera démontrée.

III.1) Sélection des peptides à cystéine constitutifs de la bibliothèque de spectres

En amont de la réalisation de la bibliothèque, il a été répertorié via Peptide Atlas tous les peptides à cystéine correspondant à des protéines kinases humaines ayant été observés en spectrométrie de masse. Une restriction a été faite au niveau de la taille des peptides à cystéine trypsiques retenus : celle-ci ne devait être comprise qu'entre 9 et 25 acides aminés. De même, une seule cystéine par peptide à cystéine a été tolérée en vue d'éviter tout problème stœchiométrique lors de l'étape de dérivation. Cela a abouti au recensement de 446 peptides à cystéine rapporteurs de 446 protéines kinases humaines soit 1 peptide à cystéine rapporteur d'1 protéine kinase humaine.

Pour cette étude, nous avons souhaité nous baser sur des identifications au préalablement validées par la communauté scientifique ainsi qu'à nos critères de sélection, acceptant par là-même une non exhaustivité par rapport à la totalité des protéines kinases humaines connues.

III.2) Construction de la bibliothèque de spectres LID à 473 nm

Pour construire la bibliothèque de spectres LID à 473 nm, les 446 peptides à cystéine synthétiques commandés chez Thermo Scientific ont été regroupés dans une matrice aqueuse puis dérivés à l'aide du protocole type, séparés à l'aide du gradient 2 et analysés en DDA-LID à 473 nm en vue d'obtenir leurs spectres de photo-fragmentation.

Le gradient 2, utilisé ici, sera le gradient avec lequel seront analysés les fluides biologiques par la suite (les échantillons E.Coli dopés ou non et l'extrait cellulaire mammaire cancéreux). L'information qui sera extraite concernera à la fois les spectres de photo-fragmentation mais aussi

les RT de tous les peptides à cystéine dérivés. En procédant de la sorte, il sera alors possible de prédire les RT des peptides à cystéine dérivés ciblés lors de l'analyse des fluides biologiques grâce à la réutilisation du même gradient. Les peptides à cystéine dérivés ne seront ainsi recherchés que dans leur zone respective attendue pour chacun d'entre eux. Cela permet à la fois de minimiser les faux positifs et de gagner du temps lors du traitement de données.

Une fois les spectres de photo-fragmentation acquis expérimentalement, la bibliothèque de spectres a pu être construite computationnellement sur Mascot et Proline. Cette étape informatique a été réalisée à Strasbourg au sein de l'équipe du LSMBO. Les spectres de photo-fragmentation expérimentaux des peptides à cystéine dérivés synthétiques détectés ont été validés en rapport aux spectres théoriques leur correspondant présent sur Mascot en utilisant les critères de validation classiquement appliqués lors du traitement de données d'une analyse DDA :

Une déviation en masse maximale sur les ions précurseurs et les ions fragments a été fixée à 10 ppm et 0,05 Th respectivement. Aucune modification n'a été tolérée outre celle sur l'acide aminé cystéine correspondant au chromophore DABCYL-C2-maléimide. Une valeur de 0 « missed cleavage » a été admise. Aucune stringence ou contrôle sur d'éventuels faux positifs n'a été réalisé en raison de la nature de l'échantillon (un mélange de peptides à cystéine connus dans une matrice aqueuse) et de l'analyse DDA (les spectres de photo-fragmentation expérimentalement obtenus peuvent être reliés à la masse et à la charge des ions précurseurs desquels ils sont issus). Il n'y a quasiment aucun risque d'erreurs.

Cette étape a mené au recensement et à la validation de 354 spectres de photo-fragmentation expérimentaux. Pour les peptides à cystéine 92 manquants, les spectres de photo-fragmentation n'ont pu être validés. Cette absence peut s'expliquer de deux manières différentes. Tout d'abord, la qualité des peptides à cystéine achetés chez Thermo Scientific. Ceux-ci sont de statut « Peptide grade 2 » signifiant une qualité moindre. Une multitude de composés sont donc susceptibles d'être présents en plus des peptides à cystéine attendus au sein des échantillons : les peptides à cystéine attendus tronqués, des acides aminés résiduels, des éléments inconnus etc…De ce fait, les peptides à cystéine pouvaient être présent à de trop faible concentrations au sein des échantillons fournis (l'information fournie par Thermo Scientific sur les concentrations théoriques des peptides à cystéine achetés n'est pas la bonne) pour être détectés comme nous pouvons le constater pour le peptide à cystéine dérivé EGNIQYLFLEYCSGGELFDR (Figure 44) :

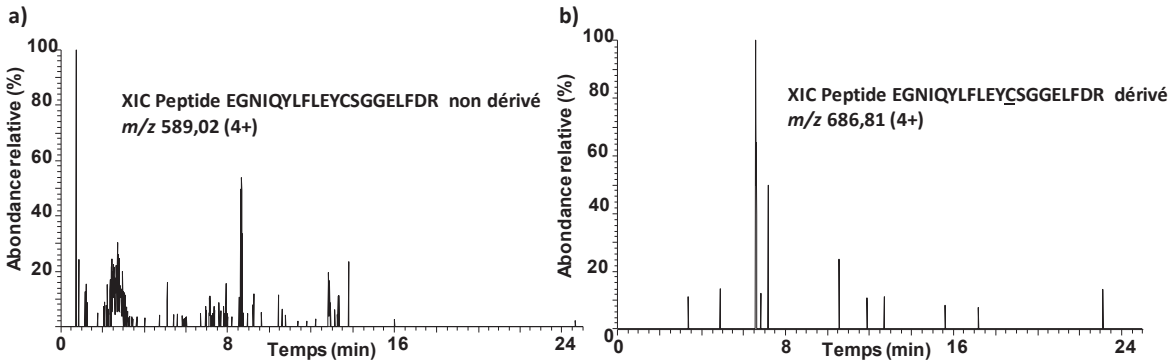


Figure 44 : XIC de l'ion quadrichargés du peptide à cystéine EGNIQYLFLEYCSGGELFDR sous forme non dérivée (a) et dérivée (b) lors de l'analyse DDA-LID à 473 nm

Dans cette figure est présenté à la fois les XIC du peptide à cystéine EGNIQYLFLEYCSGGELFDR sous forme non dérivée et dérivée (Figure 44 a et b) : il n'est présent sous aucune des deux formes. L'absence de potentiel dimère a aussi été contrôlée. Une autre raison supposée serait, cumulée à la raison précédente, une mauvaise réponse à l'ionisation ESI.

Pour remédier à cela, les quantités injectées ont été augmentées pour ces 92 peptides à cystéine jusqu'à une valeur théorique de 1 µg déposée sur la colonne chromatographique pour chaque peptide à cystéine dérivé mais cela n'a rien donné, ils n'ont pas été observés. Ces 92 peptides à cystéine ont donc été considérés comme non détectables par l'appareillage et la bibliothèque de spectres s'est vu constituée de 354 spectres de photo-fragmentation LID à 473 nm correspondant à 354 peptides à cystéine dérivés rapporteurs de 354 protéines kinases humaines (Annexe 2).

Une fois construite, la bibliothèque peut être utilisée. Le logiciel qui a été choisi pour utiliser cette bibliothèque de spectres et traiter toutes les analyses DIA-LID à 473 nm est le logiciel Skyline. L'interface de la bibliothèque sous ce logiciel se présente de la sorte (Figure 45) :

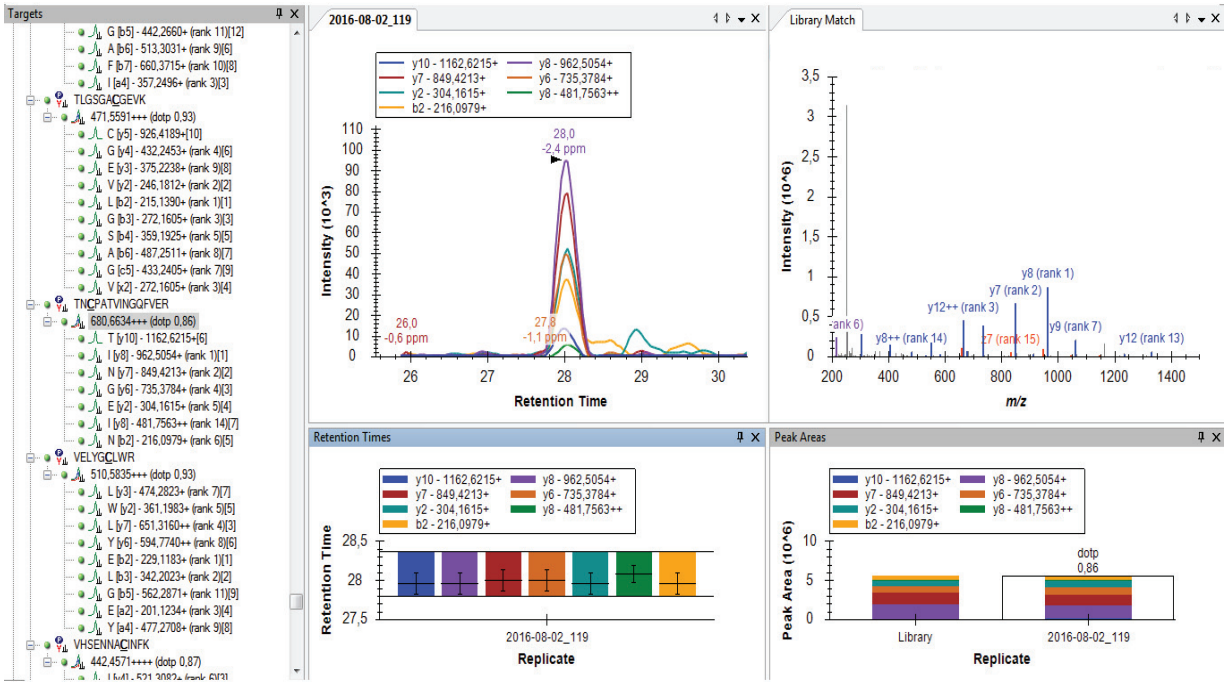


Figure 45 : Interface sous le logiciel Skyline de la bilbiothèque de spectres LID à 473 nm contenant les 354 peptides à cystéine dérivés rapporteurs de 354 protéines kinases humaines

Au niveau de la colonne de gauche se situe la liste des 354 peptides à cystéine dérivés présents au sein de la bibliothèque de spectres LID à 473 nm. On y retrouve le *m/z* de tous leurs ions fragments qui ont été détectés dans les spectres expérimentalement acquis lors de l'analyse DDA-LID à 473 nm. Ce sont ces fragments qui seront utilisés pour rechercher ces peptides à cystéine dans les matrices biologiques.

Pour traiter les données d'une expérimentation, l'interface a été utilisée de la sorte :

- Le 1er carré en haut à gauche, présentait les XIC de tous les ions fragments putatifs détectés par peptide à cystéine dérivé avec leurs déviations en masse respectives et leur RT expérimentaux.

- Le 2nd carré en haut à droite présentait le spectre expérimental contenu dans la bilbiothèque de spectres du peptide à cystéine dérivé traité avec le RT prédit.

- Le 3ème carré en bas à gauche présentait l'alignement en RT de l'apex de tous les fragments putatifs détectés.

- Le 4ème carré en bas à droite affichait la comparaison en intensité relative entre les fragments putatifs détectés issus de l'échantillon analysé et les fragments non putatifs contenu dans la bibliothèque de spectres. Cette correspondance sera matérialisée quantitativement par le biais d'un dopt score compris entre 0 et 1 avec pour zéro une correspondance nulle et pour 1 une correspondance parfaite.

Ce fut à l'aide de ce logiciel et de cette interface que furent traitées toutes les données issues des échantillons analysés par la suite.

III.3) Contrôle de l'information structurale contenue dans la bibliothèque de spectres LID à 473 nm

Une fois la bibliothèque de spectres LID à 473 nm construite, nous avons effectué une étude concernant l'information structurale détenue dans ses spectres dans le but de nous assurer qu'elle permettait d'identifier sans équivoque les peptides à cystéine dérivés qu'elle contenait.

Dans cette optique, tout d'abord, nous avons extrait le pourcentage moyen d'ions fragments b et y non redondant contenu dans chaque spectre de photo-fragmentation pour tous les peptides à cystéine dérivés de la bibliothèque afin d'avoir un aperçu de la couverture de séquence moyenne offerte par celle-ci (Equation 15 et Figure 46). Nous nous sommes limités aux ions b et y car il représente en moyenne plus de 90 % des ions fragments générés dans chaque spectre de photo-fragmentation comme nous le verrons dans le chapitre III :

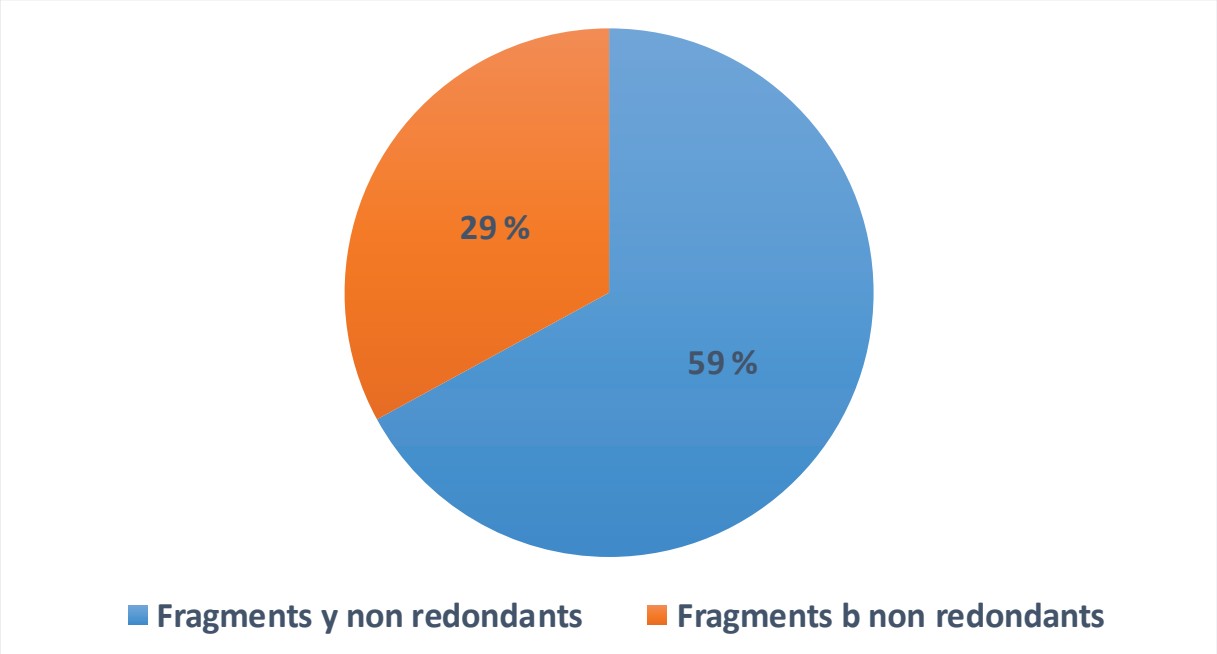


Figure 46 : Couverture de séquence moyenne obtenue via les ions b et y pour les 354 peptides à cystéine dérivés de la bibliothèque de spectres LID à 473 nm

Ici, un pourcentage de 100 % signifie une couverture totale. Par exemple, si un peptide à cystéine possède 14 acides aminés, alors 100 % équivaut aux 14 fragments y non redondants. Il en va de même pour les ions b. Nous pouvons observer ici, un pourcentage moyen d'ion y de 59 % et d'ion b de 29 %. Cela se traduit par une moyenne de 12 fragments (b et y mélangés) par peptide à cystéine pour une taille moyenne de 14 acides aminés. Nous pouvons donc estimer que si la

couverture de séquence n'est pas totale, elle est conséquente et suffisante pour identifier sans ambiguïté des peptides à cystéine dérivés dans des matrices biologiques.

Ensuite, nous avons pu voir précédemment (Chapitre II.II.3) que l'ion rapporteur (*m/z* 252,11) est l'ion fragment majoritairement produit lors de la photo-fragmentation des peptides à cystéine dérivés à 473 nm. Beaucoup d'autres fragments le sont, mais à moindre intensité. Or en DIA, en raison de la (co)-fragmentation, l'ion rapporteur sera généré par tous les peptides à cystéine et ne pourra pas jouer le rôle de fragment discriminant. Nous devons donc nous assurer, qu'en LID à 473 nm, tous les autres fragments produits, le sont suffisamment en termes d'intensité et peuvent servir à identifier les peptides à cystéine dérivés.

Dans cette optique, nous avons étudié la proportion de l'intensité des ions fragments b, y et de l'ion rapporteur *m/z* 252,11 les uns par rapport aux autres pour 40 peptides à cystéine dérivés de la bibliothèque (Equation 16 et Figure 47) :

$$\text{\% fragment (b ou y ou 252,11)} = \frac{\text{E I b ; y ; 252,11}}{\text{E I tous fragments}} \text{ X 100} \qquad \text{Equation 16}$$

Avec $I_{b\ ;\ y\ ;\ 252,11}$: l'intensité de tous les ions fragments b, y ou du 252,11 et $I_{tous\ fragments}$: l'intensité de tous les fragments détectés.

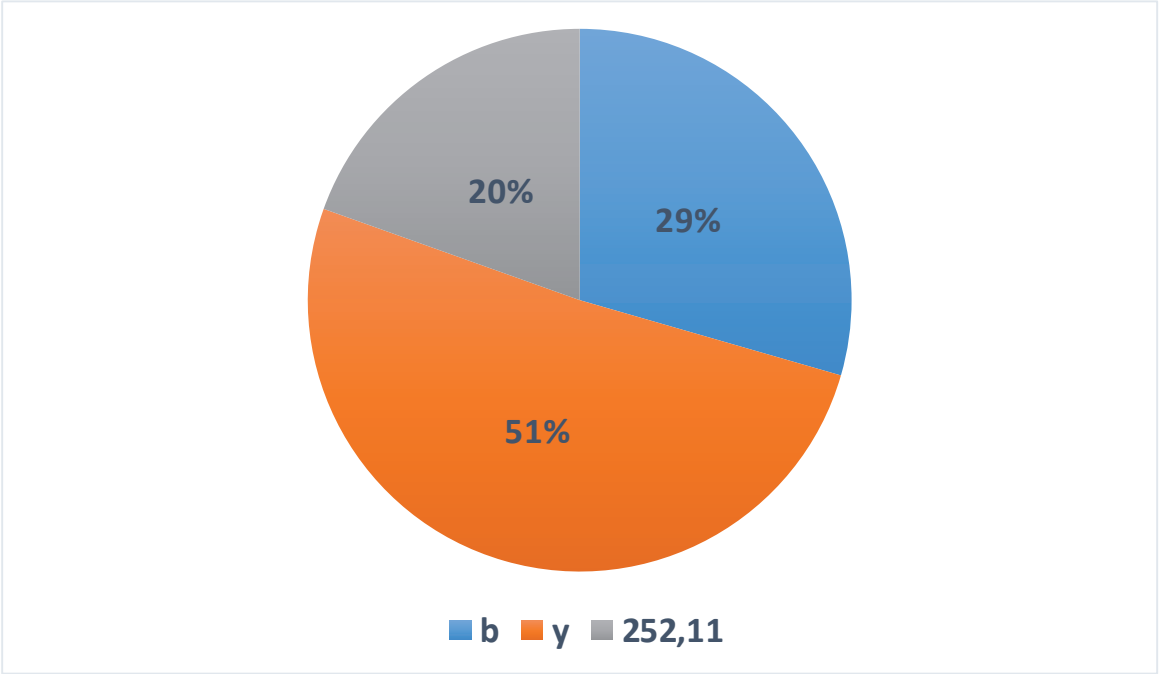


Figure 47 : Proportion moyenne, sur un spectre de photo-fragmentation, de l'intensité de l'ion rapporteur *m/z* 252,11 et des ions fragments b et y pour 40 peptides à cystéine dérivés de la bibliothèque de spectres LID à 473 nm

Comme nous pouvons l'observer (Figure 47), lors de la photo-fragmentation de 40 peptides à cystéine dérivés à 473 nm, l'ion rapporteur représente en moyenne 20 % de l'intensité totale des fragments détectés. 80 % de l'intensité totale des fragments générés produite lors de la photo-fragmentation peut donc servir à identifier les peptides à cystéine dérivés au sein de spectres de (co)-fragmentation en DIA-LID à 473 nm.

Nous pouvons donc conclure que les fragments b et y produits lors de la photo-fragmentation de peptides à cystéine dérivés à 473 nm, le sont suffisamment en termes d'intensité pour servir à les identifier.

III.4) Comparaison entre la DDA-LID et la DIA-LID à 473 nm

Une étude comparative a également été effectuée quant aux spectres obtenus en DDA-LID et en DIA-LID à 473 nm. Les spectres de la bibliothèque ayant été acquis selon ce mode, il fallait s'assurer qu'aucune divergence n'était obtenue à ce point. En effet, les spectres compilés dans la bibliothèque vont servir pour l'identification, la similarité doit donc être de mise.

Voici les spectres du peptide à cystéine dérivé VMCDSNDLPK issu de l'analyse DDA-LID à 473 nm des 446 peptides à cystéine synthétiques commandés chez Thermo Scientific en matrice aqueuse. L'expérimentation a également été effectuée en DIA-LID à 473 nm (dérivation à l'aide du protocole type, séparation à l'aide du gradient 2) (Figure 48).

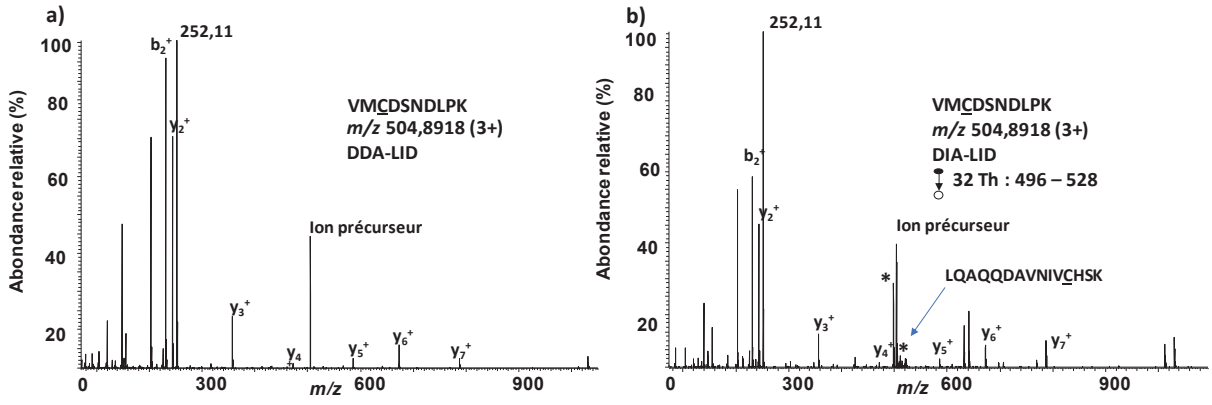


Figure 48 : Comparaison des spectres de photo-fragmentation du peptide à cystéine dérivé VMCDSNDLPK en DDA-LID à 473 nm (a) et en DIA-LID à 473 nm (b)

A travers cette figure (Figure 48) nous pouvons observer, comme souhaité, que les intensités relatives sont bien conservées entre les 2 méthodologies et qu'aucun fragment n'a été interféré. Ensuite, comme attendu, le spectre de (co)-fragmentation obtenu en DIA contient des ions précurseurs supplémentaires relativement à celui ciblé. Cela provient de la plus large plage de masse sondée (32 Th entre 496 et 528 Th). Nous pouvons également observer que l'ion rapporteur *m/z* 252,11 a vu son intensité relative augmenter entre les 2 spectres. Cela est dû à la (co)-fragmentation d'ions précurseurs dérivés avec le DABCYL C2-maléimide présent au sein de

la même plage de 32 Th et éluant au même temps de rétention (le peptide à cystéine LQAQQDAVNIVCHSK sous sa forme quadrichargée *m/z* 512,00) et rappel l'impossibilité d'utiliser l'ion rapporteur en DIA-LID à 473 nm pour identifier ou quantifier des peptides à cystéine dérivés.

Nous pouvons conclure que les spectres DDA-LID et DIA-LID à 473 nm sont similaires en termes d'intensités relatives. Les spectres de photo-fragmentation obtenus en DDA-LID à 473 nm pourront être utilisés pour l'identification de peptides à cystéine dérivés en DIA-LID à 473 nm.

IV) Preuve de concept de l'utilisation de la DIA-LID à 473 nm et d'une bibliothèque de spectres pour identifier et quantifier des peptides à cystéine dérivés rapporteur de protéines humaines au sein de matrices biologiques

Une fois la bibliothèque de spectres LID à 473 nm construite, nous sommes passés à la démonstration de la validité de cette bibliothèque utilisée avec la méthodologie la DIA-LID à 473 nm, pour identifier et quantifier des peptides à cystéine dérivés dans des matrices biologiques. Cette section sera divisée en deux temps : premièrement nous aborderons la validité du concept pour identifier des peptides à cystéine dérivés de la bibliothèque dans des fluides biologiques, la seconde sera axée sur la partie quantification de ces mêmes peptides. La matrice biologique test qui a été choisie pour ces 2 études est une matrice biologique bactérienne E.Coli. Celle-ci est une matrice biologique non humaine où les protéines kinases humaines ne sont pas exprimées.

IV.1) Preuve de concept de l'utilisation de la DIA-LID à 473 nm et d'une bibliothèque de spectres pour identifier des peptides à cystéine dérivés au sein de matrices biologiques

Pour démontrer l'efficience de la méthodologie DIA-LID à 473 nm utilisée avec une bibliothèque de spectres pour identifier les 354 peptides à cystéine dérivés la constituant dans une matrice E.Coli nous avons procédé de la manière suivante :

100 µg de E.Coli ont été dopé avec les 354 peptides à cystéine à des quantités théoriques allant de 17 ng à 14,15 µg (soit une quantité injectée allant de 4 ng à 3,33 µg). Ces quantités théoriques ont été établies en fonction des précédentes analyses effectuées pour construire la bibliothèque de spectre en DDA-LID à 473 nm. Lors de ces acquisitions une même quantité théorique de 200 ng avait été déposé pour chaque peptide à cystéine dérivé sur la colonne chromatographique mais des intensités différentes avaient été obtenues pour chaque peptide. Compte tenu du signal obtenu avec 200 ng injectés, les quantités de chaque peptide à cystéine ont été réajustées proportionnellement en vue d'obtenir un signal de 5.10^7 cps. Cette intensité a été estimée suffisante pour visualiser plus de 6 ions fragments par spectre de photo-fragmentation tout en ne saturant pas le signal.

Ainsi, le peptide à cystéine dérivé qui avait été détecté le plus intensément a été déposé à une quantité de 4 ng (au lieu de 200 ng dans l'analyse précédente). Inversement, le peptide à cystéine dérivé qui répondait le moins bien lors de l'analyse précédente a été déposé à une quantité de 3,33 µg (au lieu de 200 ng dans l'analyse précédente). Nous savons que la réponse d'un peptide en spectrométrie de masse ne varie pas proportionnellement à la concentration mais cela a été décidé pour tenter de normaliser le signal d'une certaine manière vu que nous n'avions aucune information sur les concentrations réelles initiales. Les valeurs des concentrations théoriques sont donc à prendre avec les considérations appropriées.

Cette manipulation s'est traduite par la répartition des 354 peptides à cystéine dans 3 échantillons de 100 µg d'E.Coli (environ 120 peptides à cystéine par échantillons). Les échantillons ont ensuite été dérivés suivant le protocole type, séparés à l'aide du gradient 2 et analysés en mode DIA-LID à 473 nm. Les données ont ensuite été traitées sur le logiciel Skyline. Dans la communauté scientifique, il n'y pas encore de réel consensus sur les critères à utiliser pour valider une identification par le mode DIA. Nous avons donc développé les nôtres en relation à cette étude et aux informations que nous avions pu extraire lors de recherches bibliographiques.

IV.1.a) Identification des 354 peptides à cystéine dérivés au sein d'une matrice d'E.Coli

Notre première approche a consisté à rechercher tous les peptides à cystéine de la bibliothèque à l'aide des critères que nous avions retenus et qui sont présentés ci-dessous :

- Chaque peptide à cystéine dérivé issu de la bibliothèque est recherché à l'aide de ses 6 ions fragments les plus intenses (avec 3 ions fragments détectés non interférés au minimum pour valider une identification).

- Chaque peptide à cystéine dérivé issu de la bibliothèque est recherché au sein d'une fenêtre de +/- 2,5 min par rapport au RT prédit.

- Pour chaque peptide à cystéine dérivé putatif détecté la comparaison en intensité relative entre les fragments putatifs détectés expérimentaux et les fragments non putatifs contenus dans la bibliothèque de spectres doit mener au minimum à un dopt score de 0,8 sur Skyline.

- Pour tous les fragments pour tous les peptides à cystéine dérivés putatifs détectés, une déviation en masse maximum de 15 ppm est tolérée.

Via l'utilisation de ces critères et d'un traitement manuel sur le logiciel Skyline cette manipulation a mené à la validation de l'identification de 272 peptides à cystéine dérivés sur les 354 contenus dans la bibliothèque de spectres LID à 473 nm soit 77 % de ce qui était attendu.

Voici 6 XIC issus des 272 peptides à cystéine dérivés identifiés (Figure 49) :

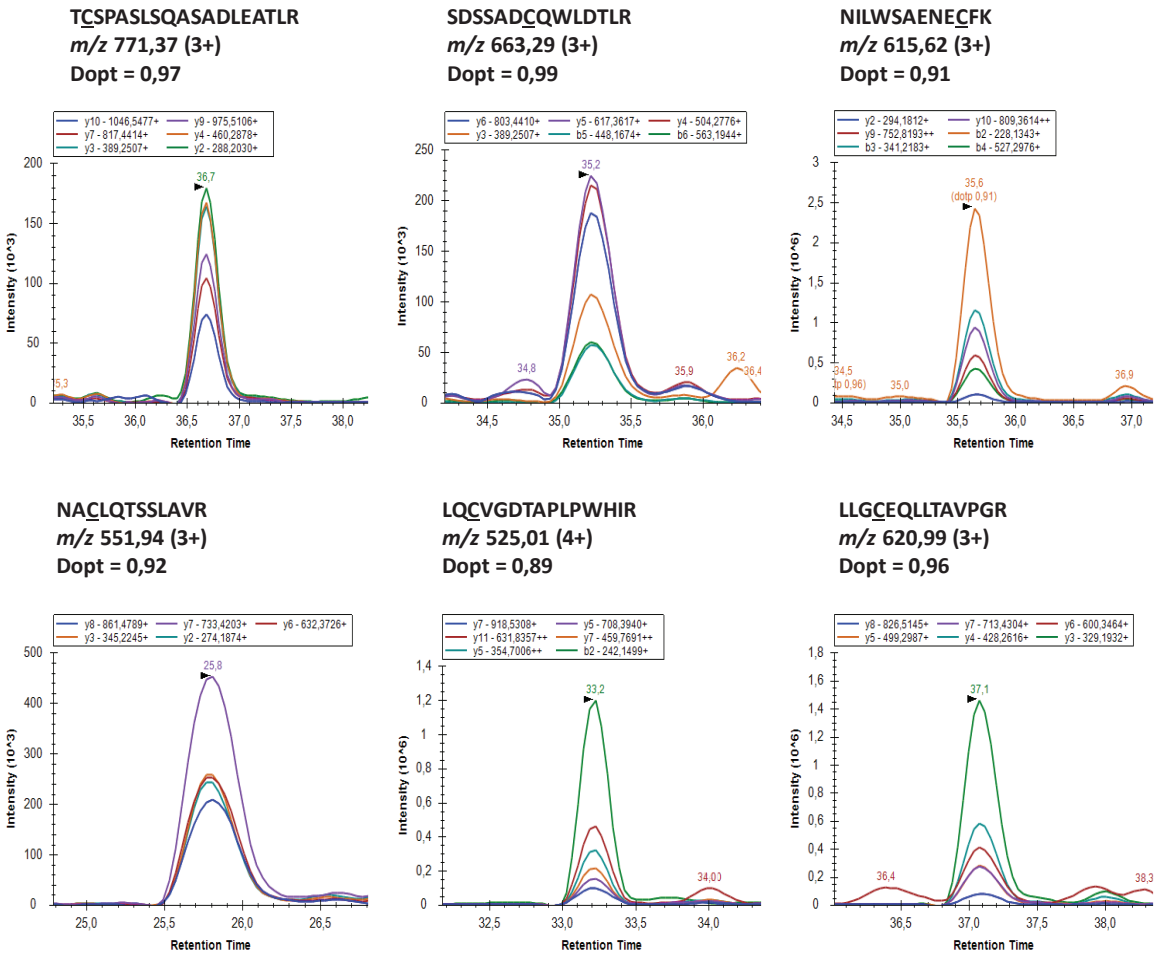


Figure 49 : XIC de 6 peptides à cystéine dérivés détectés en mode DIA-LID à 473 nm dans les échantillons d'E.Coli dopés avec les 354 peptides à cystéine dérivés

Les critères de validation, suite à cette première manipulation, sont explicités ci-dessous :

Il a été estimé par le passé qu'il était possible de traiter les données DIA similairement à des approches ciblées type SRM et que l'utilisation de 3 à 5 ions fragments par peptides pour les identifier pouvait être suffisante [117]. L'idée a été gardée et testée (mais de 3 à 6). Une sélection des fragments à suivre devait alors s'imposer. 3 choix s'offraient : soit suivre les fragments les plus spécifiques, les plus intenses ou un mélange des deux. Nous nous sommes orientés vers la sélection des 6 fragments les plus intenses par peptides à cystéine. Ce choix résulte de ce que nous avons précisé au préalable par rapport à la méthode d'activation LID à 473 nm : la photo-fragmentation de l'ion précurseur n'est pas totale, ainsi, le signal des ions fragments est parfois peu intense et les fragments les plus longs (spécifiques) sont parfois difficiles à détecter. Puis, nous avons pu observer, la formation d'isomère de conformation lors de nos analyses comme nous pouvons le voir pour les peptides à cystéine dérivés LTSVVGTILYSCPEVLK, LDEAQEAECQVLK, DQADEAQYCQLVR (Figure 50) :

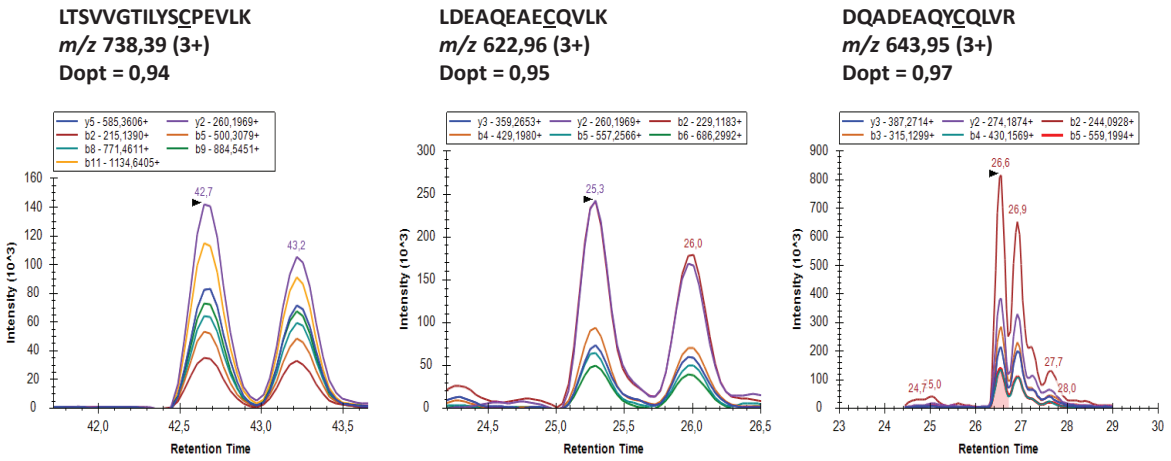


Figure 50 : XIC de 3 peptides à cystéine dérivés détectés en mode DIA-LID à 473 nm dans l'échantillon d'E.Coli dopé avec les 354 peptides à cystéine dérivés présentant une isomérie de conformation

Nous n'avons pas étudié les raisons de cette isomérie, il n'empêche que celle-ci aussi impact négativement le signal pour les peptides à cystéine concernés : il est quasiment divisé par 2.

Pour toutes ces raisons, nous nous sommes orientés sur ce choix des fragments les plus intenses tout en étant conscient de la possibilité d'augmenter le risque de faux positifs en cas de non spécificité de ces fragments. Cela a été contrôlé (Tableau 10) :

Taille du fragment (nombre d'acides aminés)	Proportion (%)	Proportion cumulée (%)
1	0,4	100,0
2	16,3	99,6
3	13,2	83,3
4	13,1	70,1
5	12,8	57,0
6	11,4	44,2
7	10,2	32,9
8	8,0	22,7
9	5,5	14,7
10	3,3	9,3
11	2,4	6,0
12	1,7	3,6
13	0,7	1,9
14	0,7	1,2
15	0,3	0,5
16	0,1	0,2
17	0,1	0,1

Tableau 10 : Proportion (cumulée ou non) de chaque taille de fragment (nombre d'acides aminés) au sein des 6 fragments les plus intenses issus des 354 spectres de photo-fragmentation de la bibliothèque de spectres LID à 473 nm.

La taille (nombre d'acides aminés) des 6 fragments les plus intenses issus de tous les spectres de photo-fragmentation de la bibliothèque a été relevée. Il a été considéré qu'à partir d'une taille de 5 acides aminés au minimum (y_5^+ par exemple), les fragments pouvaient être considérés comme spécifiques. Nous avons ensuite regardé la proportion de fragments de taille égale ou supérieure à 5 par rapport à la totalité des 6 fragments les plus intenses relevés. Cela représente 57 % de la totalité des 6 fragments les plus intenses pour tous les peptides à cystéine de la bibliothèque. La majorité des fragments utilisés sont donc plutôt spécifiques.

Lors de l'identification des 272 peptides à cystéine dérivés, le nombre de fragments utilisés pour les identifier a été vérifié (Figure 51) :

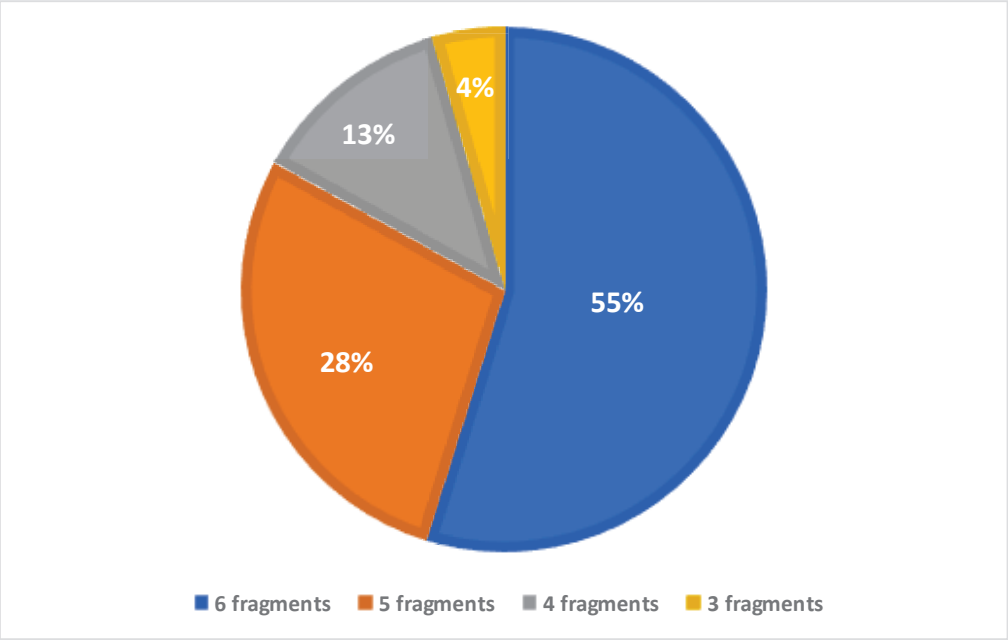


Figure 51 : % de peptides à cystéine dérivés identifiés avec 6, 5, 4 ou 3 fragments

Sur les 272 peptides à cystéine dérivés 55 % des peptides ont été identifiés à l'aide de leur 6 fragments les plus intenses et 96 % avec plus de 4 fragments. Sur les 1632 ions fragments relatifs à ces 272 peptides à cystéine dérivés (272 peptides à cystéine dérivés X 6 fragments les plus intenses = 1632) 1446 ont été observés non interférés (soit 88,6 %) démontrant par là-même le peu d'interférence générée en DIA-LID à 473 nm en raison de la spécificité de la méthode d'activation. Au vu du peu d'interférences observées, l'utilisation des 3 à 6 fragments les plus intenses semble suffisante.

Au niveau du temps de rétention, le choix s'était orienté vers des fenêtres de +/- 2,5 min. Ce fenêtrage a été validé, par rapport aux temps de rétention prédit, pour couvrir d'éventuels effets de matrice [219]. Nous avons pu observer que ce fenêtrage était adapté (Tableau 11).

	Taille de la fenêtre en RT utilisée (+/- X min)		
	1	**2.5**	**> 2,5**
Nombre de peptides	248	25	14
%	86,4	8,7	4,9
% cumulé	86,4	95,1	100,0

Tableau 11 : Nombre, pourcentage et pourcentage cumulé de peptides à cystéine dérivés identifiés en mode DIA-LID à 473 nm en fonction de la taille de la fenêtre en RT utilisée

Sans tenir compte du fenêtrage en RT et en utilisant tous les autres paramètres fixés, 287 peptides à cystéine dérivés ont été identifiés. Plus de 95 % (272) se retrouvaient dans un fenêtrage inférieur à +/- 2,5 min. Les 14 restants étaient décalés en moyenne de 3,90 min avec un maximum de 5,82 min et un minimum de 2,60 min. Ces observations ont été attribuées à des effets de matrice. Malgré la perte d'information possible, il a été décidé de garder ce fenêtrage de +/- 2,5 min en vue de minimiser le risque de faux positif tout en ayant eu la démonstration qu'on retrouvait la quasi-totalité des peptides à cystéine dérivés identifiés dans ces limites.

Le score dopt fourni par le logiciel Skyline a également été pris en compte. Il renseigne sur les intensités relatives entre tous les fragments par peptide. Il est compris entre 0 et 1 pour chaque peptide avec la valeur 1 signifiant une correspondance parfaite entre les intensités relatives de la bibliothèque et celles obtenues expérimentalement. Il n'a pas été explicitement trouvé de valeurs limites consensus à ce niveau au sein de publications scientifiques. Sur un tutoriel de Skyline une valeur de dopt de 0,80 jouait le rôle de valeur limite. Sans tenir compte du dopt et en utilisant tous les autres paramètres fixés, 275 peptides à cystéine dérivés ont été identifiés (Tableau 12).

	Dopt				
	< 0,80	**0,80-0,85**	**0,85-0,90**	**0,90-0,95**	**0,95-1**
Nombre de peptides	3	18	49	127	78
%	1,1	6,6	18,0	46,7	28,7
% cumulée	100,0	98,9	92,4	74,5	28,4

Tableau 12 : Nombre, pourcentage et pourcentage cumulé de peptides à cystéine dérivés identifiés en fonction du dopt score

La moyenne du dopt sur les 275 peptides à cystéine dérivés identifiés est de 0,92. Plus de 74,5 % (205) possèdent un dopt supérieur 0,90 et 98,9 % (272) des peptides à cystéine dérivés identifiés supérieur à 0,80. 3 peptides à cystéine dérivés identifiés possédaient un dopt inférieur à 0,80. La valeur posée comme seuil est celle de 0,80 qui permet de recouvrir plus de 98,9 % (272) des peptides à cystéine dérivés identifiés. Si une valeur de 0,85 semblait plus stringente que celle à 0,80, garder celle-ci nous permet d'avoir accès à des peptides à cystéine dérivés qui sont faiblement interférés en raison de composés présents au sein de la même fenêtre de sélection et qui génèrent 1 fragment isobare venant perturber les intensités relatives mesurées. On peut l'observer pour les peptides à cystéines dérivés MGGSFLI<u>C</u>SK, IAF<u>C</u>QGFDK (transition y_8^{++}, y_2^{+}

respectivement) qui avaient, en outre, une déviation en temps de rétention inférieure à 2,5 min (Figure 52) :

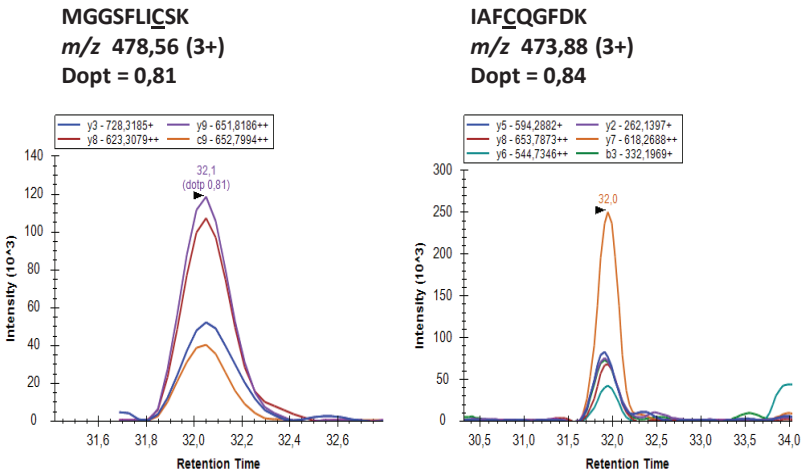


Figure 52 : XIC de 2 peptides à cystéine dérivés détectés en mode DIA-LID à 473 nm avec un dopt score inférieur à 0,85

Au niveau de la déviation en masse, celle-ci n'a pas été remaniée en raison de sa relation à l'appareillage ainsi qu'aux ions mesurés. Cependant, pour donner un ordre d'idée, sur la totalité des ions fragments gardés pour identifier les peptides à cystéine dérivés, la déviation moyenne en masse moyenne était de 2,5 ppm avec un maximum à 15 ppm et un minimum à 0 ppm.

Voici donc les critères de validation finaux :

Suivi des 6 fragments les plus intenses dont au minimum 3 non interférés pour validation
Score dopt supérieur à 0,80
Une variation en RT comprise dans un intervalle de +/- 2,5 min par rapport au RT prédit
Déviation en masse inférieure à 15 ppm
Alignement des apex et homogénéité des formes des XIC des fragments par peptide

Via cette première approche nous avons pu identifier 272 peptides à cystéine dérivés sur les 354 attendus soit 77 %. 82 peptides à cystéine dérivés sont donc manquants. Les raisons pouvant expliquer ce manque sont diverses. Tout d'abord 17 peptides à cystéine ont été exclus lors du retraitement lié au RT et à leur dopt score. Les 65 restants n'ont pas été observés. Plusieurs raisons sont possibles : tout d'abord un effet matrice supérieur à 6 min (c'est le maximum en terme de déviation en RT que j'ai sondé), quelques fragments isobares ayant complétement empêchés la détection de ces peptides ou encore une quantité trop faible déposée au sein des 100 µg de E.Coli. Cependant, ignorant fondamentalement les concentrations réelles des peptides

à cystéine fournit par Thermo Scientific, il était impossible de contrôler cette dernière supposition ou de la rendre porteuse d'informations pertinentes. Des explications supplémentaires relativement au mode DIA seront apportées dans la section V de ce chapitre. Elles aborderont les limites de ces concepts pouvant aussi expliquer la non détection de ces 65 peptides à cystéine.

IV.1.b) Echantillon blanc

En parallèle de cette manipulation, un blanc a été effectué afin de contrôler d'éventuels faux positifs. Celui-ci se constituait simplement de 100 µg d'E.Coli sans dopage de quelconques peptides à cystéine synthétiques. Il a ensuite été dérivé suivant le protocole type, séparé à l'aide du gradient 2 et analysé en mode DIA-LID à 473 nm. Cette analyse a mené à l'identification de 0 composé. Aucun faux positif n'a été détecté. Cette information se révèle réellement importante car elle démontre qu'avec la spécificité liée à la méthode d'activation LID à 473 nm couplée aux fenêtres de temps de rétention, les fausses identifications semblent quasiment inexistantes. Ce témoin reconfirme les identifications faites au-dessus et démontre la potentialité de la méthode d'activation LID à 473 nm à remédier aux soucis de (co)-fragmentation inhérent à la DIA par activation collisionnelle.

IV.1.c) Comparaison entre la DIA-LID à 473 nm et la DIA-HCD

Dans le but d'appuyer les résultats issus de l'échantillon blanc, le même échantillon a été analysé en DIA-HCD avec exactement le même paramétrage qu'en DIA-LID à 473 nm. Les résultats obtenus peuvent être représentés par les peptides à cystéine dérivés NILWSAENE<u>C</u>FK et NILVGENTL<u>C</u>K (Figure 53) :

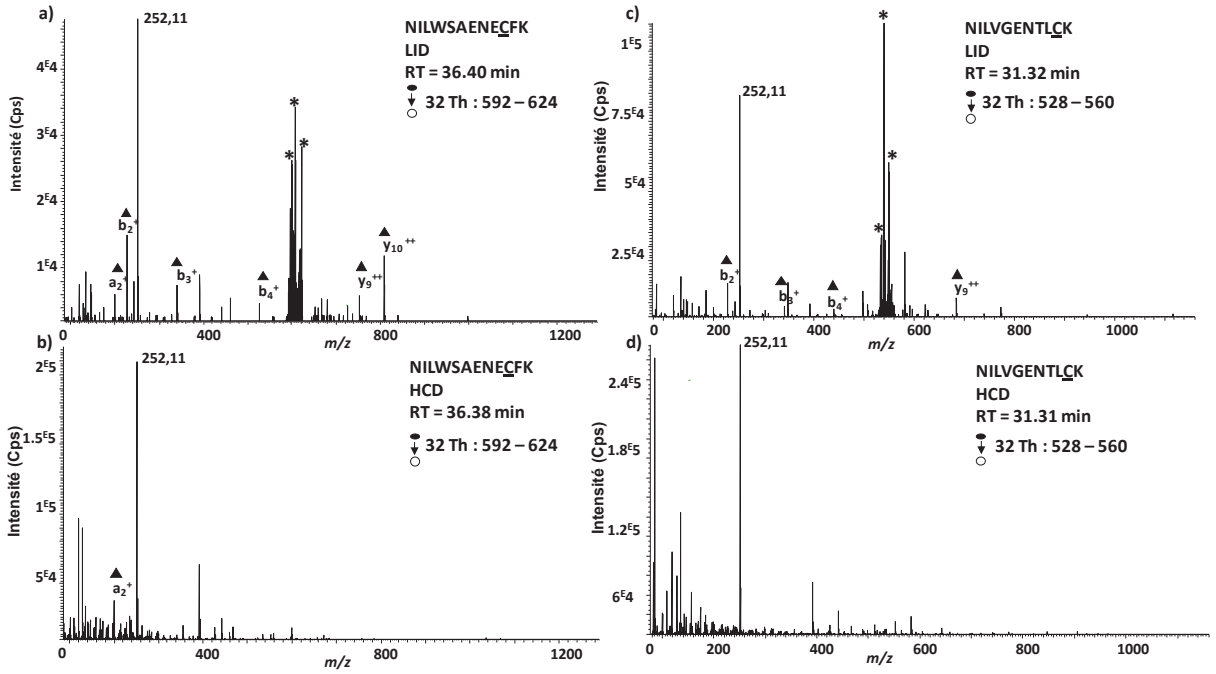


Figure 53 : Spectre de co-fragmentation des peptides à cystéine dérivés NILWSAENECFK et NILVGENTLCK en mode DIA-LID à 473 nm (a et c) et en DIA-HCD (b et d). Les * correspondent à des ions précurseurs ne contenant pas de cystéine dérivée.

A l'aide de ces 4 spectres de (co)-fragmentation nous pouvons visualiser, pour les 2 spectres DIA-LID à 473 nm, qu'au minimum 4 fragments correspondant aux peptides recherchés sont observés (Figure 53 a et c) contre 1 ou 0 en DIA-HCD (Figure 53 b et d). Cela peut être expliqué par la spécificité de la méthode d'activation LID à 473 nm. En effet, nous pouvons voir sur les spectres DIA-LID à 473 nm des ions précurseurs intacts (*), également sélectionnés dans la plage de masse correspondante. Cela vient du fait qu'ils ne possèdent pas de cystéine ou de fonction thiol et qu'ils n'ont donc pas été dérivés et ne peuvent donc pas photo-fragmenter. On observe ici la réduction de la (co)-fragmentation inhérente au mode DIA. Elle permet ainsi de simplifier les spectres obtenus et d'extraire plus facilement les informations recherchées. En revanche, cette spécificité est non présente via la méthode d'activation HCD. Il en résulte des spectres de fragmentation plus complexes ou plus d'ions fragments sont générés ce qui rend ici impossible l'identification des peptides à cystéine dérivés. L'information recherchée est noyée. Il est aussi important de rappeler qu'en HCD l'ion rapporteur est majoritairement généré et que les ions fragments b et y visualisés sont moins intenses comparativement à la méthode d'activation LID à 473 nm.

Ci-dessous il est présenté 2 XIC des peptides à cystéine dérivés LYLAENYCFDSIPK et NILWSAENECFK en DIA-HCD et LID à 473 nm pour traduire un autre effet (Figure 54) :

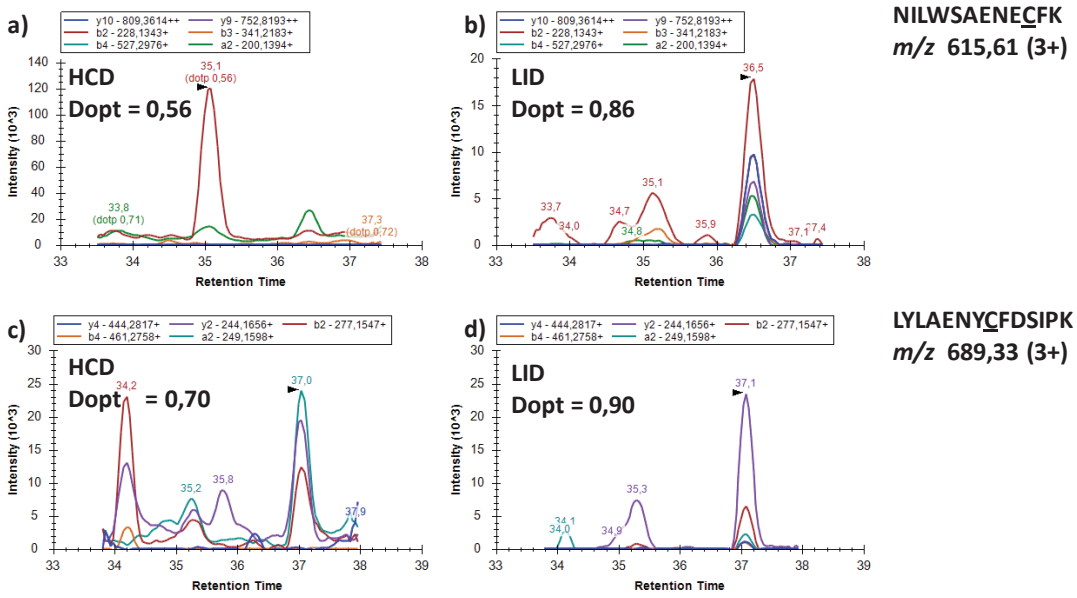


Figure 54 : XIC des 2 peptides à cystéine dérivés LYLAENYCFDSIPK et NILWSAENECFK en mode DIA-HCD (a et c) et DIA-LID à 473 nm (b et d)

Nous pouvons voir à travers ces 4 XIC que la DIA avec la méthode d'activation HCD, outre de diluer l'information, génère aussi de l'information interférente sur celle recherchée. Pour les 2 peptides à cystéine dérivés présentés ici, la majorité des fragments ont été interférés en HCD (Figure 54 a et c) rendant impossible leur identification. En revanche, cela n'apparait pas en LID à 473 nm (Figure 54 b et d) ce qui permet l'identification de ces 2 peptides à cystéine dérivés.

IV.1.d) Conclusion

Après ces observations, nous pouvons conclure plusieurs choses. Tout d'abord, repréciser la viabilité des informations extraites des spectres de photo-fragmentation obtenus en mode DDA-LID à 473 nm pour identifier les peptides à cystéine dérivés au sein de matrices biologiques en mode DIA-LID à 473 nm. Ensuite, que la spécificité liée à la méthode d'activation LID à 473 nm entraine une simplification des spectres de (co)-fragmentation en DIA. Cela permet plus de facilité pour extraire l'information recherchée et amène à plus de confiance lors de l'étape d'identification.

Avec les critères de validation utilisés, la méthodologie DIA-LID à 473 nm a permis de retrouver plus de 77 % des peptides à cystéine introduits dans une matrice biologique complexe E.Coli sans, a priori, de faux positifs.

Le potentiel résidant dans l'utilisation de la méthodologie DIA-LID à 473 nm et d'une bibliothèque de spectres pour identifier des peptides à cystéine dérivés dans des matrices biologiques a ici été démontré. Une fois cette preuve de concept établie, nous avons étudié son potentiel pour quantifier de manière répétable ces mêmes peptides à cystéine dérivés.

IV.2) Preuve de concept de l'utilisation de la DIA-LID à 473 nm et d'une bibliothèque de spectres pour quantifier de manière répétable des peptides à cystéine dérivés au sein de matrices biologiques

La méthodologie DIA-LID à 473 nm et la bibliothèque de spectres ont été utilisés pour quantifier des peptides à cystéine dérivés dans la matrice E.Coli. 65 peptides à cystéine dérivés parmi les 272 identifiés au préalable ont été sélectionnés et quantifiés de manière répétable sur une gamme dynamique de 3 décades. Ne connaissant pas les concentrations réelles de ces peptides ils ont été ajoutés à l'aveugle dans 100 µg d'E.Coli à 4 niveaux de concentration normalisés différents : 1, 0,1, 0,01 et 0,001 avec le point 0,001 point le moins concentré. Les 4 échantillons ont ensuite été dérivés suivant le protocole type, séparés à l'aide du gradient 2 et analysés en mode DIA-LID à 473 nm en triplicats. Les données ont été traitées sur le logiciel Skyline avec les critères établis au préalable. Voici les résultats obtenus pour les 65 peptides à cystéine dérivés (Tableau 13) :

Peptide	Transition	0,001 CV (%)	0,01 CV (%)	0,1 CV (%)	1 CV (%)	R^2
AAC<u>C</u>LLDGVPVALK	y5+	4,6	2,6	2,4	14,9	0,988
APEILLGLPF<u>C</u>EK	b8+	4,1	2,3	1,8	17	0,999
ASC<u>C</u>STAGSGSR	y7+	11,5	3,5	3,8	22,7	0,999
<u>C</u>AAQLASALEYIHAR	y10++	5,2	3,4	24,6	13,7	0,996
<u>C</u>EETATGLK	y5+	11,2	7,3	2,3	18,8	0,998
<u>C</u>ENALLQGFNLK	y3+	6,2	1,7	2,9	17,9	0,997
<u>C</u>ENVLLSPDER	y7+	6,4	0,8	3,8	9,4	0,958
<u>C</u>FEISDADTK	y7+	12	4,6	1,8	16,8	0,996
<u>C</u>HAPENVELALR	y4++	3,4	1,8	2	15,2	0,999
<u>C</u>LIELEPVLR	y7++	3	1,2	2,6	13,7	0,98
<u>C</u>MQEIPQEQIK	z3+	2,9	3,6	2,7	7,4	0,963
<u>C</u>PAAGNPTPSISWLK	y9++	7,9	0,8	12,3	7,5	0,953
<u>C</u>PAAGNPTPTIR	y6++	3	3,2	4,6	15,8	0,994
<u>C</u>PETPAQPDSR	y2+	4,3	8,9	9,9	6,7	0,999
<u>C</u>SESSTLLEK	y7++	9,1	9,6	1,6	15	0,995
<u>C</u>SGLLDAPR	y5++	3,4	4,3	2,4	19,9	0,998
<u>C</u>SLNPEWNETFR	y8++	5,4	2,4	2	1,6	0,984
<u>C</u>SNILLNNSGQIK	y2+	4	1,3	1,8	20,1	0,993
<u>C</u>VSAHPGLLQTLAGPLQTPR	y7++	3,3	4,3	1	26,8	0,999
<u>C</u>VSGSTAISTYPK	y2+	7,6	9,7	2,2	23	0,993
DGV<u>C</u>AQIEK	y6++	10,2	14,4	5	22,7	0,99
DLE<u>C</u>ALLLR	b2+	5,4	4,9	1,5	12,4	0,979
DLLADLIG<u>C</u>QR	a2+	9,7	8	3,8	8,7	0,995
DQ<u>C</u>IVISGESGSGK	y2+	2,9	8	2,9	8,6	0,962
EAPP<u>C</u>HLLIIHR	y4++	10,1	10,4	2,9	13,1	0,998
EASS<u>C</u>AVNLVLR	y3+	4,4	11,8	1,7	14,2	0,958
E<u>C</u>VYIIPSSK	y7++	5,8	4,8	3	14,4	0,984
EGF<u>C</u>LQNVDK	y2+	4,3	3,8	2,8	16,7	0,976
ESD<u>C</u>HYLEAR	y2+	0,9	4,2	2,8	15,4	0,998
FEVALLPEP<u>C</u>R	b2+	5	2,2	0,9	15,2	0,987
FIDFIDT<u>C</u>LIK	a2+	4,2	1,8	0,9	13,4	0,985
FLSPQLIAEEF<u>C</u>LK	a2+	2,5	1,9	0,7	18	0,999
GE<u>C</u>VITPSTDVK	y8++	9,1	3	2,3	10,8	0,989
GELLGSG<u>C</u>FGVVHR	y2+	4,5	4	28,1	10,3	0,99
GEV<u>C</u>NAVLAK	y4+	4,6	6,8	1,8	29,6	0,995
GLNLE<u>C</u>LWEK	y8++	22,8	13	5	11,5	0,981
GPLS<u>C</u>DTVLK	y8++	12,7	13	1,3	16,3	0,977
GS<u>C</u>GDLPGLR	y3+	5,7	3,2	6,2	23,8	0,965
GSLALD<u>C</u>APLLR	b3+	32,2	6,3	2,8	15,2	0,988
GT<u>C</u>IPNAEEVDVPIK	b3+	6,7	3,9	2,2	16,7	0,999
HIGSINPN<u>C</u>NVLEVIK	b6++	9,3	6	1,7	16,7	0,988
HLSS<u>C</u>AAPAPLTSAER	y4+	11	9,3	6,4	9,2	0,999
HYQTIQEAGDW<u>C</u>VPSTEPK	y6+	9,7	4,1	4,5	1,6	0,999
I<u>C</u>QAHFAPISPGFSR	y6+	8,3	7,8	1,5	20	0,999
IN<u>C</u>AAAGNPFPVR	y3+	0,8	0,9	2	16,7	0,983
IT<u>C</u>VETVVPEK	y3+	2,9	3,3	3,2	20,8	0,958
IVQTL<u>C</u>TVR	b3+	3,6	5,4	2,5	13,9	0,99
L<u>C</u>DFGISGQLVDSIAK	y3+	9,1	2,6	3,3	12	0,975
LFQQILSGVDY<u>C</u>HR	b6++	16,8	22,6	7,5	18,4	0,992

LQCVGDTAPLPWHIR	y7++	11	7,5	6,7	16,9	0,999
LTLPIPSTCPEPFAR	y12++	3,8	4,2	2	16,6	0,998
LVQCVDAFEEK	b3+	20,3	10,1	5,5	4,6	0,998
LYLAENYCFDSIPK	y2+	2,9	2,1	1	10,9	0,972
MLCDQYYLSSPELK	y7+	6,8	34,4	20,8	7,3	0,998
NILVGENTLCK	y9++	9,8	7,5	3,2	4,1	0,972
NILWSAENECFK	y9++	10,3	8,4	1,7	8,3	0,958
NTGIICTIGPASR	b3+	7,7	9	0,5	18,2	0,98
SNYQGLPVPCVK	y4+	8,1	8,4	0,6	15,3	0,998
SQCLSTLVRPVFGELK	y2+	2,6	1,2	0,7	15,8	0,998
TCVFFEAPGVR	y8+	5,4	6,5	4,1	14,6	0,998
TICGTPNYLSPEVLNK	y2+	6,3	4,4	2,4	18,7	0,998
TLGIGAFGEVCLAR	b3+	11,9	8,2	2,2	7,8	0,986
TVAACDLLQSLLHK	b3+	4,4	4,2	1,8	14,3	0,978
VHSENNACINFK	b6++	10,4	4,2	4,3	11,2	0,989
YSFGATCVK	y2+	8,4	7,7	3,2	19,9	0,996

Tableau 13 : CV (%) obtenus sur 3 réplicas en mode DIA-LID à 473 nm pour les 65 peptides à cystéine dérivés ajoutés dans 100 µg d'E.Coli pour les 4 points de concentration

Il a été possible de quantifier les 65 peptides à cystéine dérivés avec au moins une transition sur les 4 points de la gamme dynamique de 3 décades. Pour 50 peptides à cystéine dérivés, les CV de chacun des 4 points étaient inférieurs à 20 %, 10 peptides à cystéine dérivés avaient un CV compris entre 20 et 25 % pour un point de la gamme. Pour 4 autres, 1 point possédait un CV supérieur à 25 % (26, 28, 29 et 32 %). Enfin un peptide possédait un point avec un CV entre 20 et 25 % (20,8 %) et un autre supérieur à 25 % (34 %). Sur les 260 points de mesures réalisés, 244 points ont un CV inférieurs à 20 % soit 94 % de tous les points. La figure 55 présente les XIC obtenus lors des 3 réplicas du peptide à cystéine dérivé APEILLGLPFCEK au point 0,001. Tous ces résultats montrent la capacité de la méthodologie DIA-LID à 473 nm pour quantifier de manière répétable des peptides à cystéine dérivés dans des matrices biologiques ainsi que la bonne répétabilité de la méthode d'activation LID à 473 nm en elle-même.

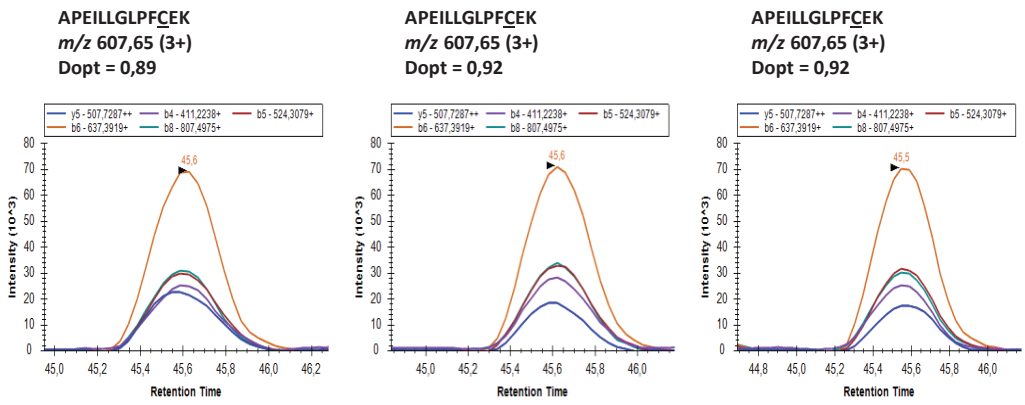


Figure 55 : XIC des 3 réplicas du peptide à cystéine dérivé APEILLGLPFCEK obtenus en mode DIA-LID à 473 nm au point de gamme 0,001

Nous pouvons constater que pour les 6 % des points de mesures restants (CV supérieur à 20 %), la variabilité est indépendante des peptides à cystéine dérivés analysés ou des points de la gamme concernés. En revanche, nous pouvons observer que tous les points de 1 ont un CV souvent nettement supérieur aux 3 autres points de la gamme. Cela pourrait être expliqué par une saturation du signal. Egalement, dans la plupart des cas, mais de moindre manière, les points à 0,001 présentent des CV supérieurs à ceux obtenus pour les points à 0,1 et 0,01 probablement en raison de la LOQ. Un point à 0,0001 a été tenté, celui-ci n'a pu être validé laissant penser que la méthodologie se limite à une gamme dynamique de 3 décades. De plus, nous pouvons observer une non homogénéité voire cohérence dans la distribution des CV entre les peptides à cystéine dérivés et parfois même d'un point à l'autre pour un même peptide. Cela semble inhérent à la méthode qui est donc moindre en termes de fidélité que la MRM où, au laboratoire, nous pouvons atteindre des CV inférieurs à 5 % sur tous les points.

Concernant la linéarité, il a été possible d'établir une courbe de calibration pour 36 peptides à cystéine dérivés avec un coefficient de corrélation (R^2) supérieur à 0,990 soit 55 % des peptides à cystéine dérivés quantifiés. Le minimum étant à 0,953. Nous sommes loin des linéarités atteintes en MRM au laboratoire supérieur à 0,990 pour la quasi-totalité des peptides quantifiés mais cela est prometteur. Ces résultats positifs nous ont poussé à approfondir dans cette voie et à estimer la LOQ de la méthodologie en utilisant non plus des concentrations normalisées mais des concentrations réelles. 1 peptide à cystéine pur à 70 % issu d'une protéine kinase humaine et présent au sein de la bibliothèque a été commandé chez JPT. Il s'agit du peptide VMCDSNDLPK issu de la protéine kinase SgK196. Celui-ci a été introduit dans 100 µg d'E.Coli à 8 niveaux de concentration différents en vue d'être quantifié sur 3 décades. Les échantillons ont été dérivés suivant le protocole type, séparés à l'aide du gradient 2 et analysés en mode DIA-LID à 473 nm en triplicats (Tableau 14 et Figure 56) :

	VMCDSNDLPK (y8 ++)	
Concentration (ng/100 µg)	**Aire**	**CV (%)**
1	200733	20
5	789212	15
10	2015548	3
50	10682414	0
100	21968041	11
250	36774976	1
500	62886917,33	5
1000	113370181,3	10

Tableau 14 : Aire et CV (%) obtenus pour 3 réplicas en mode DIA-LID à 473 nm pour le peptide à cystéine dérivé VMCDSNDLPK ajouté dans 100 µg d'E.Coli pour les 8 points de concentration

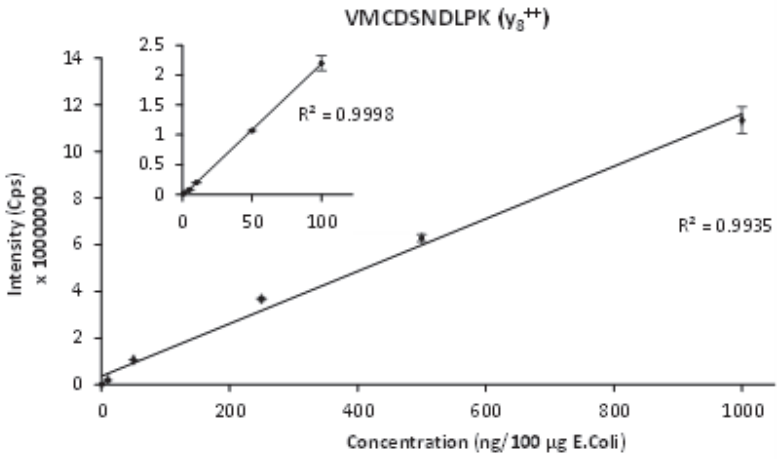


Figure 56 : Linéarité obtenue en mode DIA-LID à 473 nm pour le peptide à cystéine dérivé VMCDSNDLPK ajouté dans 100 µg d'E.Coli

Ce peptide à cystéine dérivé a ainsi pu être quantifié sur 3 décades à travers 8 points de gamme jusqu'à une valeur de LOQ de 1 ng dans 100 µg de E.Coli soit 1 dans 100 000 avec un R^2 supérieur à 0,99 et des CV inférieurs à 20 % pour tous les points à l'aide de la transition y_8^{++}. Qui plus est, une courbe de calibration entre 1 et 100 ng dans 100 µg de E.Coli a aussi été tracée (Figure 56) en vue de démontrer la linéarité réelle à petites concentrations de la méthodologie, non biaisée par les points du haut de la gamme. La linéarité est encore meilleure pour cette gamme de plus faible concentration. Des résultats similaires ont pu être obtenus pour 2 autres peptides à cystéine dérivés purs à 70 % issus de protéines humaines non kinases que nous possédions au laboratoire et que nous avons pu incorporer à notre bibliothèque de spectres : DICNDVLSLLEK et YMPQNPCIIATK (Annexe 3). Pour les transitions y_5^+ et b_2^+, des valeurs de LOQ de 1 ng et 5 ng dans 100 µg ainsi que des R^2 supérieurs à 0,99 et 0,98, respectivement, ont pu être obtenus. Ces deux résultats confirment le précédent n'en faisant pas une exception. Des tests ont été effectués à des concentrations inférieures à 1 et 5 ng dans 100 µg de E.Coli sans succès posant la LOQ à ce niveau et confirmant la gamme dynamique de 3 décades.

A travers les résultats obtenus nous pouvons établir le potentiel de la méthodologie DIA-LID à 473 nm pour quantifier de manière répétable des peptides à cystéine dérivés dans des matrices biologiques sur une gamme dynamique de 3 décades jusqu'à une LOQ de l'ordre de 1 ng dans 100 µg de matrice.

Devant ces résultats nous avons poursuivi notre étude en tentant de détecter des peptides à cystéine dérivés issus de protéines kinases humaines endogènes au sein d'un extrait cellulaire mammaire cancéreux humain.

V) Utilisation de la méthodologie DIA-LID à 473 nm et d'une bibliothèque de spectres pour identifier et quantifier des peptides à cystéine dérivés endogènes issus de protéines kinases humaines au sein d'un extrait cellulaire humain

Cette section sera consacrée à l'identification et à la quantification de peptides à cystéine dérivés endogènes issus de protéines kinases humaines au sein d'un extrait cellulaire mammaire cancéreux humain à l'aide de la méthodologie DIA-LID à 473 nm et d'une bibliothèque de spectres LID.

La raison qui nous a poussé à nous tourner vers un extrait cellulaire mammaire cancéreux humain, outre la présence attendue des kinases au sein des cellules humaines, est que les kinases sont suspectées d'avoir un rôle potentiel dans le développement des cancers du sein chez la femme comme le cancer $HER2^+$ qui représente à lui-seul entre 15 et 20 % des différents cancers du sein et implique une tyrosine kinase[220]. En vue d'approfondir cette voie, un extrait nous a été fourni par le centre de recherche en cancérologie affilié au Centre hospitalier Léon Berard à Lyon.

Cet échantillon a été dérivé à l'aide du protocole type, séparé à l'aide du gradient 2 et analysé en mode DIA-LID à 473 nm en triplicat.

Les 354 peptides à cystéine dérivés contenus dans la bibliothèque ont été sondés au sein de cet échantillon. Seulement 4 ont pu être identifiés à l'aide des critères préalablement établis dont 3 quantifiés avec des CV inférieurs à 20 %.

Voici leurs XIC ainsi que les protéines kinases auxquels ils appartiennent (Tableau 15 et Figure 57) :

Peptides	Transition	CV (%)	Proteines kinases
NTGIICTIGPASR	y_6^+	1	Pyruvate kinase PKM
SDIQDSLCYK	b_2^+	30	Hormonally up-regulated neu tumor-associated kinase
STVIENPGALCSPQSQR	y_6^+	16	Tyrosine-protein kinase FRK
VPLCINPQSAAFK	y_8^+	8	Serine/threonine protein kinase

Tableau 15 : Liste des 4 peptides à cystéine dérivés endogènes détectés au sein de l'extrait cellulaire en mode DIA-LID à 473 nm. La liste des protéines kinases humaines desquels ils sont issus, leur meilleure transition et le CV correspondant sont également présentés

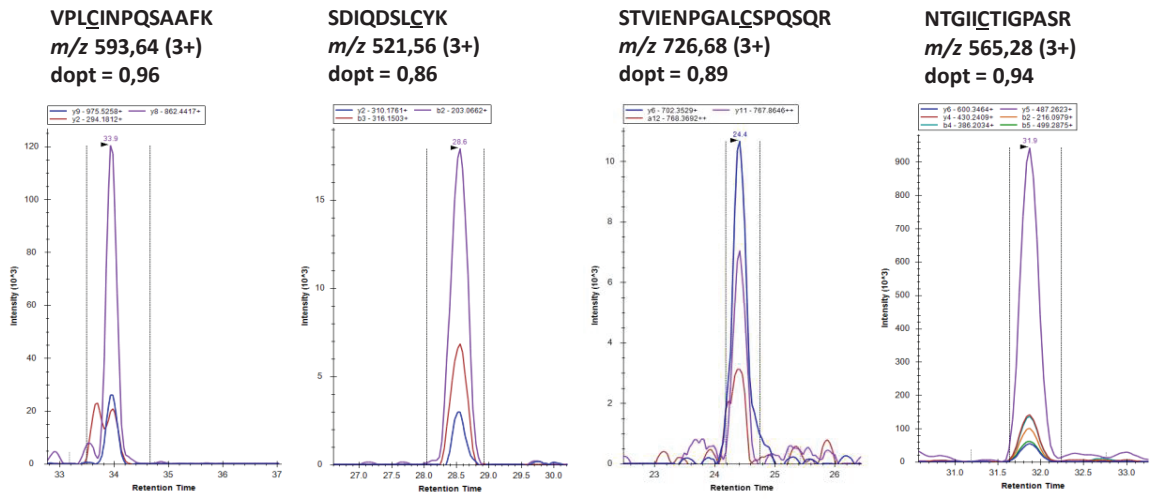


Figure 57 : XIC des 4 peptides à cystéine dérivés endogènes détectés en mode DIA-LID à 473 nm dans l'extrait cellulaire issus de protéines humaines kinases

Au vu des résultats nous pouvons faire deux observations principales : Premièrement nous avons pu identifier et quantifier des peptides à cystéine dérivés endogènes issus de protéines kinases humaines (via les critères précédemment établis) au sein d'extrait cellulaire mammaire cancéreux à l'aide la méthodologie DIA-LID à 473 nm et de la bibliothèque de spectres. Deuxièmement nous en avons détecté très peu. 4 peptides à cystéine dérivés rapporteurs de 4 protéines kinases humaines ont pu être identifiés. Pour 3 d'entre eux des CV inférieurs à 20 % ont pu être obtenus ouvrant la possibilité d'une quantification relative et du suivi de la régulation de 3 protéines kinases humaines. 1 peptide à cystéine dérivé a pu être suivi à l'aide de plus de quatre fragments. Les autres peptides à cystéine dérivés ont pu être identifiés avec 3 fragments (dont des ions y_9^+, y_8^+ y_6^+, y_{11}^{++} suffisamment spécifiques).

Peu de peptides à cystéine dérivés issus de protéines kinases ont pu être détectés et ils l'ont été avec peu de fragments. Cela peut s'expliquer de plusieurs manières : tout d'abord comme précisé précédemment, les protéines kinases semblent être faiblement exprimé au sein d'extrait cellulaire d'où l'utilisation de kit d'enrichissement pour augmenter les chances de les visualiser. Cela n'a pu être effectué au laboratoire. En outre, il faut aussi ajouter un souci inhérent au mode DIA. En effet en DDA, lors de la fragmentation, les ions sont sélectionnés 1 par 1, donc la quantité d'ions fixée dans la méthode sera majoritairement composée de l'ion précurseur d'intérêt. Or en DIA, les fenêtres étant de l'ordre de 32 Th, la même quantité d'ions sera répartie sur une multitude d'ions précurseurs différents abaissant le signal de tous et ne permettant pas de les visualiser (cela peut expliquer aussi en partie la non détection des 65 peptides à cystéine dérivés dans la section IV.1.a de ce chapitre). En raison de la diminution intrinsèque de la sensibilité liée à la DIA comparativement à la MRM-PRM, la méthodologie DIA ne semble pas adapter, sans étape d'enrichissement au préalable, pour identifier et quantifier des peptides à cystéine peu

abondants au sein de matrices biologiques (la gamme dynamique de 3 décades obtenue en DIA-LID est insuffisante).

À la vue de ces résultats, il a été décidé d'essayer de sonder d'autres peptides à cystéine dérivés issus d'autres protéines humaines susceptibles d'être exprimées plus intensément au sein de l'extrait cellulaire. Ils devraient ainsi être plus aisément détectables par DIA-LID à 473 nm. Au laboratoire nous possédions 63 peptides à cystéine supplémentaires issus de 33 protéines humaines. Ceux-ci ont été préparés selon le protocole type, séparés avec le gradient 2 puis analysés en DDA-LID à 473 nm en vue d'obtenir leurs spectres de photo-fragmentation et de les incorporer à la bibliothèque de spectres.

Sur les 63, 47 peptides à cystéine dérivés ont pu être détectés, validés et incorporés à la bibliothèque de spectres (Annexe 4). Ils ont ensuite été sondés au sein du même extrait cellulaire que précédemment. En effet, en DIA il n'y a aucune sélection en précurseur, tout ce qui est potentiellement détectable par l'appareillage est fragmenté. Il est ainsi possible d'enrichir les bibliothèques de spectres au cours du temps et de retraiter les échantillons passés avec des bibliothèques de spectres renouvelées. Ainsi il est possible d'identifier des composés qui n'étaient pas encore présents dans la bibliothèque à ce moment-là et donc non recherchés. Voici les résultats obtenus (Tableau 16 et Figure 58) :

Peptides	Transition	CV (%)	Protéines
DI<u>C</u>NDVLSLLEK	y_6^+	6	14-3-3 protein zeta/delta
EEAPDIL<u>C</u>LQETK	y_2^+	8	DNA-(apurinic or apyrimidinic site) lyase
GL<u>C</u>AIAQAESLR	y_6^+	2	40S ribosomal protein S3
I<u>C</u>SWNVDGLR	y_8^{++}	3	DNA-(apurinic or apyrimidinic site) lyase
LDYFLLSHSLLPAL<u>C</u>DSK	b_{11}^{++}	8	DNA-(apurinic or apyrimidinic site) lyase
V<u>C</u>NPIITK	y_5^+	2	Heat shock cognate 71 kDa protein
VNQIGSVTESLQA<u>C</u>K	y_{12}^{++}	3	Alpha-enolase

Tableau 16 : Liste des 7 peptides à cystéine dérivés endogènes détectés en mode DIA-LID à 473 nm au sein de l'extrait cellulaire. La liste des protéines humaines desquels ils sont issus, leur meilleure transition et le CV correspondant sont également présentés

Voici 6 des XIC des 7 peptides à cystéine dérivés détectés :

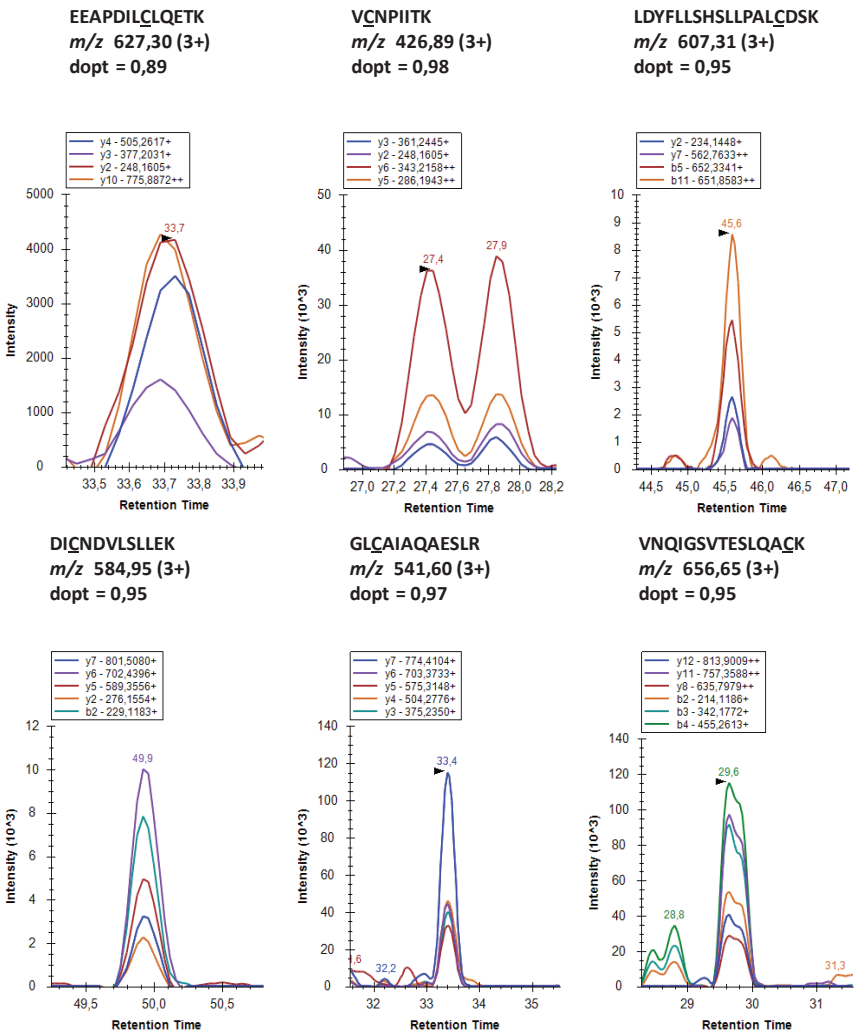


Figure 58 : XIC des 7 peptides à cystéine dérivés endogènes détectés en mode DIA-LID à 473 nm dans l'extrait cellulaire issus de protéines humaines

Via les 47 peptides à cystéine ajoutés à la bibliothèque de spectres 7 nouveaux peptides à cystéine dérivés rapporteurs de 5 protéines humaines ont pu être identifiés. Des CV inférieurs à 8 % ont été obtenus pour tous les peptides à cystéine dérivés ouvrant la possibilité d'une quantification relative et du suivi de la régulation des 5 protéines humaines. 3 peptides à cystéine dérivés sont par ailleurs rapporteurs de la même protéine pouvant alors corréler la réponse de ceux-ci pour confirmer le comportement suspecté de la protéine. Tous les peptides à cystéine dérivés ont été identifiés avec un minimum de 4 fragments (3 peptides ont pu être identifiés avec 6 fragments). Les protéines humaines correspondantes sont donc bien exprimées plus intensément que les protéines kinases humaines recherchées au début.

Si la méthodologie DIA-LID à 473 nm semble requérir une étape d'enrichissement en vue de maximiser le nombre de peptides à cystéine dérivés identifiés, elle promet une grande confidence

dans l'identification. Sur tous les peptides à cystéine dérivés identifiés très peu voire quasiment aucun fragment n'a été interféré. Cela permet d'une part d'identifier sans doute le peptide à cystéine dérivé putatif et d'autre part d'avoir un choix certains quant à la transition à sélectionner pour quantifier relativement le peptide à cystéine dérivé et donc la protéine. Il semble donc réellement nécessaire d'obtenir un gain en sensibilité car la méthodologie est prometteuse.

VI) Perspective

En vue de répondre à la problématique d'enrichissement et de perfectionner la méthodologie DIA-LID à 473 nm, 2 approches ont commencé à être développées. La première se base sur une étape de dérivation des peptides à cystéine avec le DABCYL-C2 maléimide après l'étape de digestion à la trypsine et la seconde sur l'utilisation d'un nouveau chromophore.

VI.1) Dérivation des peptides à cystéine après l'étape de digestion à la trypsine

Dans l'optique de maximiser la sensibilité lors d'une analyse DIA-LID à 473 nm il a été tenté de dériver les peptides à cystéine avec le DABCYL-C2 maléimide après l'étape digestion. En effet, comme aucun dénaturant n'est utilisé lors de l'étape de dérivation, il a été supposé que le rendement de dérivation, bien que répétable, pût ne pas être maximal en raison d'une possible non dénaturation optimale des protéines. Post-digestion, une fois tous les peptides générés, la question d'accessibilité au site de dérivation ne se posent plus. Le temps d'activité du TCEP en tant que réducteur n'était pas connu, nous avons donc simplement supposé qu'il serait encore actif après 1 nuit à digérer.

Cette expérience a été tentée sur un mélange des 3 protéines suivantes : Albumine humaine, sero-transferrine humaine et lysozyme de poulet. Les 3 protéines (conservées sous forme de poudre, Sigma Aldrich) ont été solubilisées dans une matrice aqueuse (400 µg/mL pour chaque protéine). 100 µL de cette solution a été dérivé suivant le protocole type (sauf que cette fois-ci l'étape de dérivation a eu lieu après celle de digestion) puis séparé à l'aide du gradient 2 et analysé en mode DDA-LID à 473 nm. Voici les résultats obtenus pour les 10 ions précurseurs les plus intenses issus de 10 peptides à cystéine dérivés issus de ces 3 protéines (Tableau 17) :

Peptides à cystéine	Intensité dérivation avant digestion	Intensité dérivation après digestion	Gain en signal (%)
GYSLGNWVCAAK	1,22E+07	6,92E+07	5,67
WWCNDGR	5,27E+06	1,82E+07	3,45
RPCFSALEVDETYVPK	1,37E+05	2,66E+06	19,42
QNCELFEQLGEYK	7,56E+05	5,28E+06	6,98
YICENQDSISSK	1,11E+06	2,01E+07	18,11
TCVADESAENCDK	1,04E+04	8,67E+05	83,37
AAFTECCQAADK	1,58E+05	1,17E+06	7,41
ECCEKPLLEK	2,84E+04	1,53E+06	53,87
DECMVK	4,52E+05	8,00E+05	1,77
ALVLIAFAQYLQQCPFEDHVK	8,37E+03	4,06E+05	48,51

Tableau 17 : Gain en signal de 10 ions précurseurs issus de 10 peptides à cystéine dérivés issus des 3 protéines digérées dû à une étape de dérivation réalisée post digestion

Nous pouvons voir que par le biais de cette étape de dérivation post-digestion, un gain de signal est observé pour la majorité des peptides à cystéine dérivés et aucune perte de signal n'est observée. Le gain moyen est de 25 % et donc réellement significatif. Cette étape n'a malheureusement pu être tentée directement sur l'extrait cellulaire mammaire cancéreux mais semble prometteuse et devrait être essayée prochainement.

L'enrichissement semble donc être une des solutions à développer pour améliorer la détection des peptides à cystéine dérivés endogènes issus de protéines kinases humaines ou non dans des matrices biologiques.

VI.2) Dérivation des peptides à cystéine à l'aide d'un nouveau chromophore

En vue de perfectionner la méthodologie DIA-LID à 473 nm et de l'amener à se rapprocher d'une application concrète hors du milieu de la recherche, nous nous sommes tournés vers l'utilisation d'un nouveau chromophore pour dériver les peptides à cystéine.

En effet, si le DABCYL C2-maléimide s'est révélé efficace pour dériver des peptides à cystéine celui-ci est peu soluble dans l'eau[204]. Cette caractéristique complique l'étape de préparation d'échantillon et rend difficile l'utilisation de la méthodologie entière en routine dans des laboratoires d'analyses ou des hôpitaux où ce sont la rapidité, la simplicité et l'efficacité qui sont demandées. Il est donc nécessaire de mettre au point, à partir de cette base de chromophore DABCYL, un chromophore pouvant à la fois avoir un bon taux de photo-fragmentation à 473 nm mais qui soit aussi, idéalement, soluble dans l'eau.

Dans cette optique, nous avons établi une collaboration avec l'équipe de Yann Bretonnière du laboratoire de chimie de l'ENS de Lyon dans le but de synthétiser ce chromophore adéquat. Le principe est de rajouter des substituants sur le chromophore qui vont améliorer sa solubilité sans perturber ses propriétés de photo-fragmentation. Durant ma thèse, j'ai pu tester un nouveau chromophore issu de ce projet, intitulé JCM 1425 A et qui s'avère prometteur.

VI.2.a) Structure du chromophore JCM 1425 A

Le chromophore JCM 1425 A (Mw = 579,30 g/mol) a été synthétisé sous forme de poudre au sein du laboratoire de chimie de l'ENS de Lyon et possède la structure suivante (Figure 59) :

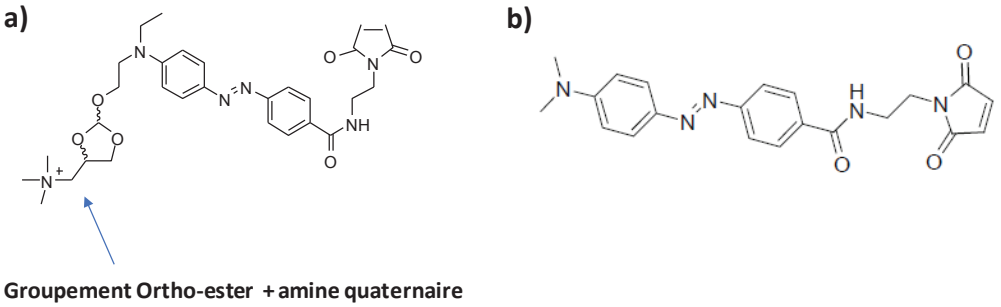


Figure 59 : Structure du chromophore JCM 1425 A (a) et du DABCYL C2-maléimide (b)

Nous pouvons voir que le chromophore JCM 1425 A possède un groupement ortho-ester + amine quartenaire en plus par rapport au DABCYL C2-maléimide. La partie amine quaternaire est supposée permettre la solubilité de JCM 1425-A en milieu aqueux et ainsi remédier au problème inhérent au DABCYL C2-maléimide. Dans le même temps, la partie ortho-ester doit induire l'hydrolyse du groupement entier à un pH acide (pH inférieur à 5) nécessaire pour la photo-fragmentation du chromophore (le chromophore sous forme non hydrolysé est censé ne pas photo-fragmenté). Le chromophore hydrolysé, lui est supposé être photo-fragmenté, il sera intitulé par la suite JCM 1425 A' (Mw = 435,20 g/mol). Cette approche permettrait donc de dériver les acides aminés cystéine directement en matrice aqueuse (pH = 7) avec le chromophore soluble puis, lors de l'acidification post digestion, de rompre ce groupement et de générer des peptides à cystéine dérivés avec un chromophore ayant une structure très proche de celui du DABCYL C2-maléimide et donc un taux de photo-fragmentation supposé similaire à 473 nm (Figure 60) :

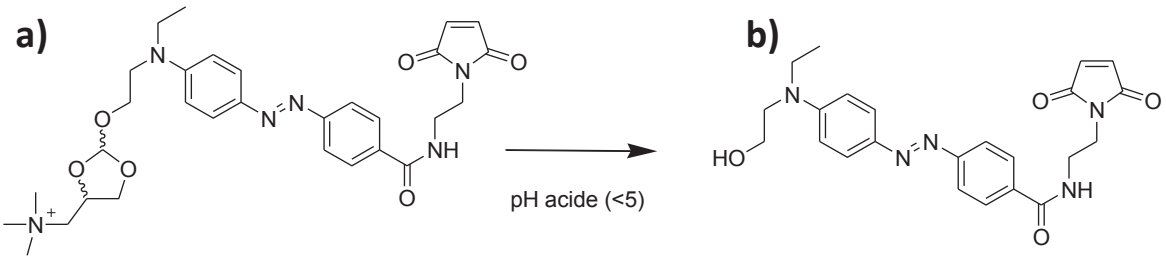


Figure 60 : Structure du chromophore JCM 1425 A attendu à pH = 7 (a) et pH inférieur à 5 (b)

VI.2.b) Photo-fragmentation du chromophore sous forme non hydrolysée (JCM 1425 A) et hydrolysée (JCM 1425 A') à 473 nm

En vue de confirmer les hypothèses faites précédemment, le chromophore JCM 1425 A a été solubilisé au sein d'une matrice 100 % aqueuse (pH = 7) à une concentration de 18 µg/mL. La solubilité est total à l'œil nu (après sonication pendant 15 min). L'ajout du groupement amine quaternaire à bien permis la solubilisation de JCM 1425 A à 100 % en matrice aqueuse. Le chromophore JCM 1425 A permet donc de remédier au problème de solubilité inhérent au DABCYL C2-maléimide. La concentration testée est peu importante mais nous n'avions pas beaucoup d'échantillon. Cette solution a ensuite été directement infusée au sein du spectromètre de masse puis analysée en mode FMS et PRM-LID à 473 nm pour visualiser la composition de l'échantillon et la photo-fragmentation des composés présents (Figure 61) :

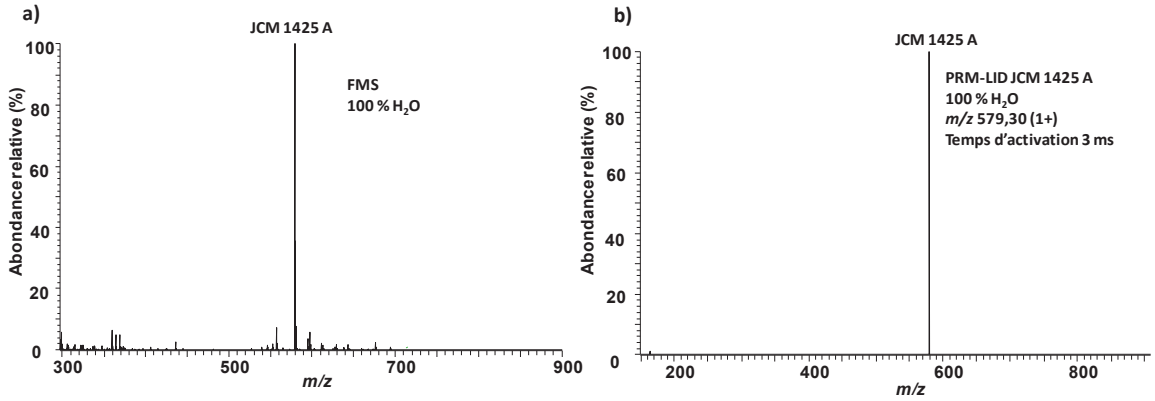


Figure 61 : Spectres FMS (a) et PRM LID à 473 nm (b) obtenus pour JCM 1425 A infusé à pH = 7 en matrice aqueuse

Nous pouvons observer qu'une seule forme [M^+] du chromophore est obtenue (JCM 1425 A) (Figure 61 a) et que cette forme est ne photo-fragmente pas en LID à 473 nm (Figure 61 b). L'hydrolyse du groupement ortho-ester + amine quartenaire semble donc essentiel si l'on veut faire photo-fragmenter les peptides à cystéine dérivés avec ce chromophore.

Une solution de JCM 1425 A a donc été préparée dans une matrice 100 % aqueuse à une concentration de 18 µg/mL (pH = 2,5 grâce à l'ajout d'AF) en vue d'obtenir JCM 1425 A'. La solution a ensuite été directement infusée au sein du spectromètre de masse puis analysée en mode FMS et PRM-LID à 473 nm (Figure 62) :

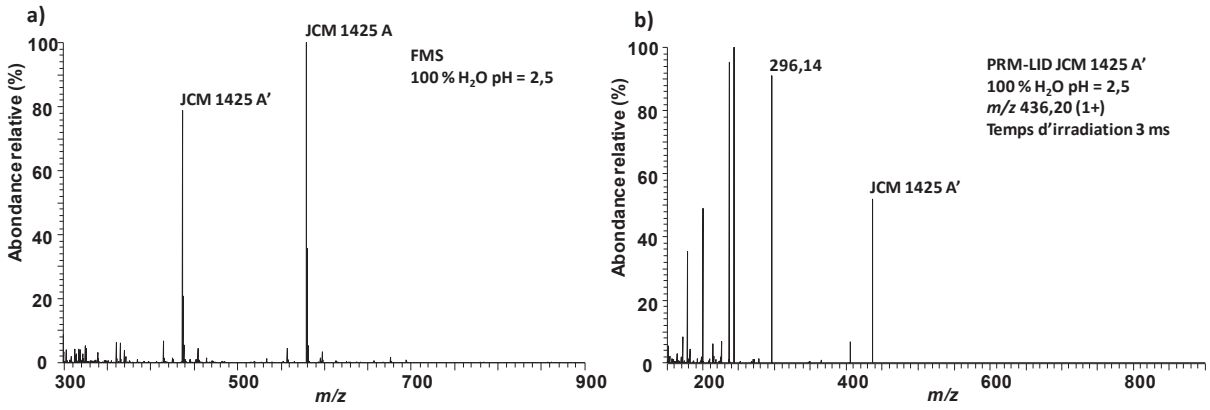


Figure 62 : Spectres FMS (a) et PRM LID à 473 nm (b) obtenus pour JCM 1425 A infusé à pH = 2,5 en matrice aqueuse

Nous pouvons tout d'abord observer qu'abaisser le pH à 2,5 n'est pas suffisant pour permettre la formation exclusive de JCM 1425 A'. A ce pH, la forme JCM 1425 A est encore intensément présente (Figure 62 a). La préparation d'échantillon doit être optimisé pour obtenir le chromophore exclusivement sous forme JCM 1425 A' après acidification. Ensuite, nous pouvons remarquer que la forme JCM 1425 A' photo-fragmente contrairement à la forme JCM 1425 A (Figures 62 b et 61 b). L'hydrolyse du groupement ortho-ester + amine quartenaire permet la photo-fragmentation du chromophore à 473 nm. Son taux de photo-fragmentation a été calculé à l'aide de l'équation 14. Le taux de photo-fragmentation obtenu est de 91 %. Ceci est un résultat comparable (même si légèrement inférieur) avec ce que l'on peut obtenir pour le DABCYL-C2-maléimide seul (97 %). Il semble donc possible de faire photo-fragmenter des peptides à cystéine dérivés avec JCM 1425 A'. Nous pouvons également noter la présence d'un ion rapporteur de *m/z* 296,14. Celui-correspond à l'ion rapporteur du DABCYL C2-maléimide (*m/z* 252,11) plus l'incrément de masse lié à la structure de JCM 1425 A' (2 atomes de carbone, 1 atome d'oxygène et 4 atomes d'hydrogène soit environ *m/z* 44) (Figure 63) :

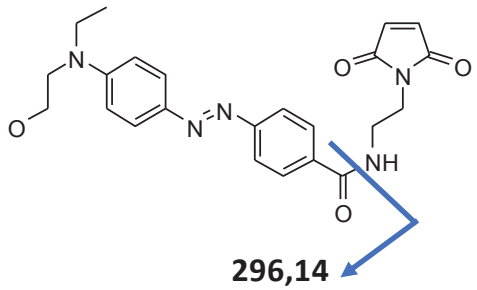


Figure 63 : Structure de JCM 1425 A' avec localisation du site de fragmentation générant l'ion rapporteur *m/z* 296,14

Le chromophore seul et tous les composés dérivés avec JCM 1425 A' (dont les peptides à cystéine) devraient produire cet ion rapporteur qui pourra agir en tant que marqueur sur un spectre de

fragmentation. En effet, dès que cet ion serait présent sur un spectre de fragmentation, alors le composé fragmenté aurait précédemment été dérivé avec JCM 1425 A'.

Néanmoins, avant d'utiliser ce chromophore pour dériver et identifier des peptides à cystéine dans des matrices biologiques, nous devons trouver un protocole expérimental permettant d'obtenir exclusivement la forme hydrolysée après acidification.

VI.2.c) Optimisation de l'étape d'hydrolyse de JCM 1425 A

En vue d'optimiser le protocole expérimental et d'hydrolyser JCM 1425 A pour former exclusivement JCM 1425 A', nous avons décidé de travailler directement avec un peptide à cystéine. Ce choix a été fait dans le but de mettre en place un protocole qui soit efficace directement pour des peptides à cystéine. Le peptide IVGGWECEK (rapporteur de l'antigène prostatique spécifique) a été choisi. Celui-ci était présent sous forme synthétique a haute pureté au laboratoire (80 % - Proteogenix) et a démontré son efficacité par le passé pour tester efficacement des chromophores[204].

Le protocole utilisé pour dériver IVGGWECEK avec JCM 1425 A est le suivant :

4 µg de IVGGWECEK ont été dérivé avec JCM 1425 A dans une solution aqueuse d'acétate d'ammonium 60 mM. Lors de cette étape, 3 équivalents molaires de TCEP ont été ajoutés pour réduire les ponts disulfures et 5 équivalents de chromophore pour dériver la cystéine. Les 2 réactifs ont été ajoutés au même moment et la solution a été gardée dans l'obscurité pendant 4h. Ensuite, le pH a été abaissé à 2,5 pour tenter d'obtenir la forme JCM 1425 A' du chromophore. L'échantillon a ensuite été infusé directement puis analysé en mode FMS.

Le peptide sous forme trichargée est attendu à une *m/z* de 533,29 s'il présente le chromophore non hydrolysé et à une *m/z* de 485,89 s'il présente le chromophore hydrolysé. Voici les résultats obtenus (Figure 64) :

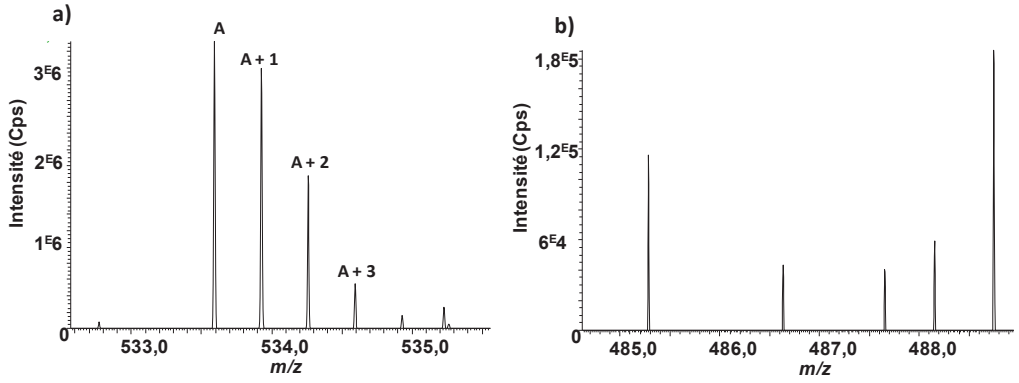


Figure 64 : Zoom sur le spectre FMS obtenu pour le peptide IVGGWE<u>C</u>EK dans les gammes de *m/z* attendues des formes trichargées avec le chromophore non hydrolysé (a) et hydrolysé (b)

Nous pouvons voir tout d'abord que la dérivation avec le chromophore a bien eu lieu : nous observons IVGGWE<u>C</u>EK dérivé avec JCM 1425 A à *m/z* 533,29 (Figure 64 a). Ceci était attendu car la fonction maléimide (spécifique des thiols) est toujours présente sur JCM 1425 A. En revanche, nous n'observons absolument pas le peptide IVGGWE<u>C</u>EK dérivé avec JCM 1425 A' (Figure 64 b). Nous nous attendions pourtant à le détecter car à pH = 2,5, JCM 1425 A' était bien formé lors de l'analyse du chromophore seul (les résultats sont identiques pour les formes dichargées). L'hydrolyse serait encore moins efficace lorsque le chromophore est greffé sur un peptide à cystéine.

Pour remédier à cela nous avons supposé qu'après dérivation il fallait chauffer l'échantillon en vue de catalyser la réaction d'hydrolyse. En effet, si les réactions d'hydrolyse sont athermiques (c'est-à-dire qu'un excédent de chaleur ne sera pas influent sur le taux d'avancement final de la réaction), elles sont plutôt lentes et une élévation de la température permet d'augmenter la cinétique de leurs réactions. Dans cette optique, le même protocole que précédemment a été réalisé suivi d'un chauffage à 40 °C pendant 15 et 30 min. Cela s'est révélé sans succès. Une autre tentative a été entreprise durant 45 min. Voici le résultat obtenu (Figure 65) :

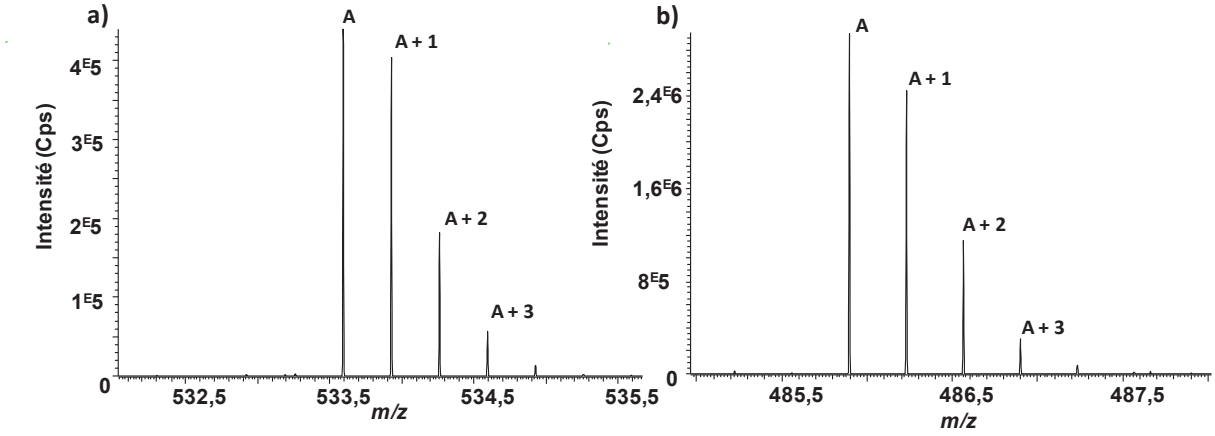


Figure 65 : Zoom sur le spectre FMS dans les gammes de *m/z* attendues des formes trichargées obtenu pour le peptide à cystéine dérivé IVGGWE<u>C</u>EK avec le chromophore non hydrolysé (a) et hydrolysé (b) après chauffage à 40°C pendant 45min

Nous pouvons voir ici que nous avons commencé à fortement hydrolyser le chromophore après un chauffage à 40°C pendant 45 min. En effet, nous pouvons voir la forme trichargée d'IVGGWE<u>C</u>EK avec le chromophore sous forme hydrolysée (JCM 1425 A') alors qu'elle était non détectée auparavant (Figures 65 b et 64 b). Dans le même temps, IVGGWE<u>C</u>EK avec le chromophore sous forme non hydrolysée (JCM 1425 A) est également présent mais bien moindrement en intensité que précédemment (4^E5 cps contre 3^E6 cps pour le mono-isotopique) (Figures 65 a et 64 a). L'impact de la température sur la cinétique de l'hydrolyse commence à

porter ces fruits. L'entreprise a été continuée avec un chauffage à 60°C pendant 6h en vue de s'affranchir définitivement de la forme non hydrolysée. Voici les résultats obtenus (Figure 66) :

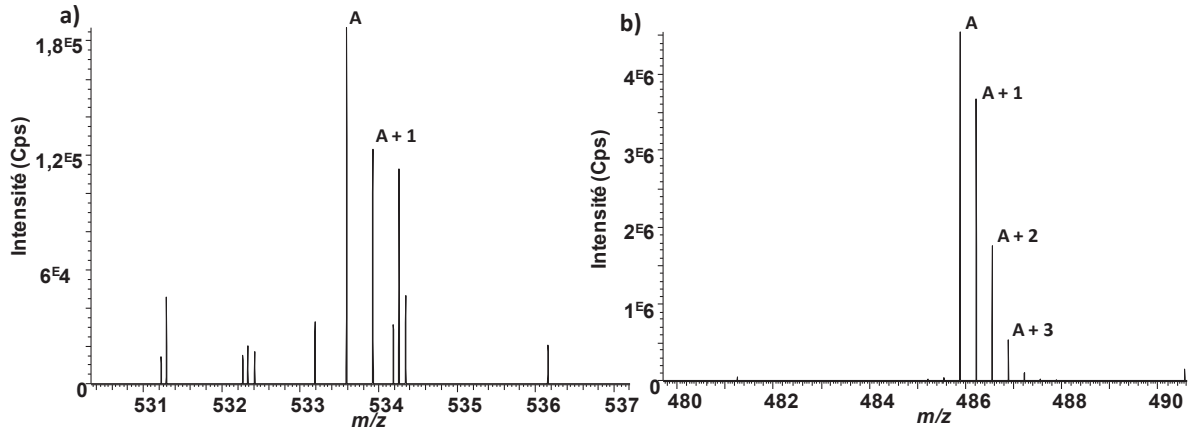


Figure 66 : Zoom sur le spectre FMS dans les gammes de *m/z* attendues des formes trichargées pour le peptide à cystéine dérivé IVGGWECEK avec le chromophore non hydrolysé (a) et hydrolysé (b) après chauffage à 60°C pendant 6h

Nous pouvons voir qu'avec l'étape de chauffage à 60°C pendant 6h nous pouvons générer majoritairement le peptide à cystéine IVGGWECEK dérivé avec le chromophore sous forme hydrolysée (JCM 1425 A') (Figure 66 b). Le gain en signal par rapport au chauffage à 45°C pendant 45 min est quasiment doublé et la forme du peptide IVGGWECEK dérivé avec le chromophore sous forme non hydrolysée (JCM 1425 A) est quasiment inexistante (Figure 66 a). Qui plus est, les formes dichargées et mono-chargées du peptide IVGGWECEK dérivé avec le chromophore sous forme non hydrolysée ne sont pas détectées.

Du fait d'un manque du chromophore nous n'avons pu pousser cette étude plus loin. Nous la considérons néanmoins comme quasiment aboutie. Un chauffage à 60°C pendant 6h après dérivation permet d'obtenir presque exclusivement la forme hydrolysée JCM 1425 A' du chromophore lorsque dérivé sur des peptides à cystéine.

VI.2.d) Comparaison de la photo-fragmentation laser du peptide à cystéine IVGGWECEK dérivé avec le DABCYL C2-maléimide et JCM 1425 A'

Une fois que nous avions réussi à obtenir majoritairement IVGGWECEK dérivé avec JCM 1425 A', nous avons tenté de comparer le taux de photo-fragmentation que nous pouvions obtenir à 473 nm avec ce chromophore comparativement au DABCYL C2-malémide. Il est important de rappeler que c'est avec leurs spectres de fragmentation que nous pouvons identifier les peptides à cystéine dérivés.

De ce fait, 4 µg de IVGGWECEK ont été dérivés selon le protocole type et selon le protocole spécifique à JCM 1425 A (dont l'étape de chauffage à 60°C pendant 6h). Les deux solutions ont

ensuite été infusées directement au sein du Qexactive et analysées en mode PRM-LID à 473 nm (Figure 67).

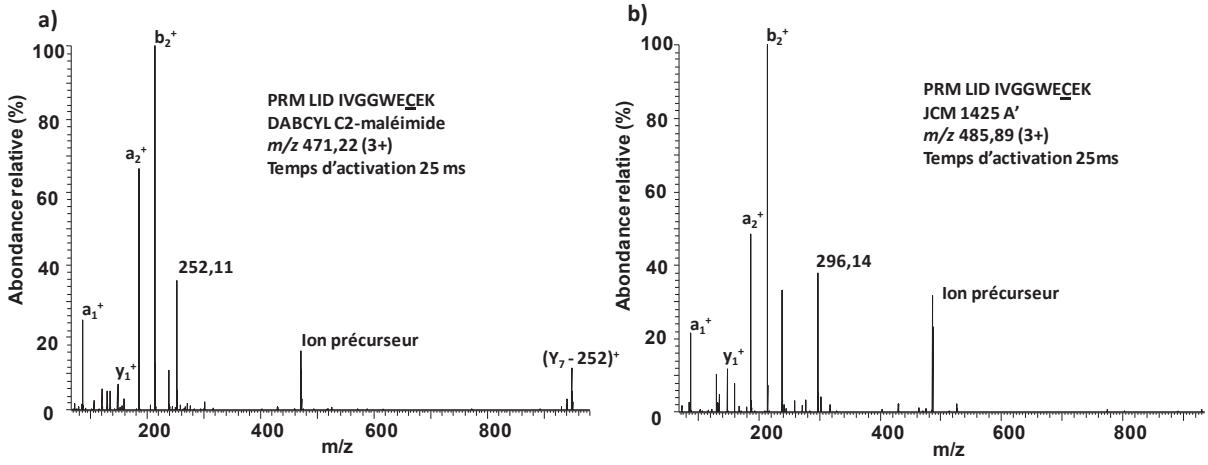


Figure 67 : Spectres de photo-fragmentation des formes trichargées du peptide IVGGWE<u>C</u>EK dérivés avec le DABCYL C2-maléimide (a) et JCM 1425 A' (b)

Nous pouvons voir tout d'abord que les spectres de photo-fragmentation obtenus avec le chromophore DABCYL C2-maléimide et JCM 1425 A' sont quasiment identiques. On peut noter quelques différences liées aux intensités relatives mais globalement ce sont les mêmes fragments qui sont produits à des intensités relatives très proches (mise à part le fragment y_7-251^+) dont les deux ions rapporteurs. Identifier des peptides à cystéine dérivés avec JCM 1425 A' semble donc possible. En termes de taux de photo-fragmentation les résultats suivants ont été obtenus (Equation 14 et Tableau 18) :

	Taux de photo-fragmentation IVGGWE<u>C</u>EK (%)
DABCYL C2-maléimide	91
JCM 1425 A'	85

Tableau 18 : Taux de photo-fragmentation du peptide à cystéine dérivé IVGGWE<u>C</u>EK avec le DABCYL C2-maléimide et JCM 1425 A'

Il semble que JCM 1425 A' induise un taux de photo-fragmentation légèrement inférieur à celui obtenu avec le DABCYL-C2 maléimide (85 % contre 91 %) (Tableau 18). Cela confirme la première observation qui avait été faite avec les deux chromophores seuls. D'autres tests ont été effectués avec des temps d'irradiation supérieurs (50, 75 et 100 ms) en vue de voir si cette observation résultait d'un temps d'activation insuffisant pour induire une photo-fragmentation totale du peptide à cystéine dérivé. Les mêmes taux de photo-fragmentation ont été obtenus. Dans le même temps il a également été effectué un test à un temps d'activation de 5 ms. Voici le résultat obtenu (Figure 68) :

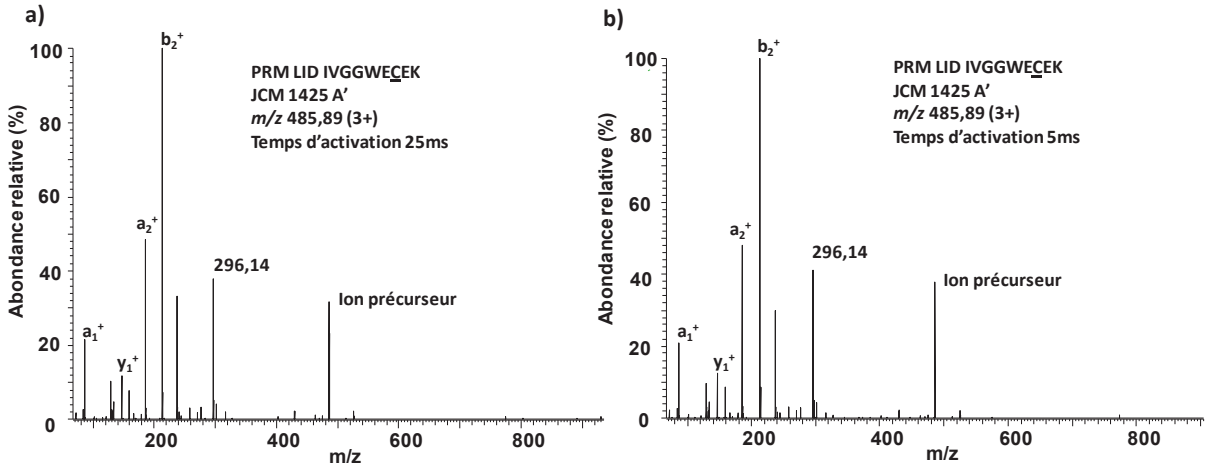


Figure 68 : Spectre de photo-fragmentation des formes trichargées du peptide à cystéine dérivé IVGGWECEK avec JCM 1425 A' avec un temps d'activation de 25 ms (a) et 5 ms (b)

Nous pouvons visualiser ici que le spectre de photo-fragmentation obtenu pour le peptide à cystéine dérivé avec JCM 1425 A' avec un temps d'irradiation de 5 ms est identique à celui de 25 ms et avec le même taux de photo-fragmentation (84 %) (Figure 68). Cela n'est pas obtenable avec le DABCYL C2-maléimide où, avec ce temps d'activation, le taux de photo-fragmentation est inférieur à celui obtenu à 25 ms[9]. Nous pourrions donc, par le biais de JCM 1425 A', photo-fragmenter plus rapidement les peptides à cystéine dérivés de manière optimal a contrario d'un taux de photo-fragmentation légèrement moindre. Cet apport n'est pas négligeable car il permettrait dans le cas d'analyse utilisant la méthode d'activation LID à 473 nm, dont la DIA, de paramétrer le Qexactive de la même manière que lors des analyses avec une activation collisionnelle type HCD (3 ms) et ainsi d'augmenter l'IT utilisé (et donc de gagner en sensibilité dans certains cas) tout en gardant la spécificité liée à la méthode d'activation LID.

VI.2.e) Conclusion

Pour conclure quant à une potentielle utilisation future du chromophore JCM 1425 A pour se substituer au DABCYL C2-maléimide pour dériver des peptides à cystéine nous pouvons dire ceci :

JCM 1425 A à démontrer qu'il remédiait au problème principal du DABCYL C2-maléimide : il est soluble en milieu aqueux. Ensuite, sa forme hydrolysée présente un taux de photo-fragmentation quasiment identique au DABCYL C2-maléimide (bien que légèrement moindre mais qui semble être suffisant pour identifier des peptides à cystéine dérivés dans des matrices biologiques) avec un temps d'activation 5 fois plus court. Ce faible temps d'activation est très attrayant dans l'optique, par exemple, des méthodologies DIA ou MRM où il permettrait, dans certains cas, de gagner en sensibilité ou de suivre plus de transition à la fois. Une inconnue quant au rendement de dérivation doit être éclairci mais celui-ci devrait être proche ou identique à celui du DABCYL C2-maléimide, JCM 1425 A' ayant quasiment la même structure. Cependant un point négatif

demeure lié à la préparation d'échantillon : il est long et difficile d'obtenir la forme hydrolysée JCM 1425 A'. En plus de l'étape de dérivation de 4h il faut au minimum 6h de chauffage à 60°C pour obtenir, presque exclusivement, cette forme. Cette étape doit finir d'être optimisée et démontrer sa répétabilité. Néanmoins, ce point négatif ne nous semble pas être un frein suffisant pour ne pas continuer à tester ce chromophore sur des digests de protéines ou des matrices biologiques et comparer les résultats obtenus avec ceux du DABCYL C2-maléimide (l'impact du temps de chauffage important sur les matrices étudiées doit être investigué). JCM 1425 A par sa solubilité et par son faible temps d'activation et taux de photo-fragmentation (JCM 1425 A') peut prétendre à remplacer le DABCYL C2-maléimide si l'étape d'hydrolyse s'avère répétable (à condition d'accepter une préparation d'échantillon plus longue et que le temps de chauffage ne nuise pas aux matrices analysées).

Conclusion

Au sein de ce chapitre nous avons démontré l'efficacité de l'utilisation du mode DIA couplé à la méthode d'activation LID à 473 nm et d'une bibliothèque de spectres pour identifier et quantifier de manière répétable des peptides à cystéine endogènes dérivés au sein de matrice biologiques humaines. 11 peptides à cystéines dérivé issus de 9 protéines humaines ont été identifiés et quantifiés au sein d'un extrait cellulaire mammaire cancéreux humain. La méthodologie peut donc être utilisée selon ce but. Il a également été démontré que l'utilisation de la méthode d'activation LID à 473 nm permettait de diminuer fortement la complexité des spectres de (co)-fragmentation en DIA comparativement à la HCD remédiant ainsi au problème principal de ce mode. Cet effet a permis une plus grande facilité au niveau de l'extraction de l'information et une plus grande confiance lors de l'identification de composés.

Cependant, notre objectif principal ciblait le suivi de 354 protéines kinases humaines en utilisant cette méthodologie : il n'est pas atteint (seulement 4 protéines kinases ont été identifiées). Plusieurs raisons peuvent justifiées cela : premièrement la faible expression des protéines kinases humaines au sein des extraits cellulaires humain. La réduction de la (co)-fragmentation n'a pas suffi à surmonter ce problème : il semble indispensable de se tourner vers l'utilisation de kit d'enrichissement. Deuxièmement, la diminution intrinsèque de la sensibilité liée au mode DIA comparativement aux autres modes (DDA, PRM). L'utilisation des fenêtres de 32 Th (ou même 25 comme en SWATH) induit une quantité d'ions maximale accumulée par ions précurseurs beaucoup plus faible que si les fenêtres étaient de 4 ou 2 Th : nous perdons en signal. Il semble donc nécessaire d'utiliser des fenêtres plus petites en DIA. Cela pourra être réalisé grâce à l'amélioration de l'appareillage.

Pour terminer, nous pouvons dire que la méthodologie DIA-LID à 473 nm est prometteuse mais qu'elle exige l'utilisation de kit ou de méthodes d'enrichissement pour être exploitée à son maximum. Ceci devrait être aisément accessible. Il semble donc important, en parallèle d'un développement d'un protocole selon ce but, de continuer à enrichir la bibliothèque de spectres en vue de couvrir le plus d'informations possibles. Il est fort probable qu'une fois le protocole optimisé, cette méthodologie nous permette d'accéder à des informations inédites.

Dans le chapitre suivant, les comportements de photo-fragmentation des peptides à cystéine dérivés avec le DABCYL C2-maléimide en LID à 473 nm seront étudiés. En effet, en éclaircissant les mécanismes mis en jeu il devient possible d'optimiser les méthodes d'identification et de quantification utilisant la LID (par anticipation du comportement de dissociation).

Chapitre III : Etude du modèle de fragmentation des peptides à cystéine dérivés avec le DABCYL C2-maléimide par LID à 473 nm

Introduction

Dans ce chapitre, il sera présenté une première investigation concernant le modèle de fragmentation des peptides à cystéine dérivés avec le DABCYL C2-maléimide par LID à 473 nm. Celle-ci a pour objectif de commencer à élucider les mécanismes mis en jeu mais également de déterminer les critères favorables à la sélection des peptides à cystéine optimaux en vue d'une détection et/ou d'une quantification par spectrométrie de masse dans une matrice déterminée. Et si aucun choix n'est possible, de pouvoir anticiper au mieux les comportements de ceux-ci.

Cette étude se concentrera principalement sur les types de fragments préférentiellement générés par cette méthode d'activation et l'impact de la présence de certains acides aminés. Nous verrons qu'un processus de redistribution d'énergie prend place et dépend de la taille (nombre d'acides aminés) des peptides à cystéine dérivés photo-fragmentés.

En vue d'étudier le modèle de fragmentation des peptides à cystéine dérivés avec le DABCYL C2-maléimide par photo-fragmentation laser à 473 nm nous nous sommes servis des 401 spectres de photo-fragmentation provenant des peptides à cystéine contenus dans la bibliothèque de spectres précédemment construite dans le chapitre II (354 spectres provenant des 354 peptides à cystéine issus des protéines kinases humaines et 47 spectres provenant des 47 peptides à cystéine issus des 33 protéines humaines).

Cette étude a été réalisée via un traitement manuel sur Excel et les informations fournies par la bibliothèque de spectres LID à 473 nm sur Skyline.

III.I) Précision quant aux résultats présentés dans ce chapitre

Il est à préciser que pour les études réalisées par la suite seulement les 6 fragments les plus intenses, pour chaque peptide à cystéine, ont été utilisés. Cette décision résulte d'une observation : les peptides à cystéine utilisés pour cette étude sont ceux achetés chez Thermo Scientific et sont de qualité seconde. De ce fait, pour certains peptides à cystéine, correctement synthétisés, un nombre important de fragments a pu être observé (parfois supérieur à 20) en raison d'une forte concentration de ces peptides à cystéine dans les solutions fournies. En revanche, pour beaucoup d'autres, la concentration était bien inférieure et ne permettait pas d'obtenir des spectres de photo-fragmentation aussi riches (inférieur à 10 fragments). Cette non homogénéité pouvait biaiser l'étude dès le départ de par le poids différent pris par chaque peptide à cystéine. Il a été observé que, tous les spectres de photo-fragmentation de la bibliothèque possédaient 6 fragments au minimum, ce chiffre a donc été gardé et les 6 fragments les plus intenses ont été retenus pour tous les peptides à cystéine de la bibliothèque en vue de cette étude. Qui plus est, en se focalisant sur les 6 fragments les plus intenses, nous pouvons mettre en avant les comportements les plus caractéristiques de la photo-fragmentation laser des peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm.

D'autre part, il est aussi important de préciser que les informations extraites concernant les fragments des peptides à cystéine dérivés avec le DABCYL C2-maléimide issus de l'activation par LID à 473 nm ne pourront être exhaustives.

En effet, voici pour exemple le spectre de photo-fragmentation du peptide ELGAALQGCLLR (Figure 69) :

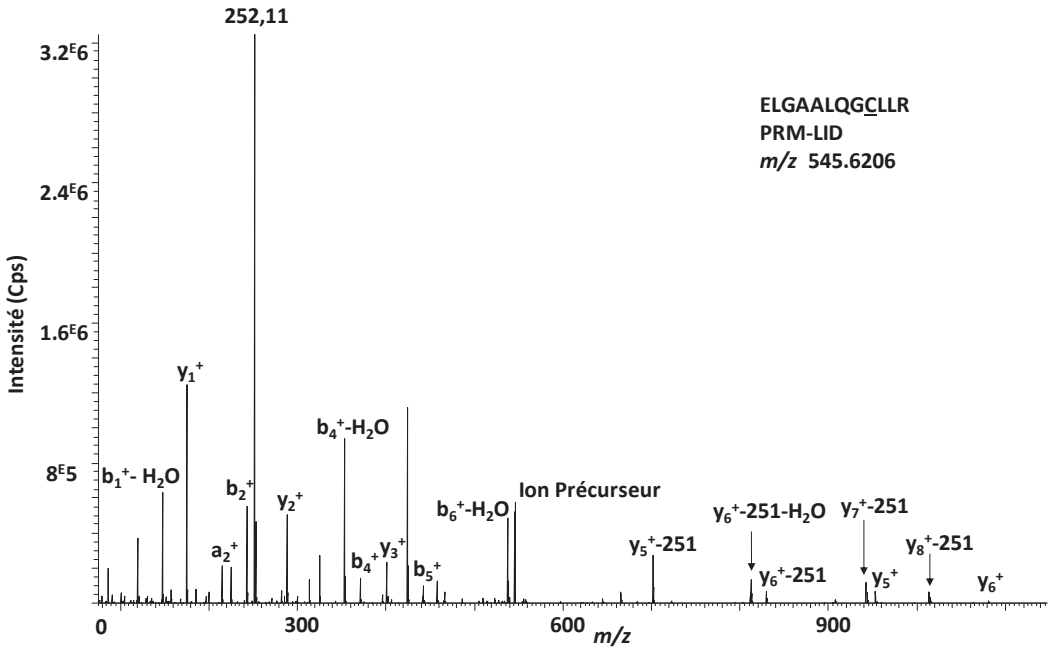


Figure 69 : Spectre de photo-fragmentation à 473 nm du peptide à cystéine dérivé ELGAALQGCLLR

Nous pouvons voir sur le spectre de photo-fragmentation obtenu, comme déjà précisé dans le chapitre II, la production majoritaire de l'ion rapporteur *m/z* 252,11 (produit de la fragmentation interne du chromophore), la fragmentation non totale de l'ion précurseur et une distribution en ion principalement composée d'ions b et y. Cependant, nous pouvons observer deux autres comportements : tout d'abord, la production de fragments avec des pertes d'H_20. C'est un phénomène qui est aussi commun à l'activation par CID/HCD. Ces ions n'ont pu être incorporés à la bibliothèque Skyline induisant une perte d'information à ce niveau. Dans le même temps, nous pouvons également observer la présence d'ions fragments imputés de la valeur 251. Ce sont des ions fragments qui ont subi une fragmentation interne au chromophore (correspondant à l'ion *m/z* 252,11) et une fragmentation de la chaine peptidique. Ces ions n'ont donc plus le chromophore intact et ne pourront pas être inclus sur Skyline (au sein du logiciel cela n'était pas réalisable), générant un manque d'information concernant le processus d'activation par LID à 473 nm en lui-même. Nous n'avons travaillé que sur les ions fragments avec le chromophore intact et sommes conscient qu'une portion non négligeable de l'information n'a pas été prise en compte.

III.II) Types de fragments générés lors de la photo-fragmentation laser de peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm

Pour démarrer notre étude, il a été décidé d'observer au sein des spectres de photo-fragmentation des 401 peptides à cystéine dérivés si des types de fragments particuliers étaient

préférentiellement générés (et par là-même des clivages de liaison). Pour ce faire, tous les types de fragments correspondants aux 6 fragments les plus intenses pour tous les peptides à cystéine dérivés de la bibliothèque ont été répertoriés et une moyenne a été réalisée. Voici les résultats obtenus (Tableau 19) :

Type de fragment	% moyen généré
y	65
b	25
a	4
x	2
z	2
c	2

Tableau 19 : % du type de fragment généré en moyenne pour les 401 peptides à cystéine dérivés de la bilbiothèque de spectres LID à 473 nm

Il peut être observé, comme il avait été explicité lors de l'introduction à la photo-fragmentation des peptides à cystéine via l'activation LID à 473 nm (Chapitre II.II.3), que ce sont majoritairement des ions b et y qui sont produits : environ 90 %. Des ruptures au niveau des liaisons C-N de la séquence peptidique sont donc essentiellement engendrées laissant entrevoir une similarité dans le type d'ion produit avec la CID, la HCD et l'IRMPD. D'autres types d'ions sont également produits (a, c, x, z) mais de manière moindre (inférieur à 5 % pour chacun). Cependant, ce serait à tort de les considérer comme mineurs, car il faut rappeler ici, que nous nous sommes focalisés sur les 6 ions fragments les plus intenses pour chaque peptide à cystéine et que plus de 10 % de ces ions, ne sont ni des ions b ni des ions y. Si la rupture au niveau des liaisons b et y est préférentiellement orientée, elle ne doit pas être assimilée à totale. Néanmoins, c'est principalement ce type d'ion qu'il faut s'attendre à trouver sur un spectre LID à 473 nm et donc cela-même qu'il faut sonder dans un spectre de photo-fragmentation ou (co)-fragmentation pour identifier des peptides à cystéine dérivés. Il en va de même pour le choix des transitions à suivre pour quantifier ces mêmes peptides.

Dans le même temps, nous avons vérifié si cette distribution se conservait quel que soit l'intensité du fragment en vue de plus ou moins généraliser la remarque précédente. Celle-ci se manifeste sur Skyline par l'attribution d'un rang aux fragments en fonction de leurs intensités. La valeur de rang va de 1 à 6 dans notre cas avec une valeur de 1 pour le fragment le plus intense et de 6 pour le 6ème fragment le plus intense pour chaque peptide à cystéine (Tableau 20) :

Rang	1	2	3	4	5	6
y%	66	65	68	66	63	61
b%	25	23	25	24	26	27
a%	2	8	4	4	5	4
x%	3	2	1	3	2	3
z%	1	1	1	1	3	2
c%	2	2	1	2	2	3

Tableau 20 : % du type de fragments généré en fonction du rang du fragment

Nous pouvons voir ici que, quel que soit l'intensité des fragments, la distribution en type de fragment est la même. Nous pouvons donc généraliser la remarque précédente.

En parallèle, nous avons étudié le type de fragment généré en fonction de l'acide aminé au niveau duquel s'effectuait la fragmentation. Nous nous posions la question de savoir si la séquence d'un peptide pouvait influencer les fragments produits et privilégiée un type de fragment. L'acide aminé lieu de la fragmentation a été obtenu de la sorte en fonction du type de fragment produit :



Lors d'une fragmentation, si ce sont des ions x, y ou z qui sont produits alors c'est l'acide aminé vers l'extrémité N-terminale qui est retenu. Si ce sont des ions a, b et c c'est l'inverse. Voici les résultats obtenus (Figure 70) :

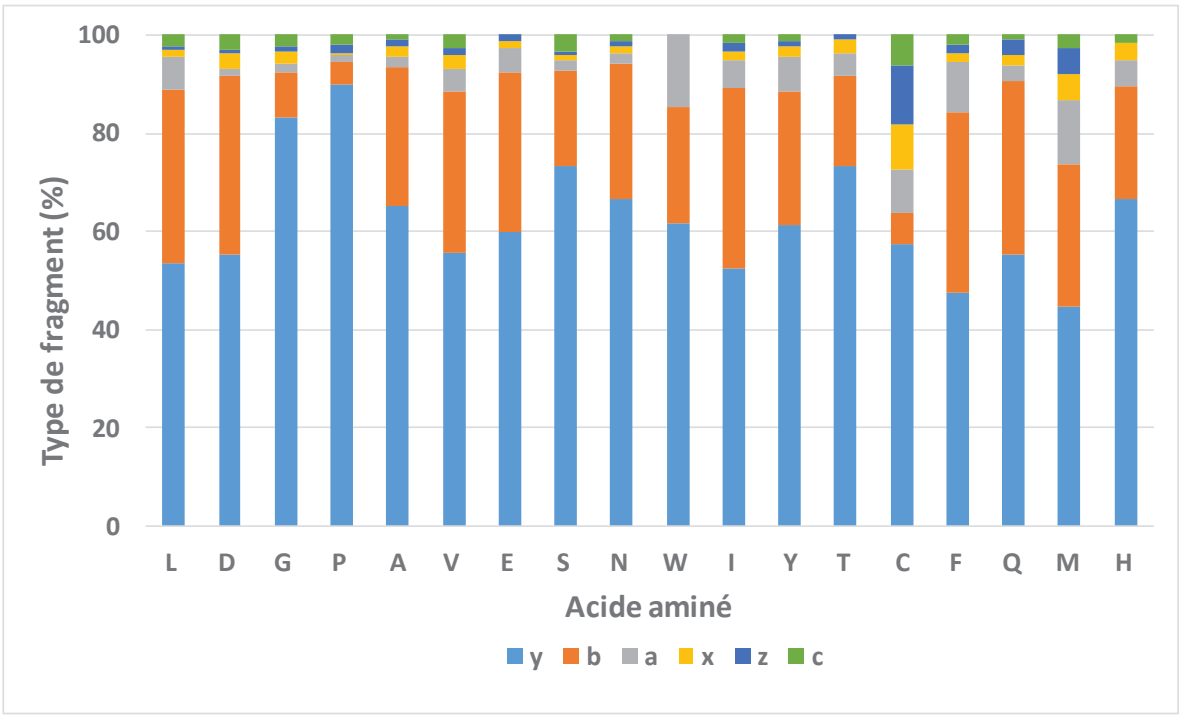


Figure 70 : Types de fragments produits en fonction de l'acide aminé qui est fragmenté

Plusieurs choses sont à observer ici. Tout d'abord, comme attendu et confirmant les observations précédentes, pour chaque acide aminé, lorsqu'il y a fragmentation à son niveau, ce sont des ions b et y qui sont majoritairement formés. Cependant, nous pouvons observer des comportements préférentiels : la fragmentation au niveau des acides aminés proline (P) et glycine (G) semble donner des ions y essentiellement (plus de 90 et 80 % respectivement). En revanche, lors d'une

fragmentation au niveau des acides aminés soufrés, une plus grande richesse de fragments est générée. En effet, 34 % et 26 % des fragments produits au niveau de la cystéine (C) et de la méthionine (M) ne sont pas des fragments b et y. Il est possible que, pour l'acide aminé C, cette diversité plus importante résulte du chromophore fixé sur elle. Un mécanisme de fragmentation radicalaire qui favorise la production d'ions radicaux $z^{\cdot}$ et c peut être supposé. 15 % des fragments produits au niveau du tryptophane (W) sont des ions a et le reste des ions b et des y exclusivement. Seulement 3 types d'ions fragments sont générés à son niveau. Lors d'une fragmentation au niveau des autres acides aminés, ce sont majoritairement des ions y qui sont produits (de 50 % à un peu plus de 70 % selon les différents acides aminés). Cette production est plus ou moins variable en fonction de l'acide aminé en question plutôt que de la famille à laquelle il appartient.

Une fois ces observations faites, nous avons poursuivi notre étude en nous penchant sur l'impact de la nature de l'acide aminé fragmenté.

III.III) Nature de l'acide aminé fragmenté lors de la photo-fragmentation laser de peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm

Dans cette section, nous avons regardé s'il y avait préférentiellement un ou plusieurs acides aminés qui étaient fragmentés. En effet, nous avons déjà vu que les ions b et y étaient en général principalement produits, mais savoir quel ion y ou b préférentiellement sélectionner en fonction de la séquence du peptide à cystéine peut s'avérer réellement utile.

Pour ce faire, le nombre de fragments obtenu pour chaque acide aminé lieu de fragmentation (cf assignement des fragments à l'acide aminé correspondant présenté au-dessus) a été calculé (toute redondance a été exclue : c'est-à-dire deux fragments de charge différente à une même location, par exemple y_5^+ et y_5^{++} pour un même peptide). Ensuite, les valeurs obtenues pour chaque acide aminé ont été pondérées en fonction de leur présence respective au sein de la bibliothèque. Voici les résultats obtenus (Tableau 21) :

Acide aminé	Nombre de fragments générés	Nombre total de l'acide aminé dans la bibliothèque	Ratio (%)	Famille
P	243	319	76	Hydrophobe Aliphatique
W	31	45	69	Aromatique
I	151	272	56	Hydrophobe Aliphatique
V	174	338	51	Hydrophobe Aliphatique
F	89	176	51	Aromatique
A	189	381	50	Hydrophobe Aliphatique
Y	75	152	49	Aromatique
L	266	554	48	Hydrophobe Aliphatique
T	120	255	47	Hydroxyle hydrophile
G	172	372	46	Hydrophobe Aliphatique
D	124	276	45	Diacide
H	51	115	44	Dibasique
S	170	394	43	Hydroxyle hydrophile
Q	94	221	43	Amide
N	77	191	40	Amide
E	171	438	39	Diacide
M	30	96	31	Soufré
C	66	402	16	Soufré

Tableau 21 : % de fragments généré par acide aminé en fonction de leur abondance dans la séquence de tous les peptides à cystéine de la bibliothèque de spectres LID à 473 nm

A travers ce résultat plusieurs observations peuvent-être faites : une localisation de la fragmentation semble être préférentiellement orientée au niveau de deux acides aminés : l'acide aminé P (76 %) et dans une moindre mesure le W (69 %). D'une manière plus générale, il semble que la fragmentation se localise préférentiellement au niveau des acides aminés possédant une chaine latérale de type aromatique ou hydrophobe aliphatique plutôt qu'au niveau des acides aminés appartenant aux familles soufrés et amides, ce qui est aussi observé en HCD[221,222]. Cependant, la plus importante des observations concerne l'acide aminé C : il y a très peu de fragments qui sont produits à son niveau. Elle n'est pas un lieu préférentiel de fragmentation. Cela signifie que malgré l'excitation initiale au niveau du chromophore DABCYL C2-maléimide (qui peut fragmenter de façon interne) il y a également redistribution de l'énergie tout au long de la séquence du peptide, et peu de fragmentation directe au niveau de C dérivée.

D'après ces résultats, nous pouvons dire que la fragmentation se localise principalement au niveau des acides aminés aromatiques et aliphatiques hydrophobes avec une préférence pour P et W. Il semble donc être à cibler préférentiellement.

En raison du processus de redistribution d'énergie supposé nous devions nous assurer que les conclusions précédentes n'avaient pas été influencées par le positionnement de C. Dans cette optique il a été regardé, pour chaque acide aminé, son positionnement moyen relativement à la position de C au sein de la bilbiothèque, dans chacun des 401 peptides à cystéine. Par peptide à cystéine, la distance entre la cystéine et chaque acide aminé a été normalisée par rapport au nombre total d'acides aminés composant la séquence du peptide à cystéine. Par exemple si un peptide à cystéine se composant de 10 acides aminés possède C en position 3 et P en position 7 alors, le positionnement obtenu pour P, dans ce peptide à cystéine, est 0,4 ((7-3)/10 = 0,4). Ceci

a été fait pour tous les acides aminés dans les 401 peptides à cystéine de la bibliothéque de spectres. Une moyenne des distances obtenues a ensuite été effectuée par acide aminé.

Ceci a été réalisé dans l'optique de voir si les acides aminés aromatiques et aliphatiques hydrophobes avaient été fragmentés préférentiellement en raison d'un positionnement différent ou particulier par rapport à l'acide aminé C et inversement pour les autres familles.

Ici est donc présenté le positionnement moyen de chaque acide aminé relativement à celui de C pour les 401 peptides à cystéine de la bilbiothèque (Tableau 22) :

Acide aminé	Position moyenne relativement à la position de la cystéine
F	0,35
P	0,34
G	0,34
V	0,34
D	0,34
L	0,34
E	0,34
I	0,34
T	0,33
Y	0,33
N	0,32
H	0,32
S	0,32
W	0,31
A	0,31
Q	0,31
M	0,30

Tableau 22 : Position moyenne relativement à la position de la cystéine pour chaque acide aminé au sein des 401 peptides à cystéine de la bibliothèque de spectres LID à 473 nm

Nous pouvons observer ici que pour les 17 acides aminés (C, l'arginine (R) et la lysine (K) ne sont pas prises en compte) la position moyenne relativement à la position de la cystéine est de 0,33 avec un minimum à 0,30 pour M, un maximum à 0,35 pour la phénylalanine (F). Nous pouvons conclure à travers ce tableau qu'aucun acide aminé ou famille d'acide aminé n'avait de placement particulier relativement à celui de la cystéine par rapport aux autres. Les résultats n'ont donc pas été biaisés par ce paramètre.

Dans le même temps, nous avons étudié la relation entre l'acide aminé lieu de fragmentation et l'intensité du fragment généré. En effet, pour quantifier des peptides par exemple, il vaut mieux sélectionner les fragments les plus intenses en vue d'atteindre une LOQ la plus basse possible. Pour ce faire, le rang en intensité (parmi les 6 fragments les plus intenses) de chaque fragment correspondant à un type d'acide aminé a été déterminé. Les rangs allaient donc de 1 pour les plus intenses à 6 pour les moins intenses. Le rang moyen a été calculé pour chaque acide aminé (Figure 71) :

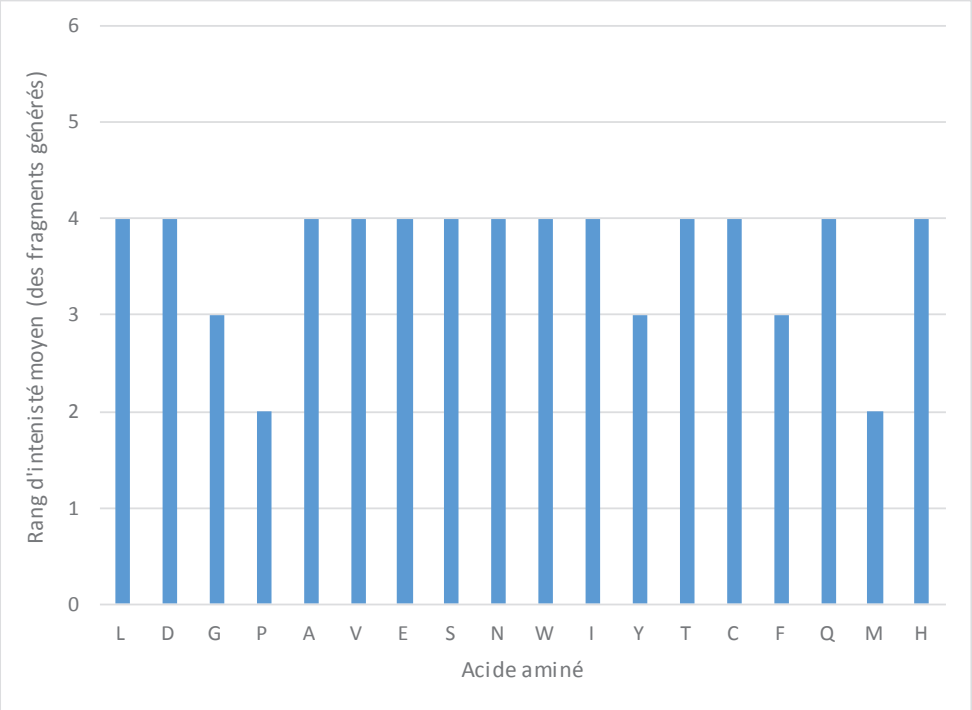


Figure 71 : Rang d'intensité moyen des fragments générés par acide aminé

D'après ces résultats, nous pouvons observer que l'acide aminé P, qui semble être une localisation préférentielle de fragmentation, est aussi une localisation de fragmentation préférentielle de fragments intenses (un rang d'intensité moyen de 2). A l'inverse, bien que M ne soit pas un lieu de fragmentation préférentiel, les fragments qui en résultent sont relativement intenses (un rang d'intensité moyen de 2 également). Pour tous les autres acides aminés, les fragments produits à leurs niveaux ont un rang d'intensité moyen de 3 ou de 4. Il semble donc nécessaire de cibler préférentiellement les acides aminés P et M au sein des séquences des peptides à cystéine dérivés pour obtenir des fragments intenses.

Une fois ces observations faites, nous avons poursuivi notre étude en nous penchant sur le processus de redistribution d'énergie lors de l'activation par LID des peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm.

III.IV) Etude du processus de redistribution d'énergie au niveau de la séquence peptidique lors de la photo-fragmentation laser de peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm

Dans cette optique, une première étude s'est axée sur l'environnement aux alentours du site de fragmentation. Pour la totalité des 6 fragments les plus intenses des 401 peptides à cystéine

dérivés de la bibliothèque de spectres LID à 473 nm la composition en acide aminé à +/- 2 acides aminés entourant le site de fragmentation a été étudiée en vue de visualiser si un schème particulier d'enchainement d'acides aminés ou un acide aminé particulier en lui-même se trouvait présent de manière redondante et pouvait nous renseigner sur le processus de redistribution d'énergie. Aucun enchainement d'acides aminés particulier ou de famille d'acide aminé a pu être détecté. En revanche, il a été vu qu'en moyenne, 30 % des fragments possédaient l'acide aminé cystéine dans une zone de +/- 2 acides aminés relativement à la localisation du site de fragmentation. Cela signifie qu'une notion de distance limite peut intervenir dans le processus de redistribution d'énergie et qu'elle doit être prise en compte et étudiée. En effet, si une distance limite est observée, la première chose à regarder lors de l'analyse d'un peptide à cystéine dérivé avec la méthode d'activation LID à 473 nm serait où se situe la cystéine dans la séquence pour ensuite établir un profil des localisations des sites de fragmentation potentiels dans un intervalle déterminé.

Nous avons donc étudié, pour toutes les séquences des peptides à cystéine dérivés de la bibliothèque, la distance (en acides aminés) entre le site de localisation de la fragmentation et la position de la cystéine. Nous avons décidé de nous limiter au fragment b et y car représentant 90 % des 6 fragments les plus intenses, ils ont été estimés comme représentatifs du comportement global de tous les fragments obtenus.

Pour traiter les valeurs acquises nous avons regroupé et classé les peptides à cystéine dérivés par taille (nombre d'acides aminés) de séquence peptidique (de 8 à 20 acides aminés) et avons moyenné les valeurs obtenues pour chaque taille. Pour chaque taille de peptide, le nombre de fragments détectés a été tracé en fonction de la distance (en acides aminés) entre la cystéine et le site de fragmentation. Voici les résultats obtenus (Figure 72) :

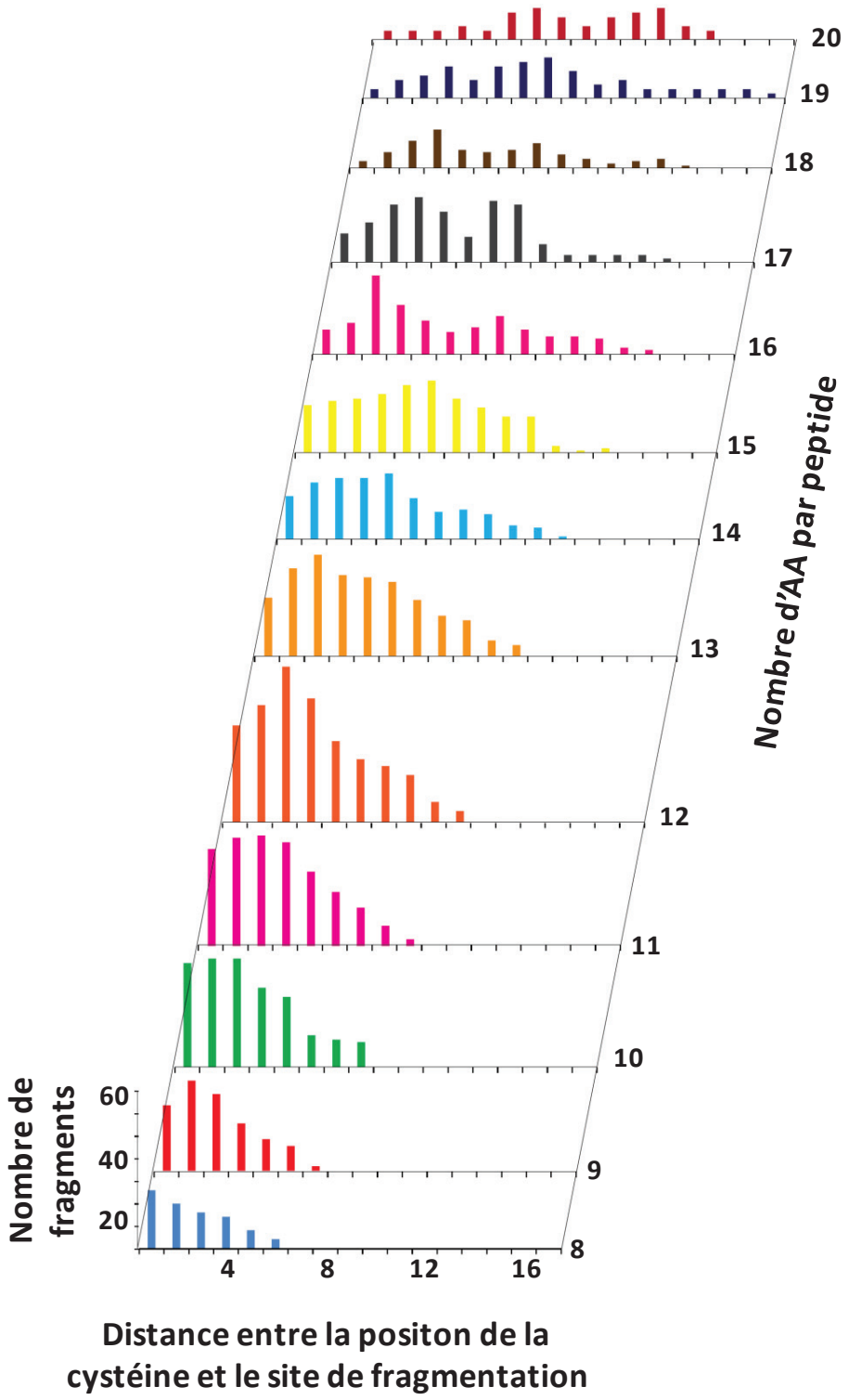


Figure 72 : Nombre de fragments détectés en fonction de la distance entre la cystéine et le site de fragmentation a été tracé pour chaque taille de peptide à cystéine

Tout d'abord, en vue de nous assurer qu'aucun biais n'était observé lors de cette étude, nous avons contrôlé le nombre de peptides à cystéine dérivés qui ont été utilisés pour chaque taille en vue d'être sûr que les comportements visualisés étaient bien issus d'un moyennage important et ne résultaient pas de simplement deux ou trois peptides à cystéine qui, avec une séquence particulière, induiraient ces observations. Voici les résultats obtenus (Tableau 23) :

Taille du peptide à cytéine (nombre d'acides aminés)	Nombre de peptides à cystéine dérivés
8	6
9	26
10	39
11	48
12	56
13	48
14	40
15	38
16	31
17	23
18	18
19	14
20	12
21	1

Tableau 23 : Nombre de peptides à cystéine dérivés issus de la bibliothèque LID à 473 nm utilisés pour chaque taille de séquence peptidique

Nous pouvons visualiser que mise à part la taille de 21 acides aminés qui n'a pas été prise en compte, un nombre de peptides à cystéine dérivés supérieur à 12 pour chaque taille (outre la taille 8 acides aminés mais dont la tendance qui sera bien affirmée, laissera peu de place au doute malgré un nombre de peptides à cystéines utilisé moindre) a été pris en compte. L'approche est estimée valide.

D'après la Figure 72, nous avons pu visualiser que le processus de redistribution d'énergie varie en fonction de la taille des peptides à cystéine dérivés étudiés. La dispersion énergétique n'est pas homogène :

Pour des peptides à cystéine dérivés de taille moindre (entre 8 et 12 acides aminés), nous pouvons observer une redistribution d'énergie ayant lieu majoritairement sur des distances proches relativement à la position de la cystéine. La taille de 8 acides aminés semble vraiment courte et engendre majoritairement une fragmentation sur l'acide aminé situé juste à côté de la cystéine. Pour les tailles 9 et 10, 3 maximas peuvent être considérés comme atteints (distance 1,2 et 3). Pour les tailles 11 et 12, le même modèle est observé mais décalé avec un acide aminé en plus (distance 2,3 et 4). Il semble que pour des peptides à cystéine dérivés de courte séquence, une fragmentation est majoritairement induite à proximité de la cystéine. Pour les peptides à cystéine dérivés très court, une distance de 1 semble préférentielle.

Pour des peptides à cystéine dérivés de taille moyenne (de 13 à 15 acides aminés) nous pouvons commencer à observer une redistribution d'énergie plus importante vers des acides aminés plus éloignés de la cystéine. Les maximas sont atteints à des localisations plus lointaines relativement au positionnement de la cystéine et avec un nombre de fragment moindre : c'est un effet qui semble s'accentuer en fonction de l'augmentation de la taille du peptide à cystéine. Si des fragmentations à des localités proches de la cystéine sont toujours observées nous pouvons visualiser un maxima à une distance de 5 pour un peptide de taille 14 acides aminés ou encore 6 pour un peptide à cystéine de taille 15 acides aminés.

En revanche, pour les peptides à cystéine de taille 16 et 17 acides aminés, deux comportements particuliers sont observés. Tout d'abord les peptides à cystéine de taille 16 acides aminés ne semblent pas suivre la tendance mise en avant précédemment : nous pouvons localiser deux maximas de fragmentation à 3 et 8 relativement à la position de la cystéine alors que nous aurions plutôt attendu un seul maxima aux alentours de 6 et 8. Pour la taille de 17 acides aminés nous pouvons observer 2 maximas aux distances 3-4 et 7-8. Ces deux tailles de peptides à cystéine dérivés ne suivent donc pas la tendance observée précédemment. Cela pourrait s'expliquer par le fait qu'à partir d'une certaine taille de peptides à cystéine dérivés, éventuellement, d'autres mécanismes interviennent pour orienter la fragmentation spécifiquement à des localités particulières relativement à la position de la cystéine ou non. Les peptides à cystéine de taille 18, 19 et 20 acides aminés semblent venir confirmer ces dires, puisque des maximas sont atteints pour des distances plus aléatoires et avec une homogénéisation beaucoup plus prononcée en termes de distances et de nombres de fragments. Ainsi, plus de fragments sont obtenus tout au long de la chaine peptidique, permettant une meilleure identification (c'est-à-dire une plus grande couverture de séquence).

Nous pouvons donc conclure que : pour des peptides à cystéine dérivés courts, la fragmentation intervient à des distances préférentielles comprises entre 1 et 4 acides aminés relativement à la position de la cystéine dérivée. Pour des peptides à cystéine dérivés s'allongeant, nous observons une redistribution d'énergie plus lointaine. Il est possible que d'autres paramètres interviennent à ce niveau, nous ne sommes pas exhaustifs mais cela nous permet néanmoins d'anticiper quelques peu les fragments à obtenir en fonction de la taille des peptides à cystéine ciblés.

III.V) Influence de la position de la cystéine lors de la photo-fragmentation laser de peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm

Outre la taille de la séquence du peptide à cystéine, nous avons décidé de choisir un autre point d'ancrage pour étudier le processus de redistribution d'énergie lors de la photo-fragmentation

des peptides à cystéine dérivés avec le DABCYL C2-maléimide via l'activation par LID à 473 nm. Le positionnement de la cystéine dans la séquence des peptides à cystéine dérivés a été choisi.

Dans cette optique, la position de la cystéine a été normalisée par rapport au nombre total d'acides aminés composant la séquence des peptides à cystéine dérivés. Par exemple si un peptide à cystéine se compose de 10 acides aminés et que sa cystéine se situe en position 3 la valeur normalisée obtenue est 0,3. Cela a été effectué pour tous les peptides à cystéine dérivés de la bibliothèque de spectres LID à 473 nm qui ont ensuite été répartis en quartiles : de 0 à 0,25 la cystéine se situe dans le premier quart de la séquence en partant de l'extrémité N-Terminale (120 peptides), de 0,25 à 0,5 dans le second (107 peptides), de 0,5 à 0,75 dans le troisième (104 peptides) et de 0,75 à 1 dans le dernier (70 peptides). Dans le même temps, nous avons contrôlé la taille moyenne de la séquence des peptides à cystéine dérivés dans chaque quartile afin de s'assurer qu'elle ne puisse pas influencer les résultats. Cela nous a été garanti en raison de l'obtention d'une taille moyenne de 13 acides aminés pour chaque quartile.

Nous avons ensuite regardé la relation entre la position de la cystéine et les types d'ions produits. Dans chaque quartile, les types de fragments produits ont été sondés (Tableau 24) :

0-0,25		0,25-0,50		0,5-0,75		0,75-1	
Type de fragment	% moyen généré	Type de fragment	% moyen généré	Type de fragment	% moyen généré	Type de fragment	% moyen généré
y	89	y	71	y	47	y	42
b	5	b	20	b	40	b	43
a	2	a	3	a	6	a	7
x	1	x	3	x	3	x	3
z	2	z	1	z	2	z	2
c	1	c	2	c	1	c	4

Tableau 24 : Influence de la position de la cystéine sur les types d'ions fragments produits

Nous pouvons tout d'abord observer que lorsque la cystéine se trouve principalement vers l'extrémité N-terminale, des ions fragments y sont majoritairement générés. Puis, plus le positionnement de la cystéine se déplace le long de la séquence peptidique vers l'extrémité C-terminale, plus le pourcentage de fragments y diminue à contrario de celui des ions b qui augmente considérablement. Si les types d'ions fragments produits sont similaires dans chaque quartile (se sont bien des ions b et y qui sont majoritairement produits et donc des liaisons C-N qui sont rompues) leurs pourcentages relatifs, lui, évoluent en relation au positionnement de la cystéine. Il est supposé que la distribution des charges le long de la séquence des peptides à cystéine dérivés puisse être en lien avec cette observation. Nous pouvons cependant conclure que, majoritairement, ce sont des ions y qui sont produits car lorsque la cystéine est placée vers l'extrémité N-terminale (quartile 0-0,25) ils le sont majoritairement (89 %) et lorsque celle-ci est placée vers l'extrémité C-terminale (quartile 0,75-1) le ratio est de 50/50 (42 % de y et 43 % de b). Dans les deux autres quartiles le pourcentage d'ions y est supérieur.

Dans la section III de ce chapitre nous avons vu que, de manière générale, lors d'une fragmentation au niveau des acides aminés P et G des ions y étaient préférentiellement générés

(Figure 70). Il est donc possible qu'une fragmentation au niveau de ces deux acides aminés, même lorsque la cystéine est localisée vers l'extrémité C-terminale, puisse produire préférentiellement des ions y et réhausser les pourcentages relatifs des ions y lors de l'observation précédente.

En vue de répondre à cette interrogation, tous les peptides à cystéine de la bibliothèque de spectres LID à 473 nm ont été répartis en quartiles comme précédemment, puis, pour chaque quartile, il a été regardé tous les types d'ions produits lors d'une fragmentation au niveau des acides aminés P et G. Voici les résultats obtenus (Figure 73) :

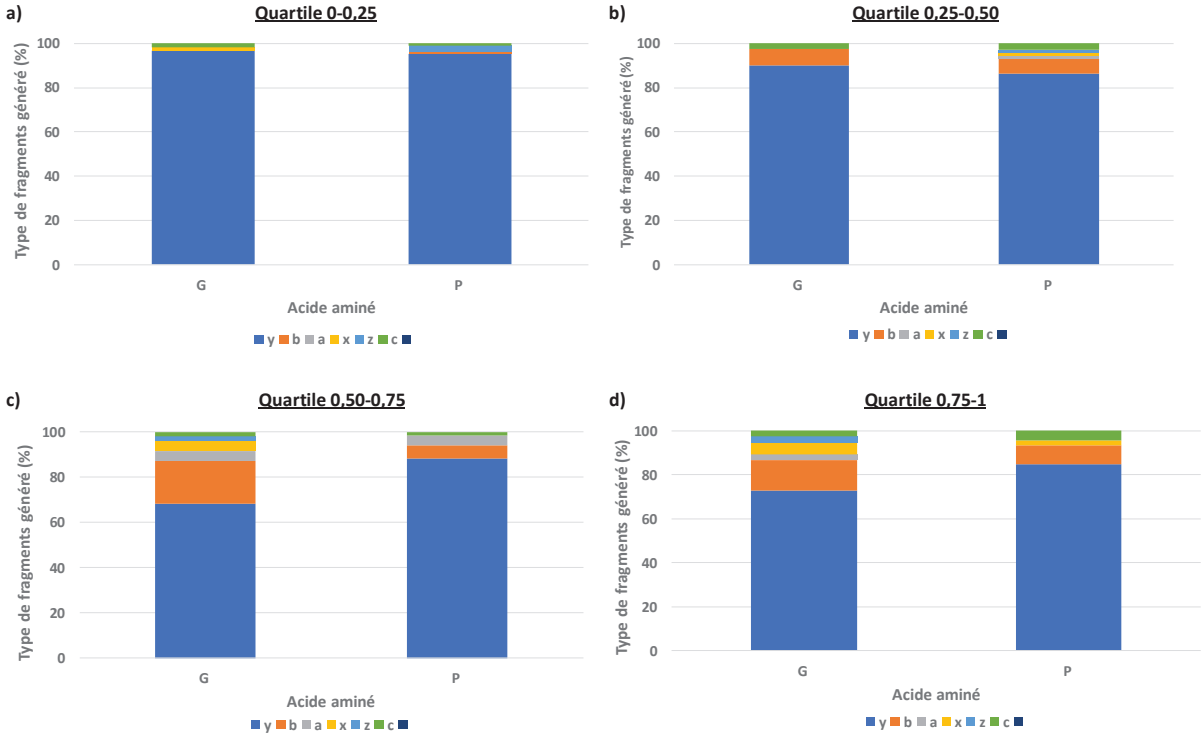


Figure 73 : Proportion du type de fragment généré au niveau des acides aminés P et G relativement à la position de la cystéine

Pour le quartile 0-0,25 (extrémité N-terminale) nous pouvons observer que, lors d'une fragmentation au niveau des acides aminés P et G, des ions y sont majoritairement produits (plus de 95 %) (Figure 73 a). Pour le quartile 0,25-0,5 un léger recul du pourcentage des fragments y est visualisé (86 % et 90 % pour P et G respectivement) (Figure 73 b). Cela est haut-delà du pourcentage moyen obtenu au niveau de tous les acides aminés pour ce quartile (Tableau 24, 71 %). Pour le quartile 0,5-0,75, le pourcentage d'ion y résultant d'une fragmentation au niveau de l'acide aminé P est toujours élevé mais, pour G, ce pourcentage diminue (88 % et 68 % respectivement) avec un pourcentage des ions b s'accentuant et une plus grande diversité de fragments produits (Figure 73 c). Néanmoins, il faut repréciser que dans ce quartile, le pourcentage moyen d'ions y obtenu est de 47 % (Tableau 24) : le pourcentage obtenu au niveau

de G est fortement au-dessus de cette valeur (+ 21 %). Concernant le dernier quartile 0,75-1, la fragmentation au niveau des acides aminés P et G engendre toujours majoritairement des ions y (plus de 85 et 72 % respectivement) (Figure 73 d). Dans ce quartile, le pourcentage moyen d'ions y obtenu au niveau de tous les acides aminés est de 42 % (Tableau 24) : les pourcentages obtenus pour G et P sont bien au-dessus de cette valeur (+ 43 % et + 30 % respectivement) : il semble donc qu'une fragmentation au niveau de ces deux acides aminés produira majoritairement des ions y (surtout pour P) quel que soit leur position relativement à celle de la cystéine.

En vue de s'assurer que la fragmentation au niveau de ces deux acides aminés a bien contribué à maintenir le taux de fragments y élevé, nous avons regardé la part de fragment générée au niveau de ces 2 acides aminés par rapport à la totalité des fragments produits au niveau de tous les acides aminés et la part d'ion y générée au niveau de ces 2 acides aminés par rapport à la totalité des ions y produits au niveau de tous les acides aminés pour chaque quartile. Nous souhaitions voir si le pourcentage relatif d'ions y obtenu au sein de chaque quartile (Tableau 24) n'était pas influencé par la fragmentation au niveau de ces 2 acides aminés (Equation 17 et 18 et Tableau 25) :

$$\%\ \text{fragments (P et G)} = \frac{Nombre\ de\ fragment\ généré\ au\ niveau\ de\ P\ et\ G}{Nombre\ total\ de\ fragments\ \ générés\ dans\ le\ quartile} \times 100 \qquad \textbf{Equation 17}$$

$$\%\ \text{ions y (P et G)} = \frac{Nombre\ de\ fragment\ y\ généré\ au\ niveau\ de\ P\ et\ G}{Nombre\ total\ de\ fragments\ y\ générés\ dans\ le\ quartile} \times 100 \qquad \textbf{Equation 18}$$

Quartile	0-0,25	0,25-0,5	0,5-0,75	0,75-1
% de fragments (P et G)	24	16	16	18
% d'ions y (P et G)	26	21	27	34

Tableau 25 : % de fragments et d'ions y générés au niveau de P et G par rapport à la totalité des fragments et ions y générés au niveau de tous les acides aminés pour chaque quartile

Nous pouvons voir tout d'abord que pour le premier quartile, plus de 24 % des fragments sont générés au niveau de P et de G et que le % d'ions y résultant est quasiment identique (26 %). Sans compter ces 2 acides aminés le pourcentage moyen d'ions y généré serait de 88 % au lieu des 89 % obtenus (Tableau 24). La fragmentation au niveau des acides aminés P et G n'a pas infléchie les résultats.

Pour le quartile 0,25-0,5, 16 % des fragments sont générés au niveau de P et de G et le % d'ions y résultant est de 21 %. Cette fois-ci, nous pouvons observer que la part d'ions y issue de la fragmentation au niveau de P et de G est plus importante que la part de fragments issue de la fragmentation au niveau de P et G. Cela signifie que, même si cette influence est encore mineure dans ce quartile (sans compter ces 2 acides aminés le pourcentage moyen d'ions y généré serait

de 70 % au lieu de 71), il peut s'agir du début d'une tendance à favoriser la production d'ions y par le biais de la fragmentation au niveau de ces 2 acides aminés et à infléchir le % du type de fragment observé.

Pour le quartile 0,5-0,75 la tendance précédente supposée est confirmée, la fragmentation au niveau de P et G représente 16 % des fragments produits mais 27 % des ions y produits sont issus d'une fragmentation au niveau de ces 2 acides aminés : sans incorporer les résultats pour ces 2 acides aminés le pourcentage moyen d'ions y généré serait de 41 % au lieu de 47 et ceux des ions b de 46 % au lieu de 40 %. Les ions b deviendraient les ions majoritairement produits pour ce quartile alors qu'ils ne le sont pas, principalement, en raison des pourcentages issus de la fragmentation au niveau de P et G. Et il en va de même et de manière plus importante pour le dernier quartile (0,75-1) où la fragmentation au niveau de P et G représente 18 % des fragments produits mais où 34 % des ions y produits sont issus d'une fragmentation au niveau de ces 2 acides aminés soit plus d'un tiers. Sans compter la fragmentation au niveau de ces 2 acides aminés le pourcentage moyen d'ions y généré serait de 33 % au lieu de 42 et ceux des ions b de 51 % au lieu de 43.

Nous pouvons donc conclure après ces observations que la fragmentation au niveau des acides aminés P et G, qu'importe le positionnement de la cystéine, engendrera principalement des ions fragments y. Cela a pu influencé les pourcentages moyens des ions b et y observés précédemment. Ci-dessous est de nouveau présenté l'influence de la position de la cystéine sur le type d'ion produit mais sans avoir incorporer les acides aminés P et G dans les calculs (Tableau 26) :

0-0,25		0,25-0,50		0,5-0,75		0,75-1	
Type de fragment	% moyen généré	Type de fragment	% moyen généré	Type de fragment	% moyen généré	Type de fragment	% moyen généré
y	88	y	70	y	41	y	33
b	5	b	21	b	46	b	51
a	2	a	4	a	6	a	8
x	1	x	3	x	3	x	2
z	2	z	1	z	2	z	2
c	1	c	2	c	1	c	4

Tableau 26 : Influence de la position de la cystéine sur le type d'ion fragment produit sans compter les acides aminés P et G

Nous pouvons aboutir aux mêmes conclusions que précédemment si ce n'est qu'à partir du quartile 0,5-0,75 le type de fragment majoritairement généré est l'ion b (il est important de repréciser que le type d'ion majoritairement produit reste, de manière générale, les ions y avec un % minimum de 33 % pour le quartile 0,75-1) au niveau de tous les acides aminés mis à part pour les acides aminés P et G où des ions y sont préférentiellement générés.

Conclusion

A travers ce chapitre nous avons pu exposer une première investigation concernant le modèle de fragmentation des peptides à cystéine dérivés avec le DABCYL C2-maléimide par LID à 473 nm. Nous avons pu voir qu'un spectre de photo-fragmentation de peptide à cystéine dérivé se compose principalement d'un ion rapporteur à 252,11 Th issu d'une fragmentation au niveau du chromophore et d'une distribution en ions fragments principalement composée d'ions b et y issus de la fragmentation de la chaine peptidique (liaison C-N). Il a été montré que très peu de fragments étaient générés au niveau de l'acide aminé cystéine impliquant un processus de redistribution d'énergie. Ce processus de redistribution s'est révélé avoir un comportement différent selon la taille de la séquence des peptides à cystéine dérivés photo-fragmentés. Il semble donc que l'énergie se redistribuera différemment selon que le peptide à cystéine dérivé soit long ou court avec des localisations préférentielles des sites de fragmentation différentes. Il a également été montré que la position de la cystéine au sein de la séquence des peptides à cystéine dérivés avait une influence sur le pourcentage relatif des ions b et y obtenus (excepté pour la proline et la glycine générant préférentiellement en permanence des ions y). Il est supposé que la distribution des charges le long de la séquence des peptides à cystéine dérivés puisse être en lien avec cette observation. Pour terminer, un tableau récapitulatif est présenté rassemblant toutes les informations qui ont été observées et pouvant servir à anticiper le spectre de photo-fragmentation qui sera obtenu en fonction du peptide à cystéine dérivé photo-fragmenté et donc, à aider à prédire les fragments principalement obtenus en vue d'une identification ou quantification par spectrométrie de masse (Tableau 27).

Il est important de repréciser ici que cette investigation n'est qu'une première approche et qu'une seconde doit être impérativement mise en place en vue d'éclaircir les mécanismes impliqués dans ce modèle de fragmentation. Premièrement, il serait nécessaire d'incorporer les fragments avec le chromophore non intact (imputés de la valeur de 251) et de travailler avec des spectres de photo-fragmentation de meilleure qualité dans le but de travailler avec plus de 6 fragments. Ensuite, il serait intéressant de continuer à approfondir la relation entre la position de la cystéine et la distribution des ions y et b observée. Dans cette optique, par exemple, une étude devrait se pencher sur la localisation des charges situées sur la séquence des peptides à cystéine dérivés et son impact sur la redistribution de l'énergie (commencer peut-être par regarder la relation entre la distance entre la cystéine dérivée et l'extrémité N-terminale chargé positivement). Dans le même temps, il serait également important de s'assurer que la normalisation effectuée dans ce chapitre et la répartition des 401 peptides à cystéine en 4 quartiles étaient judicieux et n'ont pas biaisées les résultats (une répartition plus fine est-elle nécessaire (en déciles par exemple) ? Est-ce que la manière de normaliser était trop grossière (il nous semble que nous ne devons normaliser que par taille de peptides à cystéine car une valeur de 0,3 pour la localisation d'un site

de fragmentation n'est pas absolument équivalente en termes de nombre d'acides aminés ou de distance entre un peptide se composant de 8 ou 20 acides aminés. Cela peut infléchir les résultats. Cette manière de normaliser a été effectuée en raison du faible nombre de peptides à cystéine disponibles par taille). En parallèle de ces travaux, des contacts doivent être pris avec des théoriciens pouvant expliquer par le calcul ce que nous observons expérimentalement afin de valider les conclusions. Parmi ces études théoriques, une devrait particulièrement s'attarder sur les mécanismes de fragmentation mis en jeu spécifiquement au niveau de l'acide aminé cystéine pour valider ou non le mécanisme de fragmentation radicalaire supposé.

Beaucoup de travail reste donc encore à faire relativement à cette étude mais les premiers résultats soulèvent beaucoup d'interrogations scientifiquement attrayantes ne demandant qu'à être levées.

Dans le chapitre suivant, en vue de poursuivre l'exploration des voies qui permettront d'atteindre l'exhaustivité et la confiance requise lors d'analyses non ciblées pour identifier de nouveaux biomarqueurs protéines à l'aide de la méthode d'activation LID à 473 nm, une nouvelle méthodologie intitulée C-Trap-LID et son potentiel dans cette optique seront présentés.

Poisition de la cystéine normalisée	Proportion d'ions y générée (%) sans P et G	Proportion d'ions b générée (%) sans P et G
0-0,25	88	5
0,25-0,5	70	21
0,5-0,75	41	46
0,75-1	33	51
Taille du peptide (Nombre d'acide aminés)	**Fragmentation**	
8 à 12	Préférentielle à proximité de l'acide aminé cystéine (1 à 4 acides aminés)	
13 à 15	Préférentielle éloigné de l'acide aminé cystéine (5/6 acides aminés)	
16 et >	Homogénéisation : distribution aléatoire tout au long de la séquence	
Supplément	**Observations**	
1	L'ion majoritairement produit est l'ion rapporteur m/z 252,11 Th issu d'une fragmentation au niveau du chromophore	
2	Peu de fragmentation induite au niveau de l'acide aminé C : processus de redistribution d'énergie	
3	Fragmentation préférentielle au niveau des acides aminés issus des familles aromatique et aliphatique hydrophobe contrairement à celle des soufrés	
4	La fragmentation au niveau de P et M induit les fragments les plus intenses	
5	La fragmentation au niveau des acides aminés P et G génère préférentiellement des ions y quelque soit la position de la cystéine	

Tableau 27 : Tableau récapitulatif des observations relatives au modèle de fragmentation des peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm

Chapitre IV : Mise en place d'une nouvelle méthodologie intitulée C-trap-LID

Introduction

Dans ce chapitre il sera introduit une nouvelle méthodologie intitulée C-Trap-LID. Ce procédé développé au sein du laboratoire repose sur la photo-fragmentation des ions non plus dans la cellule HCD mais directement au sein de la C-trap du Qexactive.

Premièrement, nous décrirons ce nouveau concept en détail : la réalité de la photo-fragmentation dans la C-trap sera démontrée et un point sera fait sur les avantages que peut offrir cette méthodologie.

Deuxièmement, nous nous attarderons sur les aménagements expérimentaux qui furent nécessaire pour permettre à la C-trap-LID d'être utilisée en routine au laboratoire : la mise en place d'un obturateur à l'entrée de la cellule HCD synchronisé avec un signal déclencheur provenant du Qexactive sera explicitée.

Nous terminerons ce chapitre en démontrant l'efficacité de la méthodologie C-trap-LID à 473 nm pour photo-fragmenter des peptides à cystéine dérivés : un point sera fait sur l'efficacité de l'étape de photo-fragmentation des peptides à cystéine dérivés au sein de la C-trap (taux de photo-fragmentation, richesse des spectres C-trap LID obtenus) puis, la C-trap-LID à 473 nm sera utilisée et son potentiel pour repérer et extraire toutes les masses des ions précurseurs détectables par l'appareillage ayant photo-fragmentés dans une matrice complexe sera démontré.

IV.I) Matériels et Méthodes

Cette section explicite le matériel et les méthodes utilisés tout au long de ce chapitre. Il sera d'abord présenté les échantillons qui ont été retenus en vue des différentes expérimentations, le chromophore utilisé, la préparation d'échantillons et les informations relatives au couplage LC/MS/MS utilisé et au traitement de données.

IV.I.1) Sélection des échantillons

- 10 peptides à cystéine synthétisés chez Thermo Scientific (Peptide grade 2). Ils ont été reçus déjà solubilisés dans un milieu H_2O/ACN + 0,5 % FA.

- 3 protéines conservées sous forme de poudre (Sigma Aldrich - 97 % pour les 3) : Albumine humaine, Lysozyme de poulet et Serotransferrine humaine.

IV.I.2) Chromophore utilisé

Le chromophore DABCYL C2-maléimide a été conservé pour cette étude (voir le chapitre II.1.2 pour plus de détails)

IV.I.3) Préparation d'échantillons

Le même protocole de dérivation et de préparation d'échantillons avant analyse LC/MS/MS que dans le chapitre II.I.3 a été utilisé pour les échantillons suivants :

- 2,5 µL d'une solution mère de 10 peptides à cystéine en matrice aqueuse solubilisés dans un milieu H_2O/ ACN (50/50) + 0,5 % FA (1 µg pour chaque peptide à cystéine)

- 2,5 µL d'une solution mère de 10 peptides à cystéine en matrice aqueuse solubilisés dans un mélange de 3 protéines (1 µg pour chaque peptide à cystéine et 100 µg de chaque protéine) dans un milieu H_2O/ ACN (50/50) + 0,5 % FA.

Les quantités injectées sur le système chromatographique sont les suivantes :

- 200 ng pour chaque peptide à cystéine analysé soit 2 µg pour les 10 (Vrep = 100 µL)

- 62 µg au total : 20 µg pour chaque protéine et 2 µg pour les 10 peptides à cystéine (Vrep = 100 µL)

IV.I.4) Couplage LC/MS/MS

C'est le même matériel LC/MS/MS que dans le chapitre II qui a été utilisé lors des analyses effectuées pour ce chapitre. En revanche, les gradients chromatographiques utilisés sont différents. Ils sont présentés ci-dessous et affiliés à un numéro qui sera réutilisé au long du chapitre (Tableau 28 et 29) :

Gradient 3 (Tableau 28) :

Temps (min)	% Solvant B (ACN + 0,5 % FA)
0	15
1,5	15
23	75
23,1	95
26	95
26,1	15
30	15

Tableau 28 : Variation de la composition en solvant B de la phase mobile en fonction du temps lors du gradient 3

Gradient 4 (Tableau 29) :

Temps (min)	% Solvant B (ACN + 0,5 % FA)
0	15
4	15
41	50
41,1	95
45	95
45,1	15
50	15

Tableau 29 : Variation de la composition en solvant B de la phase mobile en fonction du temps lors du gradient 4

Concernant le paramétrage du spectromètre de masse il sera explicité en détail plus loin dans ce chapitre.

IV.I.5) Logiciels

Lors de ce chapitre les données brutes auront été visualisées avec le logiciel Xcalibur 2.2.48 (Thermo Scientific). Toutes les données auront été traitées à l'aide du logiciel Excel.

IV.II) Constatation de la possibilité d'effectuer la photo-fragmentation laser au sein de la C-Trap (C-trap-LID)

Durant le courant de ma thèse, lors de la vérification journalière du bon alignement du laser au sein de la cellule HCD, j'ai pu observer qu'une étape de photo-fragmentation pouvait s'opérer directement au sein de la C-Trap du Qexactive. En effet, un alignement particulier du laser via un jeu de lentilles permet d'atteindre cette zone et ainsi induire une photo-fragmentation directement au sein de la C-trap (Figure 74) :

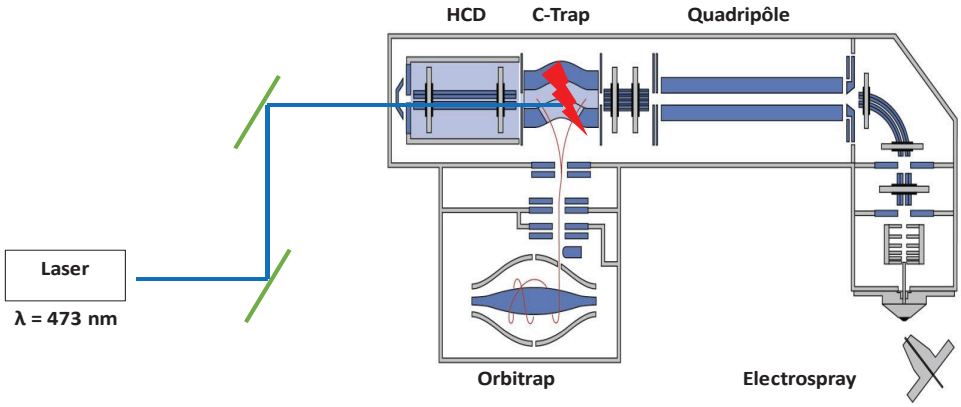


Figure 74 : Schéma du Qexactive avec un alignement particulier du laser permettant une photo-fragmentation au sein de la C-trap

Il est important de rappeler ici que, lors d'une analyse FMS classique, les ions ayant été transmis par le quadrupole seront accumulés au sein de la C-Trap puis injectés par paquets dans l'orbitrap. Ne passant par la cellule HCD, ils ne sont pas fragmentés. Dans le cas présent, avec l'alignement particulier du laser, nous pouvons faire photo-fragmenter les ions directement dans la C-trap en mode FMS sans les faire passer par la cellule HCD.

Voici le spectre du DABCYL-C2 maléimide solubilisé dans du méthanol à une concentration d'environ 50 µg/mL puis infusé seul au sein du spectromètre à un débit de 10 µL/min en mode FMS (Figure 75) :

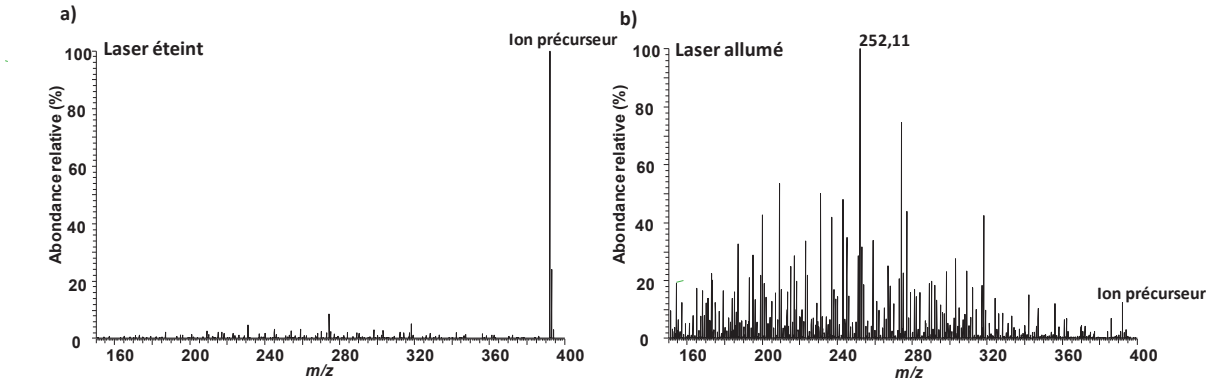


Figure 75 : Spectre FMS obtenu pour le DABCYL-C2 maléimide avec un alignement particulier du laser permettant une photo-fragmentation au sein de la C-trap lorsque le laser est allumé (a) ou éteint (b)

Nous pouvons observer que lorsque le laser est éteint, nous obtenons l'information concernant l'ion précurseur à *m/z* 392,17 (Figure 75 a), en revanche, quand le laser est allumé, nous observons la photo-fragmentation de l'ion précurseur bien qu'il ne soit pas passé dans la cellule HCD (Figure 75 b). Ici, l'ion principal observé est l'ion rapporteur caractéristique du DABCYL C2-maléimide (*m/z* 252,11), l'ion précurseur, quant à lui, est bien présent mais à une intensité

relative de 20 % par rapport à l'ion rapporteur : comme lors de la photo-fragmentation du chromophore dans la cellule HCD. Il y a donc bien photo-fragmentation des ions dans la C-trap. Ce concept sera intitulé C-Trap-LID. Si ce procédé fonctionne pour le DABCYL C2-maléimide il peut surement aussi fonctionner pour des peptides à cystéine dérivés. La question est de savoir si cette opération pourrait avoir un sens lors de l'analyse de peptides à cystéine dérivés ?

IV.III) Apport de la méthodologie C-trap-LID lors de l'analyse de peptides à cystéine dérivés

Pour prendre conscience d'un éventuel apport de cette méthodologie lors de l'analyse de peptides à cystéine dérivés il faut bien saisir ce qu'elle implique.

Si une étape de photo-fragmentation peut s'effectuer au sein de la C-Trap cela signifie qu'il n'y a plus besoin d'envoyer les ions dans la cellule HCD. De ce fait, lors d'une analyse C-Trap-LID, tous les ions analysés effectuerons le trajet suivant, c'est-à-dire celui d'une acquisition FMS classique :

Source d'ionisation → Q1 → C-trap → Orbitrap

Pour rappel, lors d'une acquisition classique MS/MS en LID ou HCD (PRM, DDA, DIA), les ions analysés effectuent ce trajet :

Source d'ionisation → Q1 → C-trap → Cellule HCD → C-trap → Orbitrap

Par conséquent, lors d'une acquisition C-Trap-LID, le temps correspondant à la fragmentation (25 ms) n'est plus à additionner à l'ITmax vu que les ions ne sont pas transmis dans la cellule HCD : il est compris dedans.

Il est important ici de rappeler un point important concernant le fonctionnement du Qexactive :

Nous devons nous souvenir que, lors d'une étape d'accumulation des ions dans la C-trap, si l'AGC Target est atteint avant la fin de l'ITmax alors la collecte sera stoppée et l'ITexp sera plus court que l'ITmax. En revanche, si l'AGC Target n'est pas atteint au bout de l'ITmax, la collecte est aussi stoppée mais cette fois-ci la quantité d'ions accumulée sera moindre que la cible fixée. En C-trap-LID, en raison de l'étape de photo-fragmentation au sein de la C-trap, nous devons tenir compte des ITexp obtenus et du temps d'activation nécessaire pour obtenir une photo-fragmentation à 473 nm avec un taux optimal pour des peptides à cystéine dérivés avec le DABCYL C2-maléimide (25 ms).

Tout d'abord, nous tenons à préciser qu'il a été démontré par le passé qu'un temps d'irradiation plus long (supérieur à 25 ms) ne modifie pas ou ne détériore pas les spectres de photo-

fragmentation LID à 473 nm obtenus. L'allongement temporel du processus de photo-fragmentation n'est pas un problème[9]. De ce fait, fixé et obtenir des IT (max et exp) supérieurs à 25 ms lors d'une analyse C-Trap-LID à 473 nm ne nuira pas à la qualité des spectres de photo-fragmentation obtenus. Cependant, dans le cas où, la quantité d'ions maximale à collectée sera atteinte rapidement (au bout de quelques ms), l'ITexp sera inférieur à l'ITmax et le temps d'activation des ions sera inférieur à 25 ms (temps où l'on sait que le taux de photo-fragmentation n'est pas optimal). La question qui se pose alors est, dans ce cas, de savoir où sont stockés les ions en attendant la fin du temps d'analyse de l'orbitrap ? En effet, il faut rappeler que l'orbitrap travaille en parallèle, de ce fait, lorsque l'orbitrap analyse des ions en vue d'un scan, les ions du prochain scan sont déjà en train d'être accumulés dans la C-trap. Pour exemple, si, lors d'une analyse, la résolution fixée de l'orbitrap est de 35 000 (temps d'analyse de 128 ms) et que pour 1 scan l'AGC Target est atteint en 10 ms, la question est : où sont stockés les ions avant d'être envoyés dans l'orbitrap ? S'ils sont stockés dans la C-trap cela ne pose aucun problème pour notre approche, s'ils ne le sont pas cela le devient. D'après Thermo Scientific les ions restent stockés dans la C-trap mais l'accumulation est stoppée. Les ions passeront donc plus de 25 ms dans la C-trap et leur taux de photo-fragmentation devrait être optimal. Cela n'est donc pas un problème.

Concernant l'apport de la méthodologie C-Trap-LID à 473 nm lors de l'analyse de peptides à cystéine dérivés, en raison de la photo-fragmentation dans la C-trap, le temps de 25 ms minimal nécessaire à cette étape (classiquement effectuée dans la cellule HCD) peut être rajouté à l'ITmax en vue d'optimiser le temps de cycle. De ce fait, lors des scans où l'ITmax est atteint nous pouvons gagner 25 ms en plus pour accumuler plus d'ions et gagner en sensibilité. Lorsque l'ITmax n'est pas atteint, en raison d'une quantité d'ions maximale collectée plus rapidement, il n'y a aucun gain en sensibilité. Cependant, une question demeure en suspens concernant la différence entre le taux de photo-fragmentation des premiers ions accumulés dans la C-trap et celui des derniers lorsque l'ITmax est atteint ? En effet, les derniers ions entrant dans la C-trap (dernières 25 ms d'accumulation) vont passer moins de temps dans celle-ci lorsque l'ITmax est atteint et, par conséquent, leurs taux de photo-fragmentation devraient être non optimaux (puisque leurs temps d'activation seront inférieurs à 25 ms) ? La question est donc de savoir si, lorsque l'ITmax est atteint, les ions sont directement transmis dans l'orbitrap ou s'ils sont encore stockés quelques ms dans la C-trap avant leur transmission. D'après Thermo Scientific les ions sont directement transmis. Le taux de photo-fragmentation des derniers ions accumulés (les dernières 25 ms) devrait donc théoriquement être non optimal lors des scans où l'ITmax est atteint.

Ils semblent donc, que le seul apport de la méthodologie C-trap-LID à 473 nm, comparativement à une approche classique, lors d'une l'analyse de peptides à cystéine dérivés soit un gain en sensibilité via un ITmax supérieur de 25 ms a contrario d'un taux de photo-fragmentation moins optimal (seulement pour les scans où l'ITmax est atteint). Il est donc délicat d'affirmer ou non dès

maintenant si cet effet pourra s'avérer, dans certains cas, déterminant. D'un autre côté, il est toujours plaisant de développer un nouveau concept.

IV.IV) Mise en place de la C-trap-LID

IV.IV.1) Installation d'un obturateur

Pour pouvoir concrétiser cette approche, il a d'abord fallu contrôler l'étape de photo-fragmentation. En effet, lorsque le laser est aligné pour faire photo-fragmenter les ions au sein de la C-trap, l'étape de photo-fragmentation est induite pour chaque scan contrairement aux approches précédentes. Les ions entrant dans la C-Trap sont donc perpétuellement photo-fragmentés et nous n'obtenons que des spectres C-trap-LID en permanence : pas de FMS. Cela implique une perte d'information concernant tous les ions précurseurs analysés et donc une limitation quant aux modes pouvant être utilisés avec cette méthodologie : si en DIA l'information concernant les ions précurseurs n'est pas vitale mais peut être désirable, elle l'est pour les modes DDA, PRM et AIF.

En vue de multiplier les modes potentiellement utilisables avec cette approche le problème a été résolu de la sorte (Figure 76) :

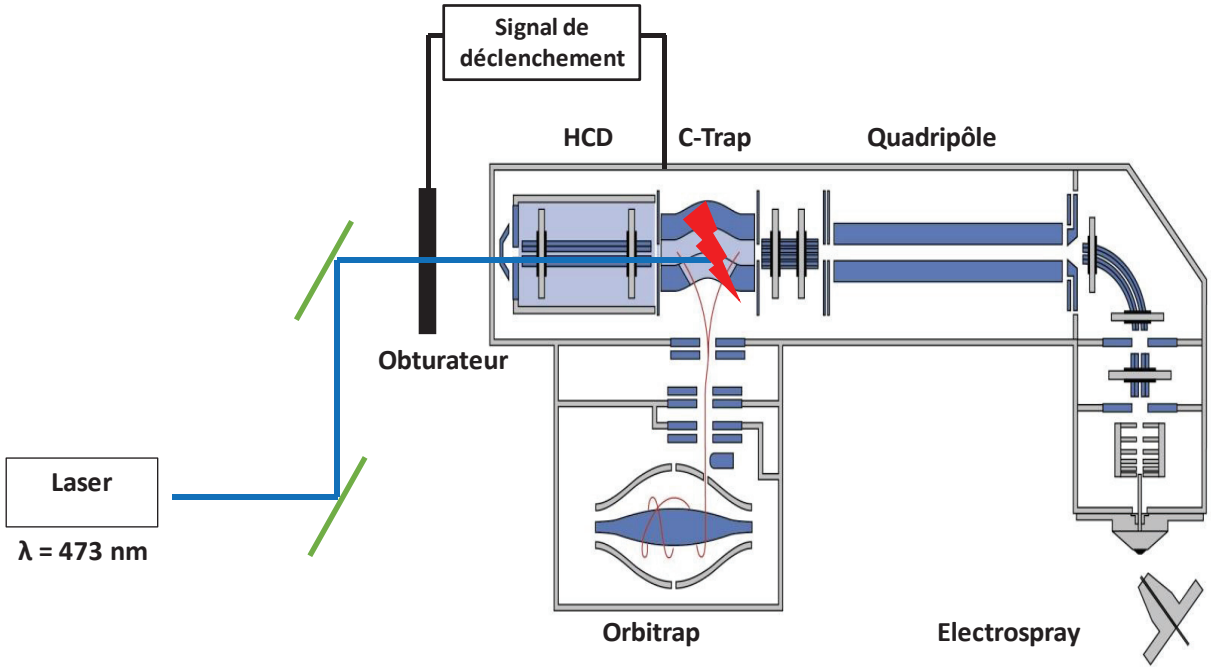


Figure 76 : Schéma du Qexactive avec un alignement particulier du laser permettant la C-trap-LID et un obturateur

Il a été décidé de disposer un obturateur à l'entrée de la cellule HCD et de le synchroniser avec le Qexactive pour que celui-ci s'ouvre (fragmentation) et se ferme (pas de fragmentation) en

fonction des scans désirés (C-trap-LID et FMS). L'ouverture et la fermeture de l'obturateur sont contrôlées par un arduino (carte matériellement libre disposant d'un microcontrôleur) et son logiciel de programmation qui reçoit « un trigger » (signal déclencheur provenant du spectromètre). Au départ, dans le but d'une approche la plus simplifiée possible, nous souhaitions obtenir le signal correspondant à l'entrée des ions dans la C-Trap. En effet, en faisant de ce signal « le » signal déclencheur du spectromètre de masse, nous pourrions ouvrir ou fermé l'obturateur à chaque fois que les ions entrent dans la C-trap.

Il est à préciser ici que Thermo Scientific n'a aucunement souhaité nous aider pour développer ce projet. De ce fait, peu de signaux sortant du Qexactive ont pu être rendus accessibles, compris et utilisables dans cette optique. Cependant, à l'aide d'un oscilloscope, en recherchant des signaux empiriques à des points tests, nous avons pu visualiser 2 signaux radiofréquence différents et inconnus semblant être rattachés au mouvement des ions dans l'appareil. Voici une représentation des 2 signaux obtenus sur l'oscilloscope lorsque le Qexactive est allumé avec le paramétrage suivant (Tableau 30 et Figure 77) :

Résolution	140 000
AGC Target	1,0E+06
IT max (ms)	250
Plage de *m/z* (Th)	150-1200

Tableau 30 : Paramétrage du Qexactive

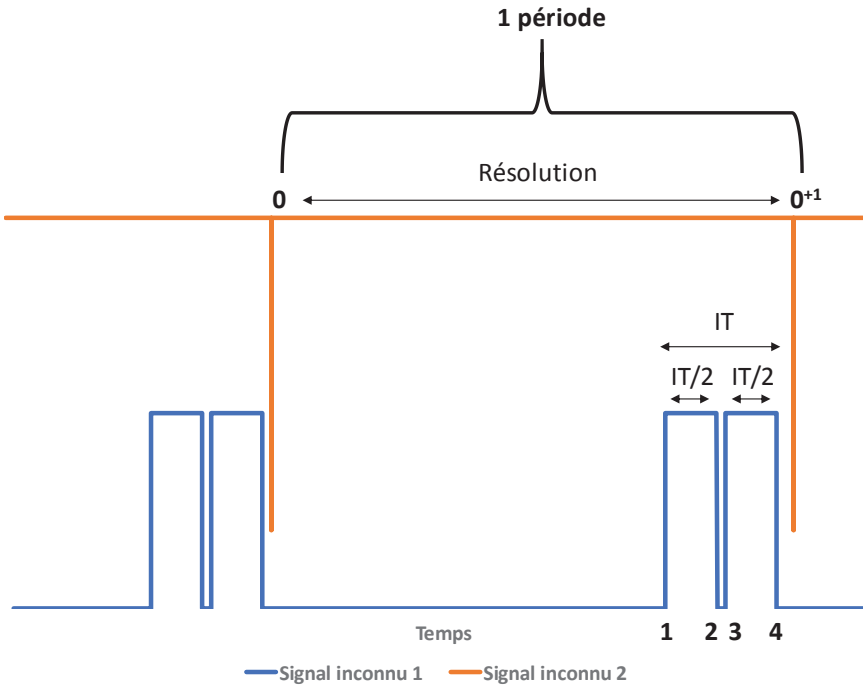


Figure 77 : Signaux inconnus détectés et observés à l'oscilloscope

Lorsque nous allumons le Qexactive avec le paramétrage précisé précédemment (Tableau 30) nous visualisons les 2 signaux inconnus exposés ci-dessus (Figure 77). En vue de tenter de les comprendre nous avons fait varier les paramètres du Qexactive lorsque celui-ci était allumé (sans être couplé à la partie chromatographique et sans analyser des composés en infusion).

Nous nous sommes rendu compte que lorsque nous faisions varier l'ITmax le signal inconnu 1 (bleu) varié aussi de manière relative. En effet lorsque nous avions un ITmax de 250 ms la distance comprise entre le point 1 et 2 correspondait à 125 ms et celle entre le point 2 et 4 aussi. Lorsque nous abaissions l'ITmax à 125 ms l'écart entre les points 1 et 2 se divisait lui aussi par deux tout comme l'écart entre les points 3 et 4. Le signal inconnu 1 semble donc correspondre ou être en relation avec l'ITmax.

Nous nous sommes également rendu compte que lorsque nous faisions varier la résolution le signal inconnu 2 (orange) varié aussi de manière relative. En effet pour une résolution de 140 000 la distance entre les points 0 et 0^{+1} correspondait à 512 ms. Lorsque nous passions à une résolution de 70 000 cette distance était divisée par deux et correspondait à 256 ms. Le signal inconnu 2 semble donc correspondre ou être en relation avec l'entrée des ions dans l'orbitrap.

Nous pouvons visualiser ici le comportement attendu concernant le travail en parallèle de l'orbitrap : les ions semblent bien être accumulés (point 1 à 4) en même temps que l'orbitrap analyse les ions accumulés lors de la période précédente (point 0 à 0^{+1}). Dans le même temps, nous pouvons également voir que l'accumulation des ions dans la C-trap ne semble pas débuter au départ de l'analyse des ions du scan précédent (point 0) mais à rebours par rapport à sa fin (point 0^{+1}). Cela est peut-être dû à des difficultés pour stocker les ions de manière stable dans la C-trap durant de longue durée.

Après ces observations nous pouvons conclure que nous avons visualisé 2 signaux significatifs : un signal semblant être en relation avec l'ITmax et un signal en relation avec l'injection des ions dans l'orbitrap. Dans le but d'une approche la plus simplifiée possible, nous nous sommes tournées vers l'utilisation du signal lié à l'injection des ions dans l'orbitrap en tant que signal déclencheur. En effet, par ce signal nous ne pouvons pas voir l'entrée des ions dans la C-trap mais leur sortie. Chaque entrée des ions dans l'orbitrap au point 0 (Figure 77) correspondrait à la fin d'une étape d'accumulation dans la C-Trap et leur injection pour analyse dans l'orbitrap en même temps que le début d'une nouvelle étape d'accumulation. Nous pouvons donc utiliser ce signal en tant que signal déclencheur en laissant entrer le laser dans la C-trap ou non entre le point 0 et 0^{+1}.

IV.IV.2) Réglage de l'obturateur

Au départ nous souhaitions tester le concept de la photo-fragmentation dans la C-trap avec une approche DIA, cependant cette approche s'est révélée impossible. En effet, pour réaliser cette

méthodologie nous sommes obligés de paramétrer le Qexactive en mode FMS pour obtenir le trajet des ions désirés :

Source d'ionisation → Q1 → C-trap → Orbitrap

Voici les variables auxquelles nous avons accès lorsque nous paramétrons le Qexactive en mode FMS (Figure 78) :

Properties of Full MS — SIM	
General	
Runtime	0 to 10 min
Polarity	positive
Full MS — SIM	
Resolution	70,000
AGC target	3e6
Maximum IT	200 ms
Scan range	150 to 2000 m/z

Figure 78 : Paramètres modifiables lors de l'utilisation du Qexactive en mode FMS

Il est possible de modifier la gamme de masse mais pas séparément au niveau du quadripôle et de l'orbitrap : cela est une limitation pour notre approche. En effet, si, par exemple, nous souhaitons sonder une fenêtre de précurseur en DIA allant de 400 à 432 Th, nous réglons la gamme de masse en fonction de ces 2 valeurs : le quadripôle laissera bien passer tous les ions compris entre 400 et 432 Th, comme souhaité, mais l'orbitrap de son côté n'enregistrera que le signal correspondant à cette plage de masse également. Nous ne pourrons pas avoir accès à toute l'information qui nous intéresse concernant les ions fragments (150-1200 Th). Sans l'aide du constructeur nous n'avons pu remédier à ce problème. Il nous a donc fallu fixer et travailler sur une plage de masse allant de 150 à 1200 Th pour avoir accès, à la fois, à tous les ions précurseurs et à tous les fragments générés lors de l'étape de photo-fragmentation dans la C-trap.

Ensuite, il nous a fallu régler l'obturateur. En effet, en fonction de l'analyse souhaitée la boucle peut être codée différemment au sein du logiciel de programmation (pour combien de scans l'obturateur est-il laissé ouvert ou fermé). Face à la limitation découlant de la gamme de masse imposée nous nous sommes tournés vers une alternance obturateur ouvert/fermé tout au long de l'analyse :

t_o = scan 1 = Obturateur fermé (FMS)

t_1 = scan 2 = Obturateur ouvert (C-trap-LID)

t_2 = scan 3 = Obturateur fermé (FMS)

t_3 = scan 4 = Obturateur ouvert (C-trap-LID)

t_4 = scan 5 = Obturateur fermé (FMS)

Etc…

Cette alternance peut être apparentée au mode AIF-LID (alternance binaire de scan FMS et LID) lors d'une photo-fragmentation laser dans la cellule HCD. Ce concept a par le passé déjà été utilisé pour identifier des peptides à cystéine dérivés photo-fragmentés grâce à l'information offerte par les ions précurseurs[9]. En effet, lors d'un scan FMS tous les ions précurseurs (relatifs à des peptides à cystéine dérivés ou non) sont détectés à une intensité donnée. Lors du scan LID les intensités relatives aux ions précurseurs issus des peptides à cystéine dérivés vont fortement diminuer en raison de leur photo-fragmentation. En revanche, les ions précurseurs qui ne sont pas issus de peptides à cystéine dérivés ne verront pas leur intensité variée. Il devient alors possible, par cette approche, de discriminer les ions précurseurs issus de peptides à cystéine dérivés des autres par le suivi des fluctuations en intensité de leurs ions précurseurs respectifs (approche différentielle).

Cette possibilité peut s'avérer fortement utile dans certains cas. En effet, lors d'analyse DDA-LID, dans l'optique de construire une bibliothèque de spectres ou non, l'information relative aux peptides de faible intensité est perdue en raison de leurs non sélections pour fragmentation. De ce fait, ces peptides de faible intensité ne pourront ni être détectés en DDA-LID, ni en DIA-LID car ils seront absents de la bibliothèque de spectres. S'il est possible de commander ces composés de faible intensité sous forme de peptides synthétiques cela n'est pas toujours possible : d'une part pour une question de prix et d'autre part par le simple fait que les analyses DDA et DIA sont principalement utilisées pour sonder des échantillons à l'aveugle et que ce qui est cherché n'est que très rarement anticipé. Par cette approche, il devient possible d'extraire tous les *m/z* des ions précurseurs ayant photo-fragmentés correspondant à tous les peptides à cystéine dérivés détectés par l'appareillage. Une analyse ciblée type PRM-LID ou DDA-LID (avec une liste d'inclusion) peut ensuite être appliquée sur ces *m/z*, permettant d'acquérir leurs spectres de photo-fragmentation pour identification ou incorporation à une bibliothèque de spectres. Nous avons donc décidé de nous tourner vers ce concept pour expérimenter la méthodologie C-trap-LID.

En vue d'être le plus résolutif et sensible possible lors de cette analyse, une résolution de 140 000 a été sélectionnée ainsi qu'un ITmax de 250 ms et une AGC Target de 1^E6. Pour mettre en place la synchronisation, nous avons codé une boucle sur le logiciel de programmation de l'arduino qui engendre une alternance de l'obturateur (ouverture/fermeture) lors d'une injection sur 2 des ions dans l'orbitrap (Figure 77 point 0).

Cette configuration sera conservée pour toutes les analyses présentées ci-dessous.

IV.V) Mise en application du concept C-Trap-LID à 473 nm

Une fois la méthodologie C-trap-LID à 473 nm mise en place, il a été décidé de l'appliquer à l'étude de peptides à cystéine dérivés en matrice aqueuse en vue d'essayer de dégager la portée et les limites intrinsèques du concept. Dans cette optique, 10 peptides à cystéine fournis par Thermo Scientific ont été solubilisés dans une matrice aqueuse, dérivés avec le protocole type, séparés avec le gradient 3 puis analysés en mode C-trap-LID à 473 nm.

Il sera présenté ci-dessous la calibration du système nécessaire à cette analyse ainsi que les résultats et informations que nous avons pu extraire à tous les niveaux.

Il est à préciser que tous les traitements effectués dans ce chapitre ont été fait manuellement dans Excel et scan par scan. En effet, nous n'avons pu automatiser le traitement de données. Celui-ci aurait nécessité une division du fichier raw initial en 2 sous-fichier : un fichier FMS contenant les scans impairs (le scan 1 est un FMS) et un fichier avec les spectres de photo-fragmentation contenant les scans pairs. Cela n'a pu être effectué et toutes les analyses ont été traitées à l'aide de leurs fichiers initiaux.

IV.V.1) Calibration avec le DABCYL-C2 maléimide

Une fois la synchronisation établie entre l'obturateur et le Qexactive et le paramétrage imposé, il a fallu, avant toute analyse, mettre en place un système de contrôle de l'alignement du laser. En effet, nous devions nous assurer que la fragmentation ait bien lieu dans la C-Trap. Cette étape est basée, comme précédemment, sur l'utilisation du DABCYL-C2 maléimide. Celui-ci a été solubilisé dans du méthanol à une concentration d'environ 50 µg/mL puis infusé seul au sein du spectromètre à un débit de 10 µL/min. Il est ensuite analysé en mode C-Trap-LID à 473 nm. Ci-dessous sont présentés les spectres de masses obtenus lorsque l'obturateur et en position ouverte et fermé (Figure 79) :

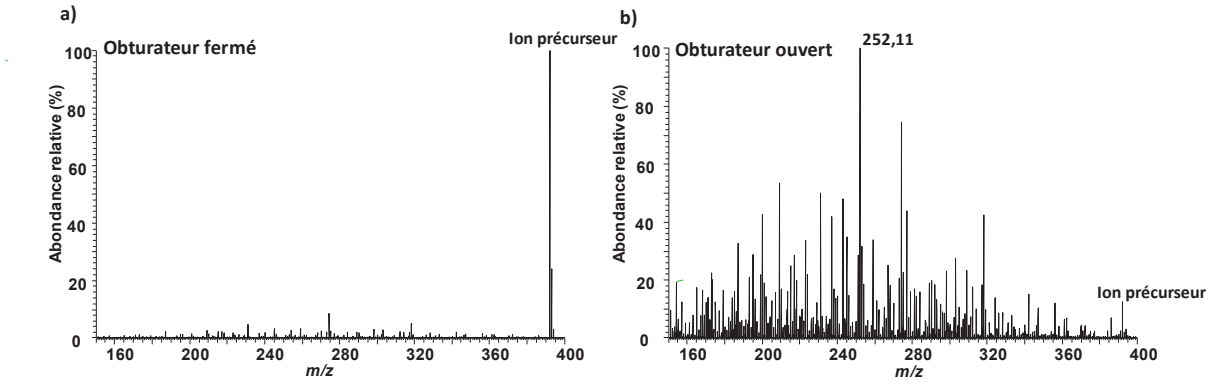


Figure 79 : Spectre de masse obtenu lors de l'infusion du DABCYL C2-maléimide lorsque l'obturateur est fermé (a) ou ouvert (b)

Nous pouvons visualiser sur les spectres obtenus, comme attendu, que lorsque l'obturateur est fermé l'ion précurseur à *m/z* 392,17 est observé sans aucun fragment (Figure 79 a) : il n'y a pas la légère présence de l'ion rapporteur qui pourrait résulter de la dégradation du chromophore en solution. En revanche, lorsque, le scan suivant, l'obturateur est ouvert nous pouvons observer la production de l'ion rapporteur *m/z* 252,11 en majorité (Figure 79 b). L'ion précurseur est, comme lors de la photo-fragmentation laser du DABCYL-C2 maléimide dans la cellule HCD, à une intensité relative inférieure à 20 % par rapport à celui-ci. Tous les autres ions observés sont, soit issus de la photo-fragmentation du DABCYL-C2 maléimide, soit ce sont des ions présents dans la source d'ionisation et/ou le spectromètre de masse. En effet, effectuant un FMS sur une plage importante de 150 à 1200 Th beaucoup d'ions sont transmis en plus du chromophore.

Via ces observations, nous pouvons démontrer la bonne synchronisation de l'obturateur et du Qexactive : quand l'obturateur est ouvert les ions précurseurs photo-fragmentent, quand il est fermé il n'y a pas de photo-fragmentation. Avant toutes les analyses qui suivent il a été fait en sorte d'obtenir ces résultats pour le chromophore indiquant un bon alignement du laser.

Ci-dessous sont présentés les XIC correspondants à l'ion précurseur du DABCYL-C2 maléimide et à l'ion rapporteur lors de la même analyse (Figure 80) :

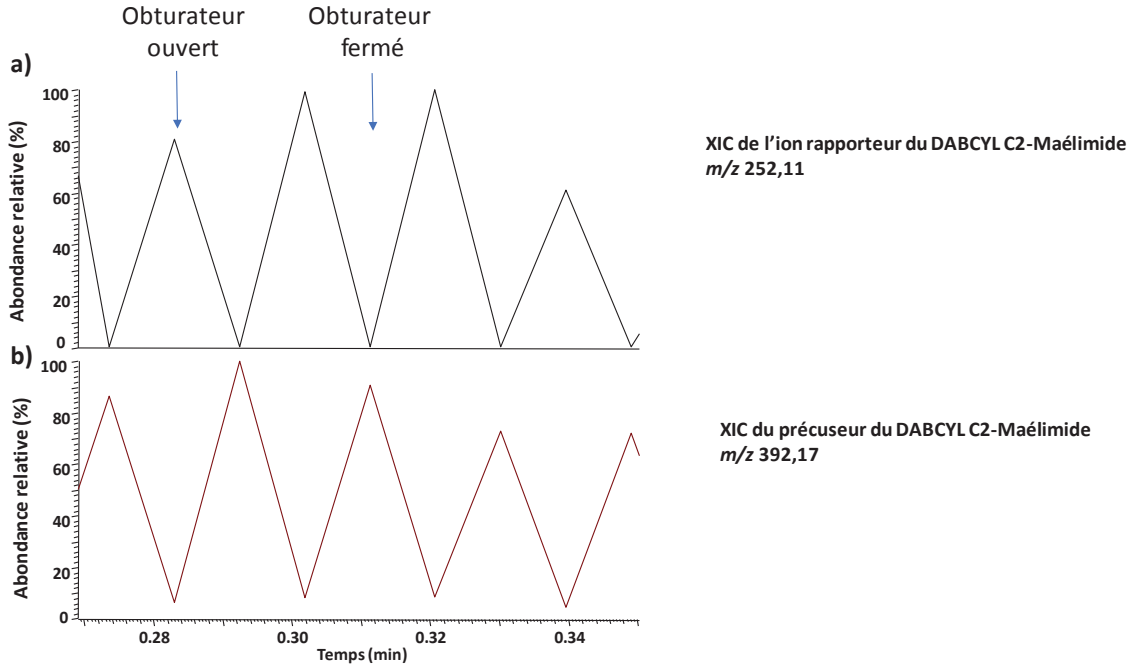


Figure 80 : XIC correspondant à l'ion rapporteur m/z 252,11 du DABCYL C2-maléimide (a) et à l'ion précurseur du DABCYL C2-maléimide (b)

Ce qu'il est important d'observer ici est le comportement alterné caractéristique de l'ion précurseur et de son fragment à travers leurs XIC respectifs : ceux-ci sont en dents de scies mais de manière décalée (Figure 80). En effet, nous savons, que l'obturateur est ouvert un scan sur

deux. De ce fait lorsque celui-ci est fermé, l'intensité de l'ion précurseur est maximale, en revanche, celle de son fragment principal est inexistante. A l'inverse, le scan suivant, lorsqu'il est ouvert nous pouvons observer une nette diminution du signal de l'ion précurseur (90 %) démontrant la photo-fragmentation de l'ion précurseur mais rappelant qu'elle n'est pas totale et une nette augmentation de l'intensité de l'ion rapporteur (80 %). Si cette observation est cohérente avec les spectres de masses obtenus, elle permet surtout la détection d'ions précurseurs ayant photo-fragmentés à l'œil nu. En effet, comme nous le verrons par la suite, un ion précurseur possédant un XIC en dents de scie sera la marque que celui-ci a été dérivé avec le DABCYL C2-maléimide et qu'il a photo-fragmenté. En revanche, si son XIC ne présente pas cette caractéristique mais une trame linéaire alors il n'aura pas été dérivé avec le DABCYL C2-maléimide et n'aura pas photo-fragmenté. Nous pourrons en faire de même avec les ion fragments issus de ces peptides à cystéine dérivés.

IV.V.2) Photo-fragmentation laser à 473 nm de peptides à cystéine dérivés avec le DABCYL C2-maléimide en utilisant la méthodologie C-Trap-LID

Une fois la calibration effectuée et la démonstration du concept C-Trap-LID à 473 nm à faire photo-fragmenter le DABCYL C2-maléimide, nous avons pu passer directement à l'analyse de peptides à cystéine dérivés en matrice aqueuse. Différentes observations ont pu être faites à travers ces premières analyses. Celles-ci s'attarderont sur l'aspect des XIC obtenus, le taux de photo-fragmentation dans la C-Trap et les informations qui peuvent être extraites des spectres C-Trap-LID obtenus.

IV.V.2.a) Etude des XIC

Différentes constatations ont pu être notées quant à l'étude des XIC et pourront être traduites, en partie, à l'aide des deux peptides à cystéine dérivés suivant LTSVVGTILYSCPEVLK et VLDANSCQSELHEK (Figure 81) :

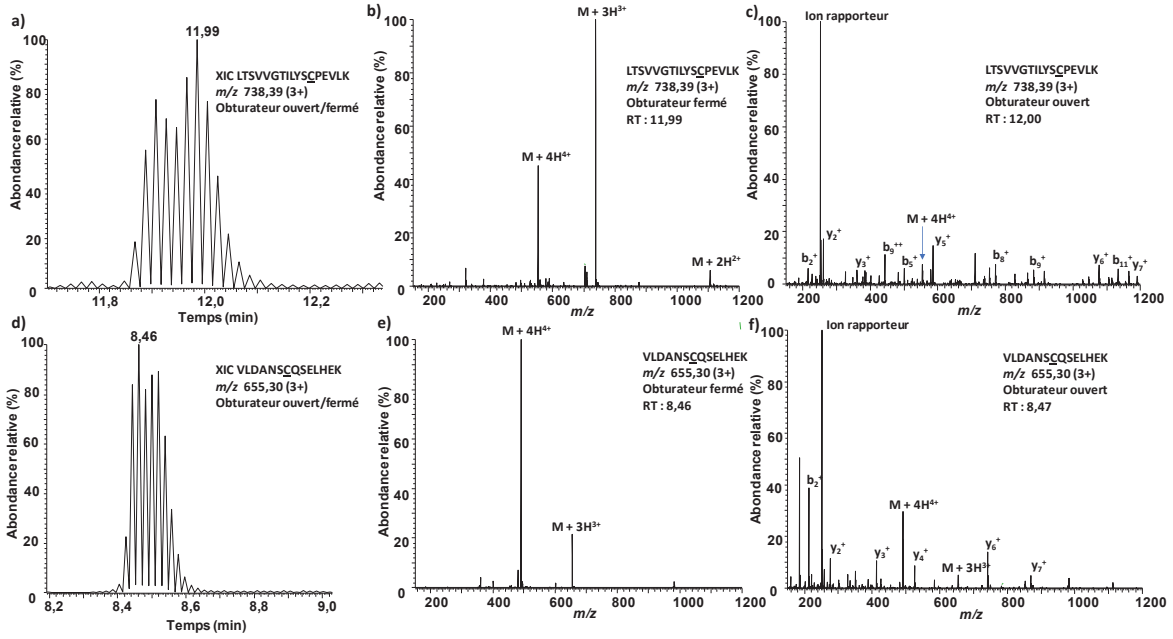


Figure 81 : XIC (a) et spectre de masse obtenu pour le peptide à cystéine dérivé LTSVVGTILYSCPEVLK lorsque l'obturateur est fermé et ouvert (b et c) ainsi que pour le peptide à cystéine dérivé VLDANSCQSELHEK (d, e, f)

Tout d'abord, nous pouvons visualiser l'aspect des XIC correspondant aux ions précurseurs de ces deux peptides à cystéines dérivés : ils sont bien en dents de scie (Figure 81 a et d). Ceux-ci sont en correspondance avec l'alternance ouverture/fermeture de l'obturateur engendrant une diminution ou non du signal correspondant aux ions précurseurs. Nous pouvons voir ici plusieurs états de charge pour les 2 peptides en raison de la large fenêtre de sélection en *m/z* (Figure 81 b et e). Le spectre de photo-fragmentation obtenu correspond à celui sommé de tous les états de charge (Figure 81 c et f). Nous voyons que lorsque l'obturateur est fermé les masses des ions précurseurs correspondant sont bien visualisées (Figure 81 b et e) en revanche lorsque celui-ci est ouvert l'ion rapporteur *m/z* 252,11 et tous les ions b et y correspondant aux peptides fragmentés sont détectés ainsi que le signal de leurs ions précurseurs respectifs nettement diminués en raison de la photo-fragmentation (Figure 81 c et f). Nous démontrons donc ici l'efficacité du concept C-Trap-LID à 473 nm pour faire photo-fragmenter des peptides à cystéine dérivés au sein du spectromètre de masse.

D'autres observations peuvent être faites au niveau de ces analyses tout d'abord traduite par les peptides à cystéine dérivés LTSVVGTILYSCPEVLK et LQGINCGPDFTPSFANLGR (Figure 82).

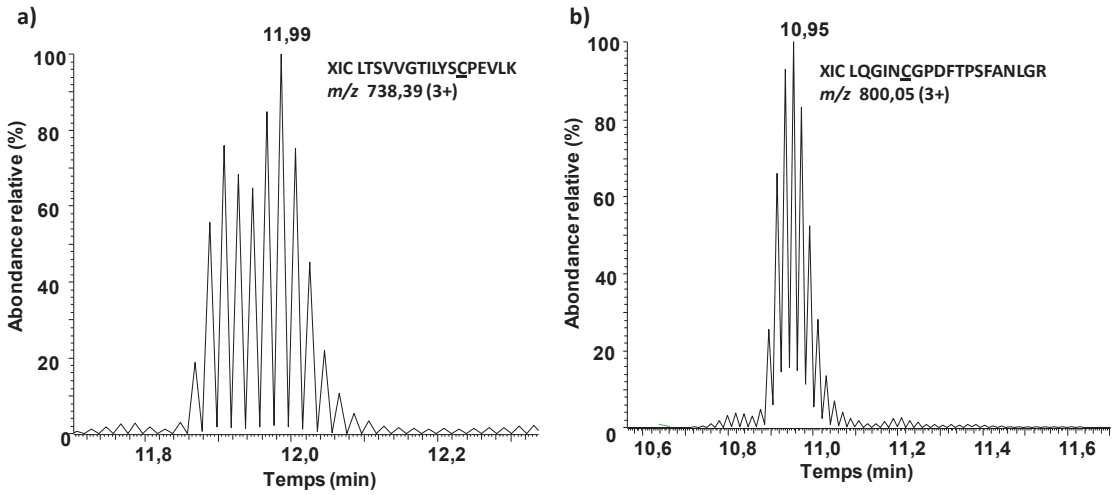


Figure 82 : XIC des peptides à cystéine dérivés LTSVVGTILYSCPEVLK (a) et LQGINCGPDFTPSFANLGR (b)

D'après les 2 XIC extraits il est possible à l'œil nu d'avoir une estimation du taux de photo-fragmentation des peptides à cystéine dérivés comparativement les uns aux autres. En effet, pour le peptide à cystéine dérivé LQGINCGPDFTPSFANLGR nous constatons un retour à la ligne de base moindre du signal correspondant à l'ion précurseur trichargé comparativement à celui de LTSVVGTILYSCPEVLK (Figure 82 b et a). Cela signifie que LQGINCGPDFTPSFANLGR semble moins bien photo-fragmenter que LTSVVGTILYSCPEVLK.

Une autre observation relative à l'esthétique des XIC et allant dans le même sens a pu également être faite. Ci-dessous est présenté le TIC durant un laps de temps particulier (Figure 83) :

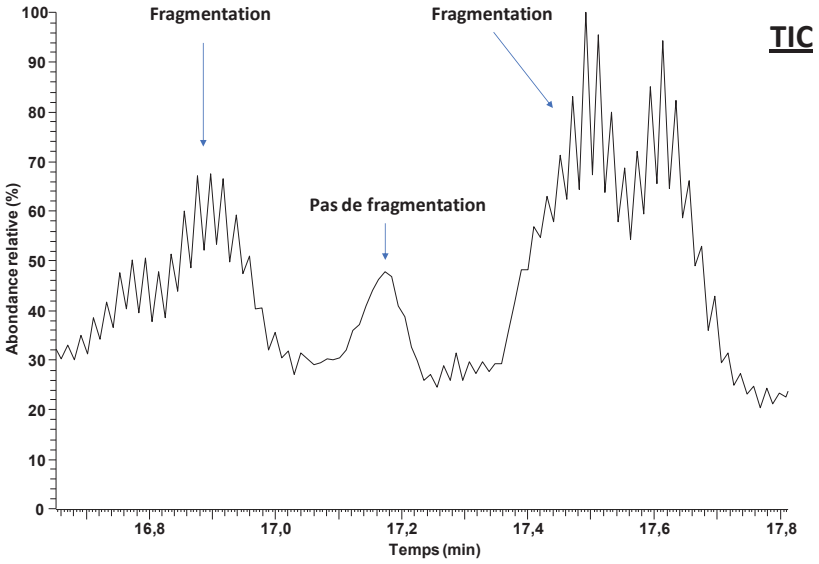


Figure 83 : TIC extrait entre 16,6 et 17,8 min

A l'aide de cette figure nous pouvons remarquer deux zones spécifiques : grossièrement l'une comprise entre 17 et 17.25 min et l'autre occupant le reste du temps. Dans la plus grosse des deux zones, nous pouvons observer un TIC en dents de scie indiquant la présence d'ions précurseurs correspondants à des composés dérivés et donc photo-fragmentant 1 scan sur 2. En revanche, dans la plus petite des deux zones, entre 17 et 17.25 min, nous pouvons constater l'absence de cet aspect en dents de scies. Cela signifie que dans cet espace du chromatogramme aucun composé n'est dérivé avec le DABCYL-C2 maléimide et donc, ne photo-fragmente pas.

Cette figure révèle la potentialité, à l'œil nu de visualiser les domaines du TIC et les temps de rétention auquel des composés intéressants (car dérivés et photo-fragmentant) se trouvent. Dans ces mêmes zones, nous pouvons extraire les XIC de tous les ions précurseurs présents à chaque scan où l'obturateur est fermé et visualisé si ceux-ci présentent un aspect en dents de scie ou non. Ci se travail manuel est long il permet néanmoins de discriminer avec certitude si un ion précurseur est dérivé ou non.

IV.V.2.b) Taux de photo-fragmentation

Dans le même temps il a été tenté d'estimer le taux de photo-fragmentation de ces peptides à cystéine dérivés en mode C-Trap-LID à 473 nm. Pour estimer le taux de photo-fragmentation la formule suivante a été utilisée (Equation 19)

$$\textbf{Taux de photo-fragmentation} = \frac{\mathbf{E\ Ifrag - I252{,}11}}{\mathbf{E\ Ifrag + Iparent}} \mathbf{\ X\ 100} \qquad \textbf{Equation 19}$$

Voici les résultats obtenus pour les 10 peptides à cystéine dérivés (Tableau 31) :

Peptides à cystéine dérivés	Taux de photo-fragmentation (%)
GLCAIAQAESLR	91
LQGINCGPDFTPSFANLGR	77
TAAQCLEHYEFLLDK	69
WCSWSLSQAR	95
EAPPCHLLIIHR	96
SNCITENPAGPYGQR	89
LTSVVGTILYSCPEVLK	89
ALEAGGCAVVR	90
VLDANSCQSELHEK	72
FVDDVNNNYYEAPSCPR	89

Tableau 31 : Taux de photo-fragmentation obtenus pour les 10 peptides à cystéine dérivés

Nous pouvons observer ici qu'un taux de photo-fragmentation moyen de 86 % est obtenu pour les 10 peptides à cystéine dérivés avec un minimum à 69 % confirmant la photo-fragmentation des peptides à cystéine dérivés dans la C-Trap. Cependant, on dénote des taux plus faibles pour 3 peptides à cystéine dérivés (inférieur à 80 %) contrairement à ceux obtenus lors de la photo-

fragmentation laser dans la cellule HCD : cela est facilement explicable. En effet, durant cette analyse beaucoup de peptides à cystéine dérivés ont co-élués (les 10 sont sortis entre 8,5 et 12 min) et, en raison du paramétrage utilisé, il n'y a eu aucune sélection en précurseur : toute la gamme de *m/z* allant de 150 à 1200 Th a été photo-fragmentée en même temps à chaque instant. Or tous les peptides à cystéine dérivés vont produire majoritairement l'ion rapporteur à 252,11 Th. Ici, il n'a pu être pris en compte dans le calcul du taux de photo-fragmentation car nous ne pouvons pas le réattribuer proportionnellement à tel ou tel ion précurseur. Avec cet ion incorporé, il est assuré de dépasser les 80 % de taux de photo-fragmentation pour tous les peptides à cystéine dérivés. Dans le même temps, nous pouvons voir que même sans l'inclure, nous obtenons des taux de photo-fragmentation élevés : une moyenne à 86 % avec un maxima à 96 % et 7 peptides à cystéine dérivés sur 10 présentent un taux de photo-fragmentation supérieur à 88 %.

Aucun scan n'a été trouvé où l'ITmax fut atteint ne permettant pas encore de trancher sur l'éventuel diminution du taux de photo-fragmentation dans ces cas.

Ces résultats démontrent que le taux de photo-fragmentation lors d'une photo-fragmentation dans la C-Trap est en correspondance avec celui obtenu lors de la photo-fragmentation laser dans la cellule HCD. Il est donc optimal et devrait être suffisant pour mener à l'identification de peptides à cystéine dérivés au sein de matrice aqueuses et biologiques.

IV.V.2.c) Identification de peptides à cystéine dérivés en utilisant la méthodologie C-Trap-LID à 473 nm

Une fois démontré que les peptides à cystéine dérivés pouvaient être photo-fragmenté dans la C-Trap avec des taux de photo-fragmentation optimaux, nous nous sommes attardés sur les différentes manières d'identifier ces mêmes peptides. Bien que le principal intérêt lié au paramétrage choisi réside dans l'étude des signaux relatifs aux ions précurseurs, nous avons également étudié les signaux issus des ions fragments.

IV.V.2.c.1) Identification par les fragments

Le premier procédé que nous avons choisi pour identifier les peptides à cystéine dérivés photo-fragmentés au sein de la C-trap reposait sur l'identification par les fragments. Cette voie est la plus classique et nous voulions montrer qu'en photo-fragmentant les ions dans la C-trap, suffisamment d'ions fragments étaient générés, stockés et transmis dans l'orbitrap pour permettre l'identification de peptides à cystéine dérivés photo-fragmentés. Pour ce faire, nous nous sommes attardés sur la diversité des fragments obtenue pour chaque peptide à cystéine dérivé : nous avons extrait manuellement le nombre de fragments détectés pour chacun des peptides à cystéine dérivés. Nous nous sommes exclusivement focalisés sur les ions b et y mono et di-chargés. L'ion rapporteur n'a pas été incorporé dans ces résultats car jugé non discriminant. Les valeurs sont issues du scan de photo-fragmentation le plus intense obtenu pour chaque peptide à cystéine dérivé (à l'apex du XIC). Voici les résultats obtenus (Tableau 32) :

Peptides à cystéine dérivés	Nombre de fragments détectés
LTSVVGTILYSCPEVLK	20
LQGINCGPDFTPSFANLGR	23
SNCITENPAGPYGQR	16
FVDDVNNNYYEAPSCPR	15
VLDANSCQSELHEK	15
EAPPCHLLIIHR	16
GLCAIAQAESLR	9
ALEAGGCAVVR	11
TAAQCLEHYEFLLDK	8
WCSWSLSQAR	11

Tableau 32 : Nombre de fragments détectés pour les 10 peptides à cystéine dérivés

Nous pouvons remarquer que tous les peptides à cystéine dérivés ont été observés par l'intermédiaire de leurs fragments : 8 fragments au minimum par peptide à cystéine dérivé avec un maximum à 23 pour le peptide LQGINCGPDFTPSFANLGR. La diversité de fragment produite, stockée et transmise lors d'une étape de photo-fragmentation à 473 nm dans la C-trap est donc suffisante pour identifier des peptides à cystéine dérivés. Cette méthodologie peut donc potentiellement servir à identifier des peptides à cystéines dérivés dans des matrices diverses à l'aide de leurs fragments.

IV.V.2.c.2) Identification par les précurseurs

La seconde voie que nous avons choisie pour identifier des peptides à cystéine dérivés photo-fragmentés au sein de la C-trap repose sur l'information offerte par les ions précurseurs. En effet, lorsque l'obturateur est fermé tous les ions précurseurs (relatifs aux peptides à cystéine dérivés ou non) sont détectés à une intensité donnée. Lorsque l'obturateur est ouvert, les intensités relatives aux ions précurseurs issus des peptides à cystéine dérivés vont fortement diminuer en raison de leur photo-fragmentation. En revanche, les ions précurseurs qui ne sont pas issus de peptides à cystéine dérivés ne verront pas leurs intensités varier. Il devient alors possible de discriminer les ions précurseurs issus de peptides à cystéine dérivés des autres par le suivi des fluctuations en intensité de leurs ions précurseurs respectifs (approche différentielle) lorsque l'obturateur est ouvert et fermé.

Pour aborder cette voie, les 10 peptides à cystéine dérivés ont été introduits dans un mélange de 3 protéines. L'échantillon a été solubilisé dans une matrice aqueuse puis dérivé avec le protocole type, séparé avec le gradient 4 puis analysé en mode C-trap-LID à 473 nm. Comme précisé auparavant, nous n'avions aucune automatisation pour le traitement de donnée. Cela a engendré d'une part, un traitement manuel et d'autre part, une extraction de l'information scan par scan. Le traitement présenté est précaire mais suffisant pour démontrer le potentiel de la méthodologie pour cette approche. Il est voué à être développé, amélioré et automatisé pour

pouvoir être utilisé dans des approches protéomiques concrètes. Voici comment nous avons procédé :

Premièrement, 2 scans successifs ont été sélectionnés : 1 scan FMS (obturateur fermé) et 1 scan C-trap-LID à 473 nm (obturateur ouvert) correspondant au RT de l'apex du pic chromatographique du peptide à cystéine dérivé SNCITENPAGPYGQR. Nous souhaitions travailler sur deux scans où nous étions sûr de détecter au moins 1 ion précurseur correspondant à 1 peptide à cystéine dérivé.

Secondement, les 1000 points les plus intenses de chaque spectre de masse ont été extraits du fichier raw. Tous les pics correspondants à des masses inférieures à 200 Th ainsi que les pics où la charge n'a pu être attribuée par le logiciel Xcalibur ont été supprimés car nous les avons considérés comme non informatifs.

Troisièmement, pour chaque spectre (FMS et C-trap-LID à 473 nm) il n'a été gardé que les ions ayant un signal supérieur ou égal à 3 fois le signal sur bruit qui était attribué à leur pic respectif par le logiciel Xcalibur (en termes d'intensité). Les ions n'ayant pas ce minimum ont été considérés comme étant du bruit et par là-même non significatifs.

Quatrièmement, toutes les masses des 2 spectres (FMS et C-trap-LID à 473 nm), gardées après ses premiers traitements, ont été cherchées mutuellement dans l'autre spectre en vue de visualiser leur apparition, disparition, (co)-présence au sein de ces 2 spectres (toutes les masses gardées issues du spectre FMS ont été recherchées dans le spectre LID à 473 nm et inversement). Pour ce faire, chaque masse s'est vu affilier un Δm de +/- 0,007 Th. Si une masse présente dans le FMS est retrouvée dans le spectre LID à 473 nm dans l'intervalle +/- 0,007 Th, elles étaient assimilées comme identiques et inversement. Si non, elles étaient assimilées comme appartenant uniquement au spectre dont elles étaient issues.

Cinquièmement nous avons effectué le calcul suivant pour l'approche différentielle :

Intensité des ions présents dans le spectre C-trap-LID à 473 nm (obturateur ouvert) – Intensité des ions présents dans le spectre FMS (obturateur fermé)

De cette manière, tous les ions précurseurs dérivés auront un signal négatif car ils ne seront pas totalement photo-fragmentés, les ions précurseurs n'ayant pas photo-fragmentés auront une valeur proche de 0 car leurs intensités n'auront quasiment pas bougé entre les 2 spectres et tous les ions fragments issus de la photo-fragmentation d'un ion précurseur dérivé avec le DABCYL-C2 maléimide auront un signal positif. Nous avons vu au laboratoire après de nombreuses analyses que l'intensité d'un ion précurseur rapporteur d'un peptide à cystéine dérivé lors d'une photo-fragmentation LID à 473 nm perdait entre 95 % et 65 % de son intensité. De ce fait, il a été considéré que tous les ions précurseurs qui perdaient moins de 65 % de leurs signaux entre le spectre FMS et C-trap-LID à 473 nm n'étaient pas issus de peptides à cystéine dérivés.

Voici la représentation finale obtenue (Figure 84) :

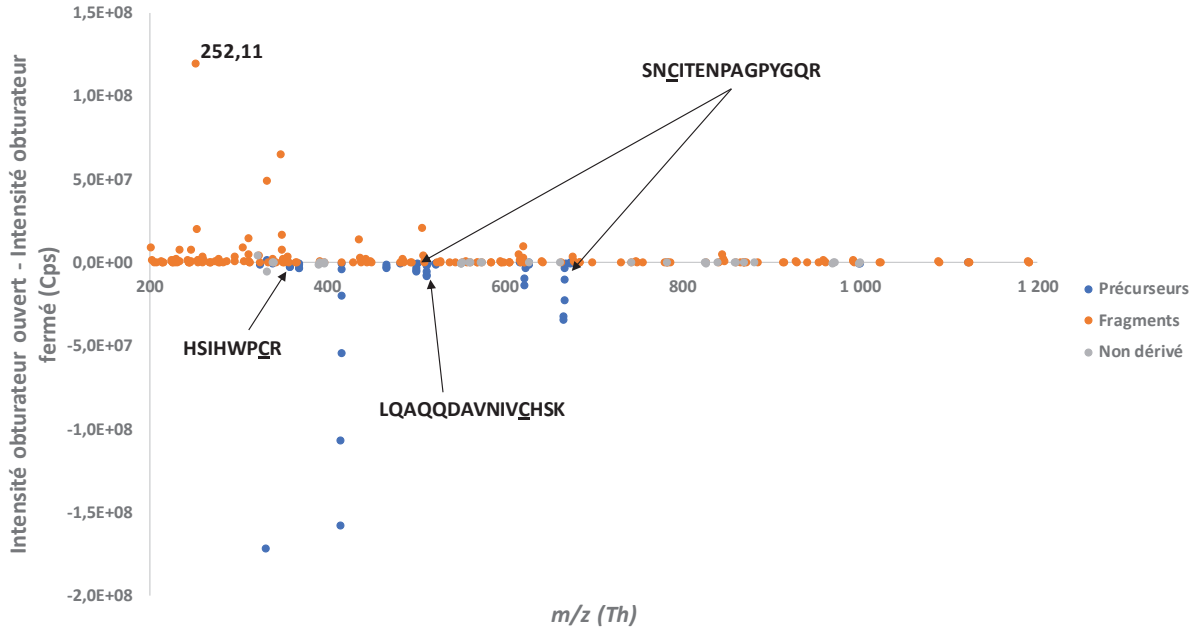


Figure 84 : Soustraction des scans FMS et C-Trap-LID à 473 nm au RT correspondant à l'apex du pic chromatographique du peptide à cystéine SNCITENPAGPYGQR dans un digest de 3 protéines

Sans exclure les isotopes 34 masses issues de composés non dérivés ont été extraites (signal = 0 et coloré en gris sur Figure 84), 58 appartenant à des ions précurseurs (signal négatif et coloré en bleu sur Figure 84) et 178 à des fragments (signal positif coloré en orange sur Figure 84) à ce temps de rétention. Afin d'extraire une liste de précurseur correspondant aux masses monoisotopiques, les isotopes ont été exclus manuellement. Via ce traitement une liste de 16 masses correspondant à des ions précurseurs ayant photo-fragmentés a pu être extraite (Tableau 33) :

m/z	Identification possible	Charge
324,119	?	?
357,42	HSIHWPCR	4^+
367,67	?	?
415,191	?	?
466,554	?	?
482,231	?	?
500,227	SNCITENPAGPYGQR	4^+
512,005	LQAQQDAVNIVCHSK	4^+
522,255	?	?
617,270	?	?
622,280	?	?
626,504	?	?
666,633	SNCITENPAGPYGQR	3^+
673,961	?	?
999,446	SNCITENPAGPYGQR	2^+
1122,779	?	?

Tableau 33 : Masses extraites correspondants à des peptides à cystéine dérivés putatifs

Parmi ces 16 m/z, 5 (relatifs aux peptides à cystéine introduits dans le mélange des 3 protéines) ont pu être identifiés directement (Figure 84 et Tableau 33). Le *m/z* à 357,42 correspond à l'ion précurseur quadrichargé du peptide à cystéine dérivé HSIHWPCR, ceux aux *m/z* 500,23 et 666,63 et 999,45 aux ions précurseurs di, tri et quadrichaghés du peptide à cystéine dérivé SNCITENPAGPYGQR et enfin celui à *m/z* 512,01 à l'ion précurseur quadri-chargé du peptide à cystéine dérivé LQAQQDAVNIVCHSK. Nous pouvons également voir 5 points très proches à proximité du *m/z* correspondant à l'ion $M+3H^{3+}$ du peptide à cystéine dérivé SNCITENPAGPYGQR : ce sont les isotopes. Tous les autres m/z sont inconnus (ceux-ci peuvent être issus des peptides à cystéine dérivés appartenant aux 3 protéines, d'un relargage de la colonne chromatographique ou de la présence d'un composé annexe non attendu en raison de l'impureté de l'échantillon du pool de peptides à cystéine de départ) et peuvent être incrémentés au sein d'une liste d'inclusion en vue d'une analyse PRM ou DDA-LID à 473 nm pour être identifiés à l'aide de leurs spectres de photo-fragmentation. Il est donc bien possible par cette approche différentielle et de la méthodologie C-Trap-LID à 473 nm de retrouver des peptides à cystéine dérivés au sein d'un échantillon à l'aide de l'information fournit par les ions précurseurs. En effet, si cette opération avait été effectuée à l'aveugle sans connaissance de l'échantillon, ces peptides à cystéine auraient pu être identifiés à l'aide de l'approche ciblée qui aurait suivie. Qui plus est, un contrôle visuel de tous les XIC correspondants aux 16 masses peut être effectué pour contrôler leurs aspects en dents de scie et confirmer leurs photo-fragmentation et donc dérivation avec le DABCYL C2-malémide.

Des masses non dérivées ont également été observées avec une valeur différentielle en intensité de 0 (surement des peptides sans cystéine issus de la digestion des 3 protéines) ainsi que de nombreux fragments (valeur différentielle en intensité positive). Parmi ces fragments nous avons pu retrouver l'ion rapporteur *m/z* 252,11 et 7 issus du peptide à cystéine dérivé SNCITENPAGPYGQR (nous n'avons pas regardé les ions fragments extraits pour les 2 autres peptides à cystéine car à ce RT nous ne sommes pas du tout à l'apex de leurs pics chromatographiques respectifs mais à leurs extrémités). Manuellement, nous en retrouvions 11 ce qui implique une perte d'information durant le traitement de données en raison, certainement, de la limitation des 1000 pics les plus intenses. En effet, par la méthode d'extraction actuelle des *m/z* nous ne sommes pas exhaustifs : nous n'avons pas accès aux *m/z* des pics les plus faiblement intenses du spectre FMS (ainsi que celui du C-Trap-LID à 473 nm) alors que c'est ce que nous cherchons à obtenir. Il nous faut donc absolument trouver un moyen qui nous permette d'extraire tous les *m/z* et de les traiter librement ensuite sans contrainte (cette démarche est en cours actuellement au laboratoire). Cependant, rien qu'avec cette approche, l'information extraite est bien plus exhaustive qu'avec une approche DDA où l'on se limite, le plus souvent, à 10 ions précurseurs différents photo-fragmentés et donc identifiables par scan. Ici, une liste potentielle bien plus conséquente de peptides à cystéine dérivés putatifs peut être extraite pour chaque scan.

Conclusion

Dans ce chapitre nous avons introduit une nouvelle méthodologie permettant de réaliser l'étape de photo-fragmentation directement au sein de la C-trap du Qexactive et non plus dans la cellule HCD. Intitulé C-trap-LID, nous avons démontré l'efficacité de ce concept pour faire photo-fragmenter de manière optimale et identifier des peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm. Concernant les apports de cette méthodologie comparativement à une approche classique (fragmentation dans la HCD), ils sont minces mais notables. Grâce à l'utilisation de la C-trap-LID, il est possible de discriminer à l'œil nu, de par l'aspect de son XIC, si un ion précurseur à photo-fragmenté ou non, de peut-être gagner légèrement en sensibilité dans certains cas et d'extraire une liste exhaustive des *m/z* correspondant à tous les peptides à cystéine dérivés au sein d'un échantillon par une approche différentielle.

L'objectif établi en résultante fut d'utiliser la C-trap-LID à 473 nm en mode DIA. Cependant, en raison de l'impossibilité de pouvoir contrôler séparément les gammes de masse relatives au quadrupole et à l'orbitrap en mode FMS sur le Qexactive, cela n'a pu être réalisé. Il semble indispensable, en vue d'exploiter tous le potentiel de ce concept, d'avoir accès à ce paramètre sans quoi nous ne pourrons aller plus loin expérimentalement. Si cela est à un moment rendu possible, il sera nécessaire de mettre en place en parallèle une approche computationnelle spécifique à la C-trap-LID. En effet, tous les scans sont acquis en mode FMS sur le Qexactive et sont donc assimilés comme tel par les logiciels classiques traitant les données en protéomique. Il faut donc développer une approche qui serait capable d'extraire toutes les informations pour chaque scan, de discriminer les scans où l'obturateur est fermé (FMS) et où il est ouvert (C-trap-LID) puis de rechercher les informations relatives aux ions fragments et aux ions précurseurs dans les bons groupes.

Conclusion générale

L'objectif de cette thèse était de contribuer à l'établissement et au développement d'une nouvelle méthodologie permettant d'identifier et de quantifier des biomarqueurs dans l'optique de détecter le plus précocement possible des altérations de l'homéostasie de l'organisme chez l'homme.

Dans le premier chapitre de ce manuscrit nous avons dressé un état de l'art relatif à cet objectif. Nous avons constaté que différentes familles de molécules biologiques pouvaient jouer le rôle de biomarqueurs dont les macromolécules protéines et que la spectrométrie de masse était un outil analytique adapté pour suivre leurs régulations. Il a été exposé que, les protéines kinases, de par leur implication dans une multitude de processus biologiques dont ceux régulant l'activité du cœur, du système cérébral ou encore intestinal, pouvait jouer un rôle de biomarqueurs putatifs pertinents et qu'elles étaient dosables par spectrométrie de masse. Nous avons donc décidé de nous concentrer sur la détection de ces protéines et d'examiner les différentes approches existantes permises par cette technique analytique pour les doser. Il a été montré que ces approches se scindaient principalement en 2 blocs, celles intitulées ciblées (SRM/PRM) et non ciblées types (DIA/DDA), et que si les approches ciblées permettaient une robustesse et une sensibilité supérieures aux approches non ciblées, elles n'étaient pas adaptées pour des approches globales de par leurs multiplexages limités. Il a dans le même temps été rapporté que l'utilisation du mode DIA permettait un multiplexage important (plusieurs milliers de peptides) et une fiabilité proche de la MRM rendant cette approche intéressante. Cependant, le traitement de données propre à ce mode est délicat et source d'erreurs en raison de l'utilisation d'une méthode d'activation non spécifique (CID). Pour remédier à cette difficulté, nous avons mis en lumière que l'utilisation d'une méthode d'activation reposant sur la photodissociation laser pouvait surmonter ce problème. En effet, en activant les molécules biologiques en fonction de leur propriété optique, il est possible d'obtenir une spécificité lors de l'étape de photo-fragmentation du fait que les composés biologiques absorbent à différentes longueurs d'ondes. En suivant ce chemin nous avons abouti à la conclusion que la méthode d'activation LID à 473 nm, en dérivant spécifiquement les peptides à cystéine avec un chromophore absorbant à cette longueur d'onde, était la plus adaptée pour répondre à cette problématique. En effet, l'acide aminé cystéine étant un acide aminé rare mais présent dans 89 % des protéines humaines, il est envisageable de couvrir tous le protéome humain et d'être spécifique en ne faisant photo-fragmenter que les peptides à cystéine dérivés compte-tenu du fait que les molécules biologiques n'absorbent pas naturellement à cette longueur d'onde.

Dans le second chapitre de ce manuscrit, il est présenté tout le développement relatif à l'établissement de la méthodologie DIA-LID à 473 nm pour identifier et quantifier des

biomarqueurs protéines kinases putatifs sur un spectromètre de masse haute résolution Qexactive. Dans cette optique, une bibliothèque de spectres de photo-fragmentation à 473 nm de 354 peptides à cystéine dérivés rapporteurs de 354 protéines kinases humaine a été construite à l'aide de spectres de photo-fragmentation expérimentaux acquis au laboratoire. Cette bibliothèque a ensuite été utilisée et a démontré son efficience pour identifier et quantifier de manière répétable les peptides à cystéine contenu en son sein. Cette approche a aussi démontré qu'elle permettait bien de diminuer fortement la complexité des spectres de (co)-fragmentation obtenus en DIA, par rapport à la HCD classique non discriminante, en raison de la spécificité de la méthode d'activation, remédiant ainsi au problème principal de ce mode. Cet effet a permis une plus grande facilité au niveau de l'extraction de l'information et une plus grande confidence lors de l'identification de composés. L'objectif semble donc être atteint. Cependant, si des protéines kinases endogènes ont pu être détectées lors de l'analyse d'un extrait cellulaire mammaire cancéreux humain, seulement très peu ont pu l'être. Cela peut être principalement attribué à la faible expression des protéines kinases humaines qui semblent requérir l'utilisation de kit d'enrichissement pour pouvoir être détectables en spectrométrie de masse. En vue d'être exploitée à son maximum, cette approche doit donc incorporer cette étape d'enrichissement. En parallèle de ce développement, la bibliothèque de spectres devra continuer à être enrichie en vue d'étendre son exhaustivité (cela a déjà commencé à être fait via l'ajout de 47 peptides à cystéine rapporteur de 33 protéines humaines).

Dans le troisième chapitre de ce manuscrit, nous nous sommes servis des 401 spectres de photo-fragmentation contenus dans la bibliothèque pour commencer à investiguer le modèle de photo-fragmentation des peptides à cystéine dérivés avec le DABCYL C2-maléimide en LID à 473 nm. Les premières observations ont été effectuées. Tout d'abord au sein des spectres de photo-fragmentation sont présents majoritairement des ions b et y ainsi qu'un ion rapporteur caractéristique issu de la fragmentation du chromophore. Il a été montré que la majorité des fragments sont générés au niveau des acides aminés de type aromatique ou hydrophobe aliphatique (surtout la proline) et que très peu de fragments sont produits au niveau de l'acide aminé cystéine, impliquant un processus de redistribution d'énergie le long de la chaine peptidique après absorption initiale au niveau du chromophore. Ce processus de redistribution s'est révélé avoir un comportement différent selon la taille de la séquence des peptides à cystéine dérivés photo-fragmentés. Il a également été montré que la position de la cystéine au sein de la séquence des peptides à cystéine dérivés photo-fragmentés avait une influence sur le pourcentage relatif des ions b et y produits. Il est supposé que la distribution des charges le long de la séquence des peptides à cystéine dérivés puisse être en lien avec cette observation. En vue de continuer l'éclaircissement des mécanismes impliqués dans ce modèle de fragmentation il serait pertinent de continuer à approfondir l'influence de la position de la cystéine et de voir s'il est bien la cause des résultats observés ou si cette observation n'est qu'une contingence d'un

autre phénomène. Dans le même temps il serait également judicieux de prendre des contacts avec des théoriciens pouvant expliquer par le calcul les observations faites et à venir.

Dans le quatrième chapitre de ce manuscrit, nous avons introduit une nouvelle méthodologie permettant de réaliser l'étape de photo-fragmentation directement au sein de la C-trap du Qexactive et non plus dans la cellule HCD. Intitulé C-trap-LID, nous avons démontré l'efficacité de ce concept pour faire photo-fragmenter de manière optimale et identifier des peptides à cystéine dérivés avec le DABCYL C2-maléimide à 473 nm. Les apports de cette méthodologie comparativement à une approche classique (fragmentation dans la cellule HCD) se traduisent par la possibilité de discriminer à l'œil nu, de par l'aspect de son XIC, si un ion précurseur a photo-fragmenté ou non, la capacité à extraire une liste exhaustive des *m/z* correspondant à tous les peptides à cystéine dérivés au sein d'un échantillon par une approche différentielle, et, dans certains cas, de peut-être gagner légèrement en sensibilité. Cependant, l'impossibilité de pouvoir contrôler séparément les gammes de masse relatives au quadripôle et à l'orbitrap sur le Qexactive ne rend pas ce concept applicable pour la DIA. Il est donc indispensable, en vue d'exploiter tout le potentiel de ce concept, d'avoir accès à ce paramètre. Il est également important de préciser que le traitement de données qui a été présenté dans ce quatrième chapitre a depuis été amélioré, permettant une extraction rapide de la liste des *m/z* correspondant à tous les peptides à cystéine dérivés au sein d'un échantillon lors de l'approche différentielle (et non plus scan par scan). Par ailleurs, il serait nécessaire d'avoir accès et de comprendre plus de signaux sortant du Qexcative en vue de multiplier les approches possibles.

D'une manière plus générale, il serait avantageux de se pencher sur la possibilité d'incorporer des diodes électroluminescentes (LED) émettant à 473 nm directement au sein de la cellule HCD (ou de la C-trap) du Qexactive[223]. Ce montage permettrait, une fois la synchronisation avec le mouvement des ions dans l'appareil établit et la démonstration de l'obtention d'un taux de photo-fragmentation optimal, de s'affranchir de toutes les contraintes relatives à l'utilisation d'un laser extérieur au couplage LC/MS/MS (contrôle du bon alignement du laser de manière quotidienne ou encore perforation de l'appareil Qexactive pour laisser entrer le laser). Dans l'optique d'utiliser cette méthodologie en routine dans des laboratoires d'analyses ou des hôpitaux, cela est presque indispensable. De plus, les LED pouvant couvrir une gamme de longueur d'onde allant de 350 à 2000 nm, il est possible de démultiplier les méthodes d'activations accessibles en fonction des LED utilisées (IRMPD, UVPD), voire d'incorporer plusieurs LED émettant à différentes longueurs d'ondes au sein d'un même appareil et ainsi user de différentes méthodes d'activations lors d'une même analyse. En parallèle, il serait intéressant, au laboratoire, d'utiliser la méthode d'activation LID pour cibler d'autres composés que des peptides (lipides, métabolites) en vue d'étendre l'usage de cette méthode d'activation à des domaines autres que la protéomique.

Bibliographie

(1) Ebhardt, H. A.; Root, A.; Sander, C.; Aebersold, R. Applications of Targeted Proteomics in Systems Biology and Translational Medicine. *Proteomics* **2015**, *15* (18), 3193–3208.

(2) Engvall, E.; Perlmann, P. Enzyme-Linked Immunosorbent Assay (ELISA). Quantitative Assay of Immunoglobulin G. *Immunochemistry* **1971**, *8* (9), 871–874.

(3) O'Farrell, P. H. High Resolution Two-Dimensional Electrophoresis of Proteins. *J. Biol. Chem.* **1975**, *250* (10), 4007–4021.

(4) HAINQUE, B.; BAUDIN, B.; LEFEBVRE, P. *Appareils et Méthodes En Biochimie et Biologie Moléculaire*; LAVOISIER MSP, Ed.; 2008.

(5) Mayne, J.; Ning, Z.; Zhang, X.; Starr, A. E.; Chen, R.; Deeke, S.; Chiang, C.-K.; Xu, B.; Wen, M.; Cheng, K.; Seebun, D.; Star, A.; Moore, J. I.; Figeys, D. Bottom-Up Proteomics (2013–2015): Keeping up in the Era of Systems Biology. *Anal. Chem.* **2016**, *88* (1), 95–121.

(6) Kuhn, E.; Addona, T.; Keshishian, H.; Burgess, M.; Mani, D. R.; Lee, R. T.; Sabatine, M. S.; Gerszten, R. E.; Carr, S. A. Developing Multiplexed Assays for Troponin I and Interleukin-33 in Plasma by Peptide Immunoaffinity Enrichment and Targeted Mass Spectrometry. *Clin. Chem.* **2009**, *55* (6), 1108–1117.

(7) Hoofnagle, A. N.; Becker, J. O.; Wener, M. H.; Heinecke, J. W. Quantification of Thyroglobulin, a Low-Abundance Serum Protein, by Immunoaffinity Peptide Enrichment and Tandem Mass Spectrometry. *Clin. Chem.* **2008**, *54* (11), 1796–1804.

(8) Tu, C.; Rudnick, P. A.; Martinez, M. Y.; Cheek, K. L.; Stein, S. E.; Slebos, R. J. C.; Liebler, D. C. Depletion of Abundant Plasma Proteins and Limitations of Plasma Proteomics. *J. Proteome Res.* **2010**, *9* (10), 4982–4991.

(9) Girod, M.; Biarc, J.; Enjalbert, Q.; Salvador, A.; Antoine, R.; Dugourd, P.; Lemoine, J. Implementing Visible 473 Nm Photodissociation in a Q-Exactive Mass Spectrometer: Towards Specific Detection of Cysteine-Containing Peptides. *Analyst* **2014**, *139* (21), 5523–5530.

(10) Reilly, J. P. Ultraviolet Photofragmentation of Biomolecular Ions. *Mass Spectrom. Rev.* **2009**, *28* (3), 425–447.

(11) Antoine, R.; Dugourd, P. Visible and Ultraviolet Spectroscopy of Gas Phase Protein Ions. *Phys. Chem. Chem. Phys.* **2011**, *13* (37), 16494–16509.

(12) Williams, E. R.; Furlong, J. J. P.; McLafferty, F. W. Efficiency of Collisionally-Activated Dissociation and 193-Nm Photodissociation of Peptide Ions in Fourier Transform Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **1990**, *1* (4), 288–294.

(13) Brodbelt, J. S. Photodissociation Mass Spectrometry: New Tools for Characterization of Biological Molecules. *Chem. Soc. Rev.* **2014**, *43* (8), 2757–2783.

(14) Enjalbert, Q.; Girod, M.; Simon, R.; Jeudy, J.; Chirot, F.; Salvador, A.; Antoine, R.; Dugourd, P.; Lemoine, J. Improved Detection Specificity for Plasma Proteins by Targeting Cysteine-Containing Peptides with Photo-SRM. *Anal. Bioanal. Chem.* **2013**, *405* (7), 2321–2331.

(15) Wang, S.; Zhang, X.; Regnier, F. E. Quantitative Proteomics Strategy Involving the Selection of Peptides Containing Both Cysteine and Histidine from Tryptic Digests of Cell Lysates. *J. Chromatogr. A* **2002**, *949* (1–2), 153–162.

(16) Strimbu, K.; Tavel, J. A. What Are Biomarkers? *Curr. Opin. HIV AIDS* **2010**, *5* (6), 463–466.

(17) Biomarkers and Surrogate Endpoints: Preferred Definitions and Conceptual Framework. *Clin. Pharmacol. Ther.* **2001**, *69* (3), 89–95.

(18) Tasevska, N. Urinary Sugars—A Biomarker of Total Sugars Intake. *Nutrients* **2015**, *7* (7), 5816–5833.

(19) Zhao, Y.-Y.; Cheng, X.; Lin, R.-C. Lipidomics Applications for Discovering Biomarkers of Diseases in Clinical Chemistry. *Int. Rev. Cell Mol. Biol.* **2014**, *313*, 1–26.

(20) Batchelor, E.; Loewer, A.; Mock, C.; Lahav, G. Stimulus-Dependent Dynamics of p53 in Single Cells. *Mol. Syst. Biol.* **2011**, *7*, 488.

(21) PAULING, L.; ITANO, H. A. Sickle Cell Anemia a Molecular Disease. *Science* **1949**, *110* (2865), 543–548.

(22) ABELEV, G. I.; PEROVA, S. D.; KHRAMKOVA, N. I.; POSTNIKOVA, Z. A.; IRLIN, I. S. Production of Embryonal Alpha-Globulin by Transplantable Mouse Hepatomas. *Transplantation* **1963**, *1*, 174–180.

(23) Bast Jr, R. C.; Feeney, M.; Lazarus, H.; Nadler, L. M.; Colvin, R. B.; Knapp, R. C. Reactivity of a Monoclonal Antibody with Human Ovarian Carcinoma. *J. Clin. Invest.* **1981**, *68* (5), 1331–1337.

(24) McDermott, J. E.; Wang, J.; Mitchell, H.; Webb-Robertson, B.-J.; Hafen, R.; Ramey, J.; Rodland, K. D. Challenges in Biomarker Discovery: Combining Expert Insights with Statistical Analysis of Complex Omics Data. *Expert Opin. Med. Diagn.* **2013**, *7* (1), 37–51.

(25) Marks, D. S.; Colwell, L. J.; Sheridan, R.; Hopf, T. A.; Pagnani, A.; Zecchina, R.; Sander, C. Protein 3D Structure Computed from Evolutionary Sequence Variation. *PLoS One* **2011**, *6* (12), e28766.

(26) Towbin, H.; Staehelin, T.; Gordon, J. Electrophoretic Transfer of Proteins from Polyacrylamide Gels to Nitrocellulose Sheets: Procedure and Some Applications. *Proc. Natl. Acad. Sci.* **1979**, *76* (9), 4350–4354.

(27) Leng, S. X.; McElhaney, J. E.; Walston, J. D.; Xie, D.; Fedarko, N. S.; Kuchel, G. A. ELISA and Multiplex Technologies for Cytokine Measurement in Inflammation and Aging Research. *J. Gerontol. A. Biol. Sci. Med. Sci.* **2008**, *63* (8), 879–884.

(28) Rifai, N.; Gillette, M. A.; Carr, S. A. Protein Biomarker Discovery and Validation: The Long and Uncertain Path to Clinical Utility. *Nat Biotech* **2006**, *24* (8), 971–983.

(29) Cummings, J.; Ward, T. H.; Greystoke, A.; Ranson, M.; Dive, C. Biomarker Method Validation in Anticancer Drug Development. *Br. J. Pharmacol.* **2008**, *153* (4), 646–656.

(30) Levene, P. A.; Alsberg, C. L. THE CLEAVAGE PRODUCTS OF VITELLIN. *J. Biol. Chem.* **1906**, *2* (1), 127–133.

(31) BURNETT, G.; KENNEDY, E. P. The Enzymatic Phosphorylation of Proteins. *J. Biol. Chem.* **1954**, *211* (2), 969–980.

(32) Université d'Angers http://biochimej.univ-angers.fr/Page2/COURS/7RelStructFonction/2Biochimie/2ModifPOSTtraduc/4Phosphorylation/1PhosPhoRylation.htm.

(33) Cohen, P. The Origins of Protein Phosphorylation. *Nat Cell Biol* **2002**, *4* (5), E127–E130.

(34) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The Protein Kinase Complement of the Human Genome. *Science* **2002**, *298* (5600), 1912–1934.

(35) KREBS, E. G.; FISCHER, E. H. The Phosphorylase B to a Converting Enzyme of Rabbit Skeletal Muscle. *Biochim. Biophys. Acta* **1956**, *20* (1), 150–157.

(36) Crosby, M. E. Cell Cycle: Principles of Control. *Yale J. Biol. Med.* **2007**, *80* (3), 141–142.

(37) Rodriguez, P.; Kranias, E. G. Phospholamban: A Key Determinant of Cardiac Function and Dysfunction. *Arch. Mal. Coeur Vaiss.* **2005**, *98* (12), 1239–1243.

(38) Hofmann, F.; Ammendola, A.; Schlossmann, J. Rising behind NO: cGMP-Dependent Protein Kinases. *J. Cell Sci.* **2000**, *113 (Pt 1*, 1671–1676.

(39) Eide, E. J.; Virshup, D. M. Casein Kinase I: Another Cog in the Circadian Clockworks. *Chronobiol. Int.* **2001**, *18* (3), 389–398.

(40) Wang, H.; Brown, J.; Martin, M. Glycogen Synthase Kinase 3: A Point of Convergence for the Host Inflammatory Response. *Cytokine* **2011**, *53* (2), 130–140.

(41) Olvedy, M.; Tisserand, J. C.; Luciani, F.; Boeckx, B.; Wouters, J.; Lopez, S.; Rambow, F.; Aibar, S.; Thienpont, B.; Barra, J.; Köhler, C.; Radaelli, E.; Tartare-Deckert, S.; Aerts, S.; Dubreuil, P.; van den Oord, J. J.; Lambrechts, D.; De Sepulveda, P.; Marine, J.-C. Comparative Oncogenomics Identifies Tyrosine Kinase FES as a Tumor Suppressor in Melanoma. *J. Clin. Invest.* **2017**, *127* (6), 2310–2325.

(42) Chico, L. K.; Van Eldik, L. J.; Watterson, D. M. Targeting Protein Kinases in Central Nervous

System Disorders. *Nat Rev Drug Discov* **2009**, *8* (11), 892–909.

(43) Johnson, L. N.; Noble, M. E. M.; Owen, D. J. Active and Inactive Protein Kinases: Structural Basis for Regulation. *Cell* **1996**, *85* (2), 149–158.

(44) Adams, J. A. Kinetic and Catalytic Mechanisms of Protein Kinases. *Chem. Rev.* **2001**, *101* (8), 2271–2290.

(45) Taylor, S. S.; Radzio-Andzelm, E. Protein Kinase Inhibition: Natural and Synthetic Variations on a Theme. *Curr. Opin. Chem. Biol.* **1997**, *1* (2), 219–226.

(46) Nolen, B.; Taylor, S.; Ghosh, G. Regulation of Protein Kinases: Controlling Activity through Activation Segment Conformation. *Mol. Cell* **2004**, *15* (5), 661–675.

(47) Dhillon, A. S.; Hagan, S.; Rath, O.; Kolch, W. MAP Kinase Signalling Pathways in Cancer. *Oncogene 26* (22), 3279–3290.

(48) Santarius, T.; Shipley, J.; Brewer, D.; Stratton, M. R.; Cooper, C. S. A Census of Amplified and Overexpressed Human Cancer Genes. *Nat Rev Cancer* **2010**, *10* (1), 59–64.

(49) Liegl-Atzwanger, B.; Fletcher, J. A.; Fletcher, C. D. M. Gastrointestinal Stromal Tumors. *Virchows Arch.* **2010**, *456* (2), 111–127.

(50) Bennett, B. L. C-Jun N-Terminal Kinase-Dependent Mechanisms in Respiratory Disease. *Eur. Respir. J.* **2006**, *28* (3), 651 LP-661.

(51) Wang, Y. Mitogen-Activated Protein Kinases in Heart Development and Diseases. *Circulation* **2007**, *116* (12), 1413–1423.

(52) Pandya, N.; Santani, D.; Jain, S. Role of Mitogen-Activated Protein (MAP) Kinases in Cardiovascular Diseases. *Cardiovasc. Drug Rev.* **2005**, *23* (3), 247–254.

(53) Greenman, C.; Stephens, P.; Smith, R.; Dalgliesh, G. L.; Hunter, C.; Bignell, G.; Davies, H.; Teague, J.; Butler, A.; Stevens, C.; Edkins, S.; O'Meara, S.; Vastrik, I.; Schmidt, E. E.; Avis, T.; Barthorpe, S.; Bhamra, G.; Buck, G.; Choudhury, B.; Clements, J.; Cole, J.; Dicks, E.; Forbes, S.; Gray, K.; Halliday, K.; Harrison, R.; Hills, K.; Hinton, J.; Jenkinson, A.; Jones, D.; Menzies, A.; Mironenko, T.; Perry, J.; Raine, K.; Richardson, D.; Shepherd, R.; Small, A.; Tofts, C.; Varian, J.; Webb, T.; West, S.; Widaa, S.; Yates, A.; Cahill, D. P.; Louis, D. N.; Goldstraw, P.; Nicholson, A. G.; Brasseur, F.; Looijenga, L.; Weber, B. L.; Chiew, Y.-E.; deFazio, A.; Greaves, M. F.; Green, A. R.; Campbell, P.; Birney, E.; Easton, D. F.; Chenevix-Trench, G.; Tan, M.-H.; Khoo, S. K.; Teh, B. T.; Yuen, S. T.; Leung, S. Y.; Wooster, R.; Futreal, P. A.; Stratton, M. R. Patterns of Somatic Mutation in Human Cancer Genomes. *Nature* **2007**, *446* (7132), 153–158.

(54) Sevecka, M.; MacBeath, G. State-Based Discovery: A Multidimensional Screen for Small-Molecule Modulators of EGF Signaling. *Nat Meth* **2006**, *3* (10), 825–831.

(55) Daub, H.; Specht, K.; Ullrich, A. Strategies to Overcome Resistance to Targeted Protein

Kinase Inhibitors. *Nat Rev Drug Discov* **2004**, *3* (12), 1001–1010.

(56) Inoue, A.; Saijo, Y.; Maemondo, M.; Gomi, K.; Tokue, Y.; Kimura, Y.; Ebina, M.; Kikuchi, T.; Moriya, T.; Nukiwa, T. Severe Acute Interstitial Pneumonia and Gefitinib. *Lancet* **2003**, *361* (9352), 137–139.

(57) McAllister, F. E.; Niepel, M.; Haas, W.; Huttlin, E.; Sorger, P. K.; Gygi, S. P. Mass Spectrometry Based Method to Increase Throughput for Kinome Analyses Using ATP Probes. *Anal. Chem.* **2013**, *85* (9), 4666–4674.

(58) Patricelli, M. P.; Nomanbhoy, T. K.; Wu, J.; Brown, H.; Zhou, D.; Zhang, J.; Jagannathan, S.; Aban, A.; Okerberg, E.; Herring, C.; Nordin, B.; Weissig, H.; Yang, Q.; Lee, J.-D.; Gray, N. S.; Kozarich, J. W. In Situ Kinase Profiling Reveals Functionally Relevant Properties of Native Kinases. *Chem. Biol.* **2011**, *18* (6), 699–710.

(59) Zhang, L.; Holmes, I. P.; Hochgräfe, F.; Walker, S. R.; Ali, N. A.; Humphrey, E. S.; Wu, J.; de Silva, M.; Kersten, W. J. A.; Connor, T.; Falk, H.; Allan, L.; Street, I. P.; Bentley, J. D.; Pilling, P. A.; Monahan, B. J.; Peat, T. S.; Daly, R. J. Characterization of the Novel Broad-Spectrum Kinase Inhibitor CTx-0294885 As an Affinity Reagent for Mass Spectrometry-Based Kinome Profiling. *J. Proteome Res.* **2013**, *12* (7), 3104–3116.

(60) Stensballe, A.; Andersen, S.; Jensen, O. N. Characterization of Phosphoproteins from Electrophoretic Gels by Nanoscale Fe(III) Affinity Chromatography with off-Line Mass Spectrometry Analysis. *Proteomics* **2001**, *1* (2), 207–222.

(61) Kokubu, M.; Ishihama, Y.; Sato, T.; Nagasu, T.; Oda, Y. Specificity of Immobilized Metal Affinity-Based IMAC/C18 Tip Enrichment of Phosphopeptides for Protein Phosphorylation Analysis. *Anal. Chem.* **2005**, *77* (16), 5144–5154.

(62) Mazanek, M.; Mituloviæ, G.; Herzog, F.; Stingl, C.; Hutchins, J. R. A.; Peters, J.-M.; Mechtler, K. Titanium Dioxide as a Chemo-Affinity Solid Phase in Offline Phosphopeptide Chromatography prior to HPLC-MS/MS Analysis. *Nat. Protoc.* **2007**, *2* (5), 1059–1069.

(63) Sugiyama, N.; Masuda, T.; Shinoda, K.; Nakamura, A.; Tomita, M.; Ishihama, Y. Phosphopeptide Enrichment by Aliphatic Hydroxy Acid-Modified Metal Oxide Chromatography for Nano-LC-MS/MS in Proteomics Applications. *Mol. Cell. Proteomics* **2007**, *6* (6), 1103–1109.

(64) Rush, J.; Moritz, A.; Lee, K. A.; Guo, A.; Goss, V. L.; Spek, E. J.; Zhang, H.; Zha, X.-M.; Polakiewicz, R. D.; Comb, M. J. Immunoaffinity Profiling of Tyrosine Phosphorylation in Cancer Cells. *Nat Biotech* **2005**, *23* (1), 94–101.

(65) Julien, R. Intérêts Des Marqueurs Bilogiques Dans Les Essais Cliniques, Henry Poincaré - Nancy 1, 2011.

(66) Anderson, N. L.; Anderson, N. G. The Human Plasma Proteome: History, Character, and Diagnostic Prospects . *Mol. Cell. Proteomics* **2002**, *1* (11), 845–867.

(67) Graumann, J.; Hubner, N. C.; Kim, J. B.; Ko, K.; Moser, M.; Kumar, C.; Cox, J.; Schöler, H.; Mann, M. Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) and Proteome Quantitation of Mouse Embryonic Stem Cells to a Depth of 5,111 Proteins. *Mol. Cell. Proteomics* **2008**, *7* (4), 672–683.

(68) Chen, E. I.; McClatchy, D.; Park, S. K.; Yates, J. R. 3rd. Comparisons of Mass Spectrometry Compatible Surfactants for Global Analysis of the Mammalian Brain Proteome. *Anal. Chem.* **2008**, *80* (22), 8694–8701.

(69) Brodbelt, J. S. Ion Activation Methods for Peptides and Proteins. *Anal. Chem.* **2016**, *88* (1), 30–51.

(70) Kelstrup, C. D.; Young, C.; Lavallee, R.; Nielsen, M. L.; Olsen, J. V. Optimized Fast and Sensitive Acquisition Methods for Shotgun Proteomics on a Quadrupole Orbitrap Mass Spectrometer. *J. Proteome Res.* **2012**, *11* (6), 3487–3497.

(71) Kelleher, N. L. Top-down Proteomics. *Anal. Chem.* **2004**, *76* (11), 197A–203A.

(72) Cristobal, A.; Marino, F.; Post, H.; van den Toorn, H. W. P.; Mohammed, S.; Heck, A. J. R. Toward an Optimized Workflow for Middle-Down Proteomics. *Anal. Chem.* **2017**, *89* (6), 3318–3325.

(73) Smith, L. M.; Kelleher, N. L. Proteoform: A Single Term Describing Protein Complexity. *Nature methods*. United States March 2013, pp 186–187.

(74) Catherman, A. D.; Skinner, O. S.; Kelleher, N. L. Top Down Proteomics: Facts and Perspectives. *Biochem. Biophys. Res. Commun.* **2014**, *445* (4), 683–693.

(75) Worboys, J. D.; Sinclair, J.; Yuan, Y.; Jorgensen, C. Systematic Evaluation of Quantotypic Peptides for Targeted Analysis of the Human Kinome. *Nat Meth* **2014**, *11* (10), 1041–1044.

(76) Chait, B. T. Chemistry. Mass Spectrometry: Bottom-up or Top-Down? *Science* **2006**, *314* (5796), 65–66.

(77) Hu, A.; Noble, W. S.; Wolf-Yadlin, A. Technical Advances in Proteomics: New Developments in Data-Independent Acquisition. *F1000Research* **2016**, *5*.

(78) Addona, T. A.; Abbatiello, S. E.; Schilling, B.; Skates, S. J.; Mani, D. R.; Bunk, D. M.; Spiegelman, C. H.; Zimmerman, L. J.; Ham, A.-J. L.; Keshishian, H.; Hall, S. C.; Allen, S.; Blackman, R. K.; Borchers, C. H.; Buck, C.; Cardasis, H. L.; Cusack, M. P.; Dodder, N. G.; Gibson, B. W.; Held, J. M.; Hiltke, T.; Jackson, A.; Johansen, E. B.; Kinsinger, C. R.; Li, J.; Mesri, M.; Neubert, T. A.; Niles, R. K.; Pulsipher, T. C.; Ransohoff, D.; Rodriguez, H.; Rudnick, P. A.; Smith, D.; Tabb, D. L.; Tegeler, T. J.; Variyath, A. M.; Vega-Montoto, L. J.; Wahlander, A.; Waldemarson, S.; Wang, M.; Whiteaker, J. R.; Zhao, L.; Anderson, N. L.; Fisher, S. J.; Liebler, D. C.; Paulovich, A. G.; Regnier, F. E.; Tempst, P.; Carr, S. A. Multi-Site Assessment of the Precision and Reproducibility of Multiple Reaction Monitoring-Based

Measurements of Proteins in Plasma. *Nat Biotech* **2009**, *27* (7), 633–641.

(79) Kennedy, J. J.; Abbatiello, S. E.; Kim, K.; Yan, P.; Whiteaker, J. R.; Lin, C.; Kim, J. S.; Zhang, Y.; Wang, X.; Ivey, R. G.; Zhao, L.; Min, H.; Lee, Y.; Yu, M.-H.; Yang, E. G.; Lee, C.; Wang, P.; Rodriguez, H.; Kim, Y.; Carr, S. A.; Paulovich, A. G. Demonstrating the Feasibility of Large-Scale Development of Standardized Assays to Quantify Human Proteins. *Nat Meth* **2014**, *11* (2), 149–155.

(80) Kuhn, E.; Whiteaker, J. R.; Mani, D. R.; Jackson, A. M.; Zhao, L.; Pope, M. E.; Smith, D.; Rivera, K. D.; Anderson, N. L.; Skates, S. J.; Pearson, T. W.; Paulovich, A. G.; Carr, S. A. Interlaboratory Evaluation of Automated, Multiplexed Peptide Immunoaffinity Enrichment Coupled to Multiple Reaction Monitoring Mass Spectrometry for Quantifying Proteins in Plasma. *Mol. Cell. Proteomics* **2012**, *11* (6).

(81) Stergachis, A. B.; MacLean, B.; Lee, K.; Stamatoyannopoulos, J. A.; MacCoss, M. J. Rapid Empirical Discovery of Optimal Peptides for Targeted Proteomics. *Nat Meth* **2011**, *8* (12), 1041–1043.

(82) Liebler, D. C.; Zimmerman, L. J. Targeted Quantitation of Proteins by Mass Spectrometry. *Biochemistry* **2013**, *52* (22), 3797–3806.

(83) MRM Atlas http://www.mrmatlas.org/mrmassay.php.

(84) Vidova, V.; Spacil, Z. A Review on Mass Spectrometry-Based Quantitative Proteomics: Targeted and Data Independent Acquisition. *Anal. Chim. Acta* **2017**, *964*, 7–23.

(85) Shi, T.; Song, E.; Nie, S.; Rodland, K. D.; Liu, T.; Qian, W.-J.; Smith, R. D. Advances in Targeted Proteomics and Applications to Biomedical Research. *Proteomics* **2016**, *16* (15–16), 2160–2182.

(86) Rougemont, B.; Bontemps Gallo, S.; Ayciriex, S.; Carrière, R.; Hondermarck, H.; Lacroix, J. M.; Le Blanc, J. C. Y.; Lemoine, J. Scout-MRM: Multiplexed Targeted Mass Spectrometry-Based Assay without Retention Time Scheduling Exemplified by Dickeya Dadantii Proteomic Analysis during Plant Infection. *Anal. Chem.* **2017**, *89* (3), 1421–1426.

(87) Mueller, C. A.; Weinmann, W.; Dresen, S.; Schreiber, A.; Gergov, M. Development of a Multi-Target Screening Analysis for 301 Drugs Using a QTrap Liquid Chromatography/tandem Mass Spectrometry System and Automated Library Searching. *Rapid Commun. Mass Spectrom.* **2005**, *19* (10), 1332–1338.

(88) Wei, R.; Li, G.; Seymour, A. B. High-Throughput and Multiplexed LC/MS/MRM Method for Targeted Metabolomics. *Anal. Chem.* **2010**, *82* (13), 5527–5533.

(89) Lange, V.; Picotti, P.; Domon, B.; Aebersold, R. Selected Reaction Monitoring for Quantitative Proteomics: A Tutorial. *Mol. Syst. Biol.* **2008**, *4*, 222.

(90) Selevsek, N.; Matondo, M.; Sanchez Carbayo, M.; Aebersold, R.; Domon, B. Systematic Quantification of Peptides/proteins in Urine Using Selected Reaction Monitoring.

Proteomics **2011**, *11* (6), 1135–1147.

(91) Hoofnagle, A. N.; Becker, J. O.; Oda, M. N.; Cavigiolio, G.; Mayer, P.; Vaisar, T. Multiple-Reaction Monitoring–Mass Spectrometric Assays Can Accurately Measure the Relative Protein Abundance in Complex Mixtures. *Clin. Chem.* **2012**, *58* (4), 777 LP-781.

(92) Peterson, A. C.; Russell, J. D.; Bailey, D. J.; Westphall, M. S.; Coon, J. J. Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics. *Mol. Cell. Proteomics* **2012**, *11* (11), 1475–1488.

(93) Gallien, S.; Duriez, E.; Demeure, K.; Domon, B. Selectivity of LC-MS/MS Analysis: Implication for Proteomics Experiments. *J. Proteomics* **2013**, *81*, 148–158.

(94) Ronsein, G. E.; Pamir, N.; von Haller, P. D.; Kim, D. S.; Oda, M. N.; Jarvik, G. P.; Vaisar, T.; Heinecke, J. W. Parallel Reaction Monitoring (PRM) and Selected Reaction Monitoring (SRM) Exhibit Comparable Linearity, Dynamic Range and Precision for Targeted Quantitative HDL Proteomics. *J. Proteomics* **2015**, *113*, 388–399.

(95) Majovsky, P.; Naumann, C.; Lee, C.-W.; Lassowskat, I.; Trujillo, M.; Dissmeyer, N.; Hoehenwarter, W. Targeted Proteomics Analysis of Protein Degradation in Plant Signaling on an LTQ-Orbitrap Mass Spectrometer. *J. Proteome Res.* **2014**, *13* (10), 4246–4258.

(96) Morin, L.-P.; Mess, J.-N.; Garofolo, F. Large-Molecule Quantification: Sensitivity and Selectivity Head-to-Head Comparison of Triple Quadrupole with Q-TOF. *Bioanalysis* **2013**, *5* (10), 1181–1193.

(97) Rauniyar, N. Parallel Reaction Monitoring: A Targeted Experiment Performed Using High Resolution and High Mass Accuracy Mass Spectrometry. *Int. J. Mol. Sci.* **2015**, *16* (12), 28566–28581.

(98) Aebersold, R.; Mann, M. Mass Spectrometry-Based Proteomics. *Nature* **2003**, *422* (6928), 198–207.

(99) Eng, J. K.; Fischer, B.; Grossmann, J.; Maccoss, M. J. A Fast SEQUEST Cross Correlation Algorithm. *J. Proteome Res.* **2008**, *7* (10), 4598–4602.

(100) Koenig, T.; Menze, B. H.; Kirchner, M.; Monigatti, F.; Parker, K. C.; Patterson, T.; Steen, J. J.; Hamprecht, F. A.; Steen, H. Robust Prediction of the MASCOT Score for an Improved Quality Assessment in Mass Spectrometric Proteomics. *J. Proteome Res.* **2008**, *7* (9), 3708–3717.

(101) Cox, J.; Mann, M. MaxQuant Enables High Peptide Identification Rates, Individualized P.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367–1372.

(102) Nagaraj, N.; Kulak, N. A.; Cox, J.; Neuhauser, N.; Mayr, K.; Hoerning, O.; Vorm, O.; Mann, M. System-Wide Perturbation Analysis with Nearly Complete Coverage of the Yeast Proteome by Single-Shot Ultra HPLC Runs on a Bench Top Orbitrap. *Mol. Cell. Proteomics*

2012, *11* (3), M111.013722.

(103) Ong, S.-E.; Mann, M. A Practical Recipe for Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC). *Nat. Protoc.* **2006**, *1* (6), 2650–2660.

(104) Wiese, S.; Reidegeld, K. A.; Meyer, H. E.; Warscheid, B. Protein Labeling by iTRAQ: A New Tool for Quantitative Mass Spectrometry in Proteome Research. *Proteomics* **2007**, *7* (3), 340–350.

(105) Argenzio, E.; Bange, T.; Oldrini, B.; Bianchi, F.; Peesari, R.; Mari, S.; Di Fiore, P. P.; Mann, M.; Polo, S. Proteomic Snapshot of the EGF-Induced Ubiquitin Network. *Mol. Syst. Biol.* **2011**, *7*, 462.

(106) Villén, J.; Beausoleil, S. A.; Gerber, S. A.; Gygi, S. P. Large-Scale Phosphorylation Analysis of Mouse Liver. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (5), 1488–1493.

(107) Gavin, A.-C.; Maeda, K.; Kuhner, S. Recent Advances in Charting Protein-Protein Interaction: Mass Spectrometry-Based Approaches. *Curr. Opin. Biotechnol.* **2011**, *22* (1), 42–49.

(108) Cristea, I. M.; Graham, D. Virology Meets Proteomics. *Proteomics*. Germany June 2015, pp 1941–1942.

(109) Venable, J. D.; Dong, M.-Q.; Wohlschlegel, J.; Dillin, A.; Yates, J. R. Automated Approach for Quantitative Analysis of Complex Peptide Mixtures from Tandem Mass Spectra. *Nat. Methods* **2004**, *1* (1), 39–45.

(110) Tabb, D. L.; Vega-Montoto, L.; Rudnick, P. A.; Variyath, A. M.; Ham, A.-J. L.; Bunk, D. M.; Kilpatrick, L. E.; Billheimer, D. D.; Blackman, R. K.; Cardasis, H. L.; Carr, S. A.; Clauser, K. R.; Jaffe, J. D.; Kowalski, K. A.; Neubert, T. A.; Regnier, F. E.; Schilling, B.; Tegeler, T. J.; Wang, M.; Wang, P.; Whiteaker, J. R.; Zimmerman, L. J.; Fisher, S. J.; Gibson, B. W.; Kinsinger, C. R.; Mesri, M.; Rodriguez, H.; Stein, S. E.; Tempst, P.; Paulovich, A. G.; Liebler, D. C.; Spiegelman, C. Repeatability and Reproducibility in Proteomic Identifications by Liquid Chromatography-Tandem Mass Spectrometry. *J. Proteome Res.* **2010**, *9* (2), 761–776.

(111) Purvine, S.; Eppel*, J.-T.; Yi, E. C.; Goodlett, D. R. Shotgun Collision-Induced Dissociation of Peptides Using a Time of Flight Mass Analyzer. *Proteomics* **2003**, *3* (6), 847–850.

(112) Liu, Y.; Huttenhain, R.; Collins, B.; Aebersold, R. Mass Spectrometric Protein Maps for Biomarker Discovery and Clinical Research. *Expert Rev. Mol. Diagn.* **2013**, *13* (8), 811–825.

(113) Sajic, T.; Liu, Y.; Aebersold, R. Using Data-Independent, High-Resolution Mass Spectrometry in Protein Biomarker Research: Perspectives and Clinical Applications. *Proteomics. Clin. Appl.* **2015**, *9* (3–4), 307–321.

(114) Tsou, C.-C.; Avtonomov, D.; Larsen, B.; Tucholska, M.; Choi, H.; Gingras, A.-C.; Nesvizhskii, A. I. DIA-Umpire: Comprehensive Computational Framework for Data-Independent Acquisition Proteomics. *Nat. Methods* **2015**, *12* (3), 258–64, 7 p following 264.

(115) Bern, M.; Finney, G.; Hoopmann, M. R.; Merrihew, G.; Toth, M. J.; MacCoss, M. J. Deconvolution of Mixture Spectra from Ion-Trap Data-Independent-Acquisition Tandem Mass Spectrometry. *Anal. Chem.* **2010**, *82* (3), 833–841.

(116) Ting, Y. S.; Egertson, J. D.; Payne, S. H.; Kim, S.; MacLean, B.; Kall, L.; Aebersold, R.; Smith, R. D.; Noble, W. S.; MacCoss, M. J. Peptide-Centric Proteome Analysis: An Alternative Strategy for the Analysis of Tandem Mass Spectrometry Data. *Mol. Cell. Proteomics* **2015**, *14* (9), 2301–2307.

(117) Gillet, L. C.; Navarro, P.; Tate, S.; Rost, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R. Targeted Data Extraction of the MS/MS Spectra Generated by Data-Independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Mol. Cell. Proteomics* **2012**, *11* (6), O111.016717.

(118) Rost, H. L.; Rosenberger, G.; Navarro, P.; Gillet, L.; Miladinovic, S. M.; Schubert, O. T.; Wolski, W.; Collins, B. C.; Malmstrom, J.; Malmstrom, L.; Aebersold, R. OpenSWATH Enables Automated, Targeted Analysis of Data-Independent Acquisition MS Data. *Nature biotechnology*. United States March 2014, pp 219–223.

(119) Bruderer, R.; Bernhardt, O. M.; Gandhi, T.; Miladinovic, S. M.; Cheng, L.-Y.; Messner, S.; Ehrenberger, T.; Zanotelli, V.; Butscheid, Y.; Escher, C.; Vitek, O.; Rinner, O.; Reiter, L. Extending the Limits of Quantitative Proteome Profiling with Data-Independent Acquisition and Application to Acetaminophen-Treated Three-Dimensional Liver Microtissues. *Mol. Cell. Proteomics* **2015**, *14* (5), 1400–1410.

(120) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: An Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments. *Bioinformatics* **2010**, *26* (7), 966–968.

(121) Silva, J. C.; Denny, R.; Dorschel, C. A.; Gorenstein, M.; Kass, I. J.; Li, G.-Z.; McKenna, T.; Nold, M. J.; Richardson, K.; Young, P.; Geromanos, S. Quantitative Proteomic Analysis by Accurate Mass Retention Time Pairs. *Anal. Chem.* **2005**, *77* (7), 2187–2200.

(122) Silva, J. C.; Gorenstein, M. V; Li, G.-Z.; Vissers, J. P. C.; Geromanos, S. J. Absolute Quantification of Proteins by LCMSE: A Virtue of Parallel MS Acquisition. *Mol. Cell. Proteomics* **2006**, *5* (1), 144–156.

(123) Silva, J. C.; Denny, R.; Dorschel, C.; Gorenstein, M. V; Li, G.-Z.; Richardson, K.; Wall, D.; Geromanos, S. J. Simultaneous Qualitative and Quantitative Analysis of the Escherichia Coli Proteome: A Sweet Tale. *Mol. Cell. Proteomics* **2006**, *5* (4), 589–607.

(124) Geiger, T.; Cox, J.; Mann, M. Proteomics on an Orbitrap Benchtop Mass Spectrometer Using All-Ion Fragmentation. *Mol. Cell. Proteomics* **2010**, *9* (10), 2252–2261.

(125) Weisbrod, C. R.; Eng, J. K.; Hoopmann, M. R.; Baker, T.; Bruce, J. E. Accurate Peptide Fragment Mass Analysis: Multiplexed Peptide Identification and Quantification. *J.*

Proteome Res. **2012**, *11* (3), 1621–1632.

(126) Egertson, J. D.; Kuehn, A.; Merrihew, G. E.; Bateman, N. W.; MacLean, B. X.; Ting, Y. S.; Canterbury, J. D.; Marsh, D. M.; Kellmann, M.; Zabrouskov, V.; Wu, C. C.; MacCoss, M. J. Multiplexed MS/MS for Improved Data-Independent Acquisition. *Nat. Methods* **2013**, *10* (8), 744–746.

(127) Prakash, A.; Peterman, S.; Ahmad, S.; Sarracino, D.; Frewen, B.; Vogelsang, M.; Byram, G.; Krastins, B.; Vadali, G.; Lopez, M. Hybrid Data Acquisition and Processing Strategies with Increased Throughput and Selectivity: pSMART Analysis for Global Qualitative and Quantitative Analysis. *J. Proteome Res.* **2014**, *13* (12), 5415–5430.

(128) Liu, Y.; Buil, A.; Collins, B. C.; Gillet, L. C. J.; Blum, L. C.; Cheng, L.-Y.; Vitek, O.; Mouritsen, J.; Lachance, G.; Spector, T. D.; Dermitzakis, E. T.; Aebersold, R. Quantitative Variability of 342 Plasma Proteins in a Human Twin Population. *Mol. Syst. Biol.* **2015**, *11* (1), 786.

(129) Lambert, J.-P.; Ivosev, G.; Couzens, A. L.; Larsen, B.; Taipale, M.; Lin, Z.-Y.; Zhong, Q.; Lindquist, S.; Vidal, M.; Aebersold, R.; Pawson, T.; Bonner, R.; Tate, S.; Gingras, A.-C. Mapping Differential Interactomes by Affinity Purification Coupled with Data-Independent Mass Spectrometry Acquisition. *Nat. Methods* **2013**, *10* (12), 1239–1245.

(130) Bilbao, A.; Varesio, E.; Luban, J.; Strambio-De-Castillia, C.; Hopfgartner, G.; Muller, M.; Lisacek, F. Processing Strategies and Software Solutions for Data-Independent Acquisition in Mass Spectrometry. *Proteomics* **2015**, *15* (5–6), 964–980.

(131) McIlwain, S.; Tamura, K.; Kertesz-Farkas, A.; Grant, C. E.; Diament, B.; Frewen, B.; Howbert, J. J.; Hoopmann, M. R.; Kall, L.; Eng, J. K.; MacCoss, M. J.; Noble, W. S. Crux: Rapid Open Source Protein Tandem Mass Spectrometry Analysis. *Journal of proteome research*. United States October 2014, pp 4488–4491.

(132) Deutsch, E. W.; Mendoza, L.; Shteynberg, D.; Farrah, T.; Lam, H.; Tasman, N.; Sun, Z.; Nilsson, E.; Pratt, B.; Prazen, B.; Eng, J. K.; Martin, D. B.; Nesvizhskii, A. I.; Aebersold, R. A Guided Tour of the Trans-Proteomic Pipeline. *Proteomics* **2010**, *10* (6), 1150–1159.

(133) Lam, H.; Aebersold, R. Building and Searching Tandem Mass (MS/MS) Spectral Libraries for Peptide Identification in Proteomics. *Methods* **2011**, *54* (4), 424–431.

(134) Aebersold, R. A Stress Test for Mass Spectrometry-Based Proteomics. *Nature methods*. United States June 2009, pp 411–412.

(135) Kryuchkov, F.; Verano-Braga, T.; Hansen, T. A.; Sprenger, R. R.; Kjeldsen, F. Deconvolution of Mixture Spectra and Increased Throughput of Peptide Identification by Utilization of Intensified Complementary Ions Formed in Tandem Mass Spectrometry. *J. Proteome Res.* **2013**, *12* (7), 3362–3371.

(136) Liu, Y.; Chen, J.; Sethi, A.; Li, Q. K.; Chen, L.; Collins, B.; Gillet, L. C. J.; Wollscheid, B.; Zhang, H.; Aebersold, R. Glycoproteomic Analysis of Prostate Cancer Tissues by SWATH

Mass Spectrometry Discovers N-Acylethanolamine Acid Amidase and Protein Tyrosine Kinase 7 as Signatures for Tumor Aggressiveness. *Mol. Cell. Proteomics* **2014**, *13* (7), 1753–1768.

(137) Parker, B. L.; Yang, G.; Humphrey, S. J.; Chaudhuri, R.; Ma, X.; Peterman, S.; James, D. E. Targeted Phosphoproteomics of Insulin Signaling Using Data-Independent Acquisition Mass Spectrometry. *Sci. Signal.* **2015**, *8* (380), rs6--rs6.

(138) Roepstorff, P.; Fohlman, J. Letter to the Editors. *Biol. Mass Spectrom.* **1984**, *11* (11), 601.

(139) Biemann, K. Contributions of Mass Spectrometry to Peptide and Protein Structure. *Biol. Mass Spectrom.* **1988**, *16* (1–12), 99–111.

(140) Waugh, R. J.; Bowie, J. H. A Review of the Collision-Induced Dissociations of Deprotonated Dipeptides and Tripeptides. An Aid to Structure Determination. *Rapid Commun. Mass Spectrom.* **1994**, *8* (2), 169–173.

(141) Bowie, J. H.; Brinkworth, C. S.; Dua, S. Collision-Induced Fragmentations of the (M-H)− Parent Anions of Underivatized Peptides: An Aid to Structure Determination and Some Unusual Negative Ion Cleavages. *Mass Spectrom. Rev.* **2002**, *21* (2), 87–107.

(142) Paizs, B.; Suhai, S. Fragmentation Pathways of Protonated Peptides. *Mass Spectrom. Rev.* **2005**, *24* (4), 508–548.

(143) Mitchell Wells, J.; McLuckey, S. A. Collision-Induced Dissociation (CID) of Peptides and Proteins. In *Biological Mass Spectrometry*; Enzymology, B. T.-M. in, Ed.; Academic Press, 2005; Vol. Volume 402, pp 148–185.

(144) Cunningham, C.; Glish, G. L.; Burinsky, D. J. High Amplitude Short Time Excitation: A Method to Form and Detect Low Mass Product Ions in a Quadrupole Ion Trap Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (1), 81–84.

(145) Zubarev, R. A.; Zubarev, A. R.; Savitski, M. M. Electron Capture/Transfer versus Collisionally Activated/Induced Dissociations: Solo or Duet? *J. Am. Soc. Mass Spectrom.* **2008**, *19* (6), 753–761.

(146) Mcluckey, S. A.; Van Berkel, G. J.; Glish, G. L. Tandem Mass Spectrometry of Small, Multiply Charged Oligonucleotides. *J. Am. Soc. Mass Spectrom.* **1992**, *3* (1), 60–70.

(147) University of Bristol http://www.chm.bris.ac.uk/ms/cid.xhtml.

(148) Dongré, A. R.; Jones, J. L.; Somogyi, Á.; Wysocki, V. H. Influence of Peptide Composition, Gas-Phase Basicity, and Chemical Modification on Fragmentation Efficiency: Evidence for the Mobile Proton Model. *J. Am. Chem. Soc.* **1996**, *118* (35), 8365–8374.

(149) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. Mobile and Localized Protons: A Framework for Understanding Peptide Dissociation. *J. Mass Spectrom.* **2000**, *35* (12), 1399–1406.

(150) Olsen, J. V; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M. Higher-Energy C-Trap Dissociation for Peptide Modification Analysis. *Nat Meth* **2007**, *4* (9), 709–712.

(151) Ichou, F.; Schwarzenberg, A.; Lesage, D.; Alves, S.; Junot, C.; Machuron-Mandard, X.; Tabet, J.-C. Comparison of the Activation Time Effects and the Internal Energy Distributions for the CID, PQD and HCD Excitation Modes. *J. Mass Spectrom.* **2014**, *49* (6), 498–508.

(152) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. Electron Detachment Dissociation of Peptide Di-Anions: An Electron–hole Recombination Phenomenon. *Chem. Phys. Lett.* **2001**, *342* (3), 299–302.

(153) Coon, J. J.; Shabanowitz, J.; Hunt, D. F.; Syka, J. E. P. Electron Transfer Dissociation of Peptide Anions. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (6), 880–882.

(154) Zubarev, R. A. Reactions of Polypeptide Ions with Electrons in the Gas Phase. *Mass Spectrom. Rev.* **2003**, *22* (1), 57–77.

(155) Zhurov, K. O.; Fornelli, L.; Wodrich, M. D.; Laskay, U. A.; Tsybin, Y. O. Principles of Electron Capture and Transfer Dissociation Mass Spectrometry Applied to Peptide and Protein Structure Analysis. *Chem. Soc. Rev.* **2013**, *42* (12), 5014–5030.

(156) Zubarev, R. A. Electron-Capture Dissociation Tandem Mass Spectrometry. *Curr. Opin. Biotechnol.* **2004**, *15* (1), 12–16.

(157) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry. *Proc. Natl. Acad. Sci. United States Am.* **2004**, *101* (26), 9528–9533.

(158) National High Magnetic Field Laboratory https://nationalmaglab.org/user-facilities/icr/techniques/electron-transfer-dissociation.

(159) Tureček, F. NCα Bond Dissociation Energies and Kinetics in Amide and Peptide Radicals. Is the Dissociation a Non-Ergodic Process? *J. Am. Chem. Soc.* **2003**, *125* (19), 5954–5963.

(160) Leymarie, N.; Costello, C. E.; O'Connor, P. B. Electron Capture Dissociation Initiates a Free Radical Reaction Cascade. *J. Am. Chem. Soc.* **2003**, *125* (29), 8949–8958.

(161) Ledvina, A. R.; McAlister, G. C.; Gardner, M. W.; Smith, S. I.; Madsen, J. A.; Schwartz, J. C.; Stafford, G. C. J.; Syka, J. E. P.; Brodbelt, J. S.; Coon, J. J. Infrared Photoactivation Reduces Peptide Folding and Hydrogen-Atom Migration Following ETD Tandem Mass Spectrometry. *Angew. Chem. Int. Ed. Engl.* **2009**, *48* (45), 8526–8528.

(162) Horn, D. M.; Ge, Y.; McLafferty, F. W. Activated Ion Electron Capture Dissociation for Mass Spectral Sequencing of Larger (42 kDa) Proteins. *Anal. Chem.* **2000**, *72* (20), 4778–4784.

(163) Chung, T. W.; Moss, C. L.; Zimnicka, M.; Johnson, R. S.; Moritz, R. L.; Turecek, F. Electron-

Capture and -Transfer Dissociation of Peptides Tagged with Tunable Fixed-Charge Groups: Structures and Dissociation Energetics. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (1), 13–30.

(164) Ueberheide, B. M.; Fenyo, D.; Alewood, P. F.; Chait, B. T. Rapid Sensitive Analysis of Cysteine Rich Peptide Venom Components. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (17), 6910–6915.

(165) Cassou, C. A.; Sterling, H. J.; Susa, A. C.; Williams, E. R. Electrothermal Supercharging in Mass Spectrometry and Tandem Mass Spectrometry of Native Proteins. *Anal. Chem.* **2013**, *85* (1), 138–146.

(166) Cooper, H. J.; Hakansson, K.; Marshall, A. G. The Role of Electron Capture Dissociation in Biomolecular Analysis. *Mass Spectrom. Rev.* **2005**, *24* (2), 201–222.

(167) Tsybin, Y. O.; Quinn, J. P.; Tsybin, O. Y.; Hendrickson, C. L.; Marshall, A. G. Electron Capture Dissociation Implementation Progress in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (6), 762–771.

(168) Ge, Y.; Rybakova, I. N.; Xu, Q.; Moss, R. L. Top-down High-Resolution Mass Spectrometry of Cardiac Myosin Binding Protein C Revealed That Truncation Alters Protein Phosphorylation State. *Proc. Natl. Acad. Sci.* **2009**, *106* (31), 12658–12663.

(169) Good, D. M.; Wirtala, M.; McAlister, G. C.; Coon, J. J. Performance Characteristics of Electron Transfer Dissociation Mass Spectrometry. *Mol. Cell. Proteomics* **2007**, *6* (11), 1942–1951.

(170) Fornelli, L.; Damoc, E.; Thomas, P. M.; Kelleher, N. L.; Aizikov, K.; Denisov, E.; Makarov, A.; Tsybin, Y. O. Analysis of Intact Monoclonal Antibody IgG1 by Electron Transfer Dissociation Orbitrap FTMS. *Mol. Cell. Proteomics* **2012**, *11* (12), 1758–1767.

(171) Tsybin, Y. O.; Fornelli, L.; Stoermer, C.; Luebeck, M.; Parra, J.; Nallet, S.; Wurm, F. M.; Hartmer, R. Structural Analysis of Intact Monoclonal Antibodies by Electron Transfer Dissociation Mass Spectrometry. *Anal. Chem.* **2011**, *83* (23), 8919–8927.

(172) Brodbelt, J. S.; Wilson, J. J. Infrared Multiphoton Dissociation in Quadrupole Ion Traps. *Mass Spectrom. Rev.* **2009**, *28* (3), 390–424.

(173) Zhou, W.; Håkansson, K. Structural Characterization of Carbohydrates by Fourier Transform Tandem Mass Spectrometry. *Curr. Proteomics* **2011**, *8* (4), 297–308.

(174) Stafford, G. C.; Kelley, P. E.; Syka, J. E. P.; Reynolds, W. E.; Todd, J. F. J. Recent Improvements in and Analytical Applications of Advanced Ion Trap Technology. *Int. J. Mass Spectrom. Ion Process.* **1984**, *60* (1), 85–98.

(175) Payne, A. H.; Glish, G. L. Thermally Assisted Infrared Multiphoton Photodissociation in a Quadrupole Ion Trap. *Anal. Chem.* **2001**, *73* (15), 3542–3548.

(176) Newsome, G. A.; Glish, G. L. Improving IRMPD in a Quadrupole Ion Trap. *J. Am. Soc. Mass*

Spectrom. **2009**, *20* (6), 1127–1131.

(177) Remes, P. M.; Glish, G. L. Collisional Cooling in a Quadrupole Ion Trap at Sub-Ambient Temperatures. *Int. J. Mass Spectrom.* **2007**, *265* (2–3), 176–181.

(178) Vartanian, V. H.; Goolsby, B.; Brodbelt, J. S. Identification of Tetracycline Antibiotics by Electrospray Ionization in a Quadrupole Ion Trap. *J. Am. Soc. Mass Spectrom.* **1998**, *9* (10), 1089–1098.

(179) Goolsby, B. J.; Brodbelt, J. S. Tandem Infrared Multiphoton Dissociation and Collisionally Activated Dissociation Techniques in a Quadrupole Ion Trap. *Anal. Chem.* **2001**, *73* (6), 1270–1276.

(180) Doussineau, T.; Antoine, R.; Santacreu, M.; Dugourd, P. Pushing the Limit of Infrared Multiphoton Dissociation to Megadalton-Size DNA Ions. *J. Phys. Chem. Lett.* **2012**, *3* (16), 2141–2145.

(181) Mazzitelli, C. L.; Brodbelt, J. S. Probing Ligand Binding to Duplex DNA Using KMnO4 Reactions and Electrospray Ionization Tandem Mass Spectrometry. *Anal. Chem.* **2007**, *79* (12), 4636–4647.

(182) Hashimoto, Y.; Hasegawa, H.; Waki, I. High Sensitivity and Broad Dynamic Range Infrared Multiphoton Dissociation for a Quadrupole Ion Trap. *Rapid Commun. Mass Spectrom.* **2004**, *18* (19), 2255–2259.

(183) Thompson, M. S.; Cui, W.; Reilly, J. P. Fragmentation of Singly Charged Peptide Ions by Photodissociation at Lambda = 157 Nm. *Angew. Chem. Int. Ed. Engl.* **2004**, *43* (36), 4791–4794.

(184) Madsen, J. A.; Boutz, D. R.; Brodbelt, J. S. Ultrafast Ultraviolet Photodissociation at 193 Nm and Its Applicability to Proteomic Workflows. *J. Proteome Res.* **2010**, *9* (8), 4205–4214.

(185) Halim, M. A.; Girod, M.; MacAleese, L.; Lemoine, J.; Antoine, R.; Dugourd, P. Combined Infrared Multiphoton Dissociation with Ultraviolet Photodissociation for Ubiquitin Characterization. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (9), 1435–1442.

(186) Yeh, G. K.; Sun, Q.; Meneses, C.; Julian, R. R. Rapid Peptide Fragmentation without Electrons, Collisions, Infrared Radiation, or Native Chromophores. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (3), 385–393.

(187) O'Brien, J. P.; Pruet, J. M.; Brodbelt, J. S. Chromogenic Chemical Probe for Protein Structural Characterization via Ultraviolet Photodissociation Mass Spectrometry. *Anal. Chem.* **2013**, *85* (15), 7391–7397.

(188) Parthasarathi, R.; He, Y.; Reilly, J. P.; Raghavachari, K. New Insights into the Vacuum UV Photodissociation of Peptides. *J. Am. Chem. Soc.* **2010**, *132* (5), 1606–1610.

(189) Webber, N.; He, Y.; Reilly, J. P. 157 Nm Photodissociation of Dipeptide Ions Containing N-Terminal Arginine. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (2), 196–203.

(190) Zhang, L.; Reilly, J. P. Peptide de Novo Sequencing Using 157 Nm Photodissociation in a Tandem Time-of-Flight Mass Spectrometer. *Anal. Chem.* **2010**, *82* (3), 898–908.

(191) Zhang, L.; Reilly, J. P. De Novo Sequencing of Tryptic Peptides Derived from Deinococcus Radiodurans Ribosomal Proteins Using 157 Nm Photodissociation MALDI TOF/TOF Mass Spectrometry. *J. Proteome Res.* **2010**, *9* (6), 3025–3034.

(192) Moon, J. H.; Yoon, S. H.; Kim, M. S. Photodissociation of Singly Protonated Peptides at 193 Nm Investigated with Tandem Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2005**, *19* (22), 3248–3252.

(193) Moon, J. H.; Shin, Y. S.; Cha, H. J.; Kim, M. S. Photodissociation at 193 Nm of Some Singly Protonated Peptides and Proteins with M/z 2000–9000 Using a Tandem Time-of-Flight Mass Spectrometer Equipped with a Second Source for Delayed Extraction/post-Acceleration of Product Ions. *Rapid Commun. Mass Spectrom.* **2007**, *21* (3), 359–368.

(194) Moon, J. H.; Yoon, S. H.; Bae, Y. J.; Kim, M. S. Dissociation Kinetics of Singly Protonated Leucine Enkephalin Investigated by Time-Resolved Photodissociation Tandem Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (7), 1151–1158.

(195) Choi, K. M.; Yoon, S. H.; Sun, M.; Oh, J. Y.; Moon, J. H.; Kim, M. S. Characteristics of Photodissociation at 193 Nm of Singly Protonated Peptides Generated by Matrix-Assisted Laser Desorption Ionization (MALDI). *J. Am. Soc. Mass Spectrom.* **2006**, *17* (12), 1643–1653.

(196) Racaud, A.; Antoine, R.; Dugourd, P.; Lemoine, J. Photoinduced Dissociation of Heparin-Derived Oligosaccharides Controlled by Charge Location. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (12), 2077–2084.

(197) Agarwal, A.; Diedrich, J. K.; Julian, R. R. Direct Elucidation of Disulfide Bond Partners Using Ultraviolet Photodissociation Mass Spectrometry. *Anal. Chem.* **2011**, *83* (17), 6455–6458.

(198) Zhang, X.; Li, H.; Moore, B.; Wongkongkathep, P.; Ogorzalek Loo, R. R.; Loo, J. A.; Julian, R. R. Radical-Directed Dissociation of Peptides and Proteins by Infrared Multiphoton Dissociation and Sustained off-Resonance Irradiation Collision-Induced Dissociation with Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2014**, *28* (24), 2729–2734.

(199) Moore, B.; Sun, Q.; Hsu, J. C.; Lee, A. H.; Yoo, G. C.; Ly, T.; Julian, R. R. Dissociation Chemistry of Hydrogen-Deficient Radical Peptide Anions. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (3), 460–468.

(200) Diedrich, J. K.; Julian, R. R. Facile Identification of Phosphorylation Sites in Peptides by Radical Directed Dissociation. *Anal. Chem.* **2011**, *83* (17), 6818–6826.

(201) Tao, Y.; Quebbemann, N. R.; Julian, R. R. Discriminating D-Amino Acid-Containing Peptide Epimers by Radical-Directed Dissociation Mass Spectrometry. *Anal. Chem.* **2012**, *84* (15), 6814–6820.

(202) Aponte, J. R.; Vasicek, L.; Swaminathan, J.; Xu, H.; Koag, M. C.; Lee, S.; Brodbelt, J. S. Streamlining Bottom-Up Protein Identification Based on Selective Ultraviolet Photodissociation (UVPD) of Chromophore-Tagged Histidine- and Tyrosine-Containing Peptides. *Anal. Chem.* **2014**, *86* (13), 6237–6244.

(203) Cotham, V. C.; Wine, Y.; Brodbelt, J. S. Selective 351 Nm Photodissociation of Cysteine-Containing Peptides for Discrimination of Antigen-Binding Regions of IgG Fragments in Bottom-Up Liquid Chromatography–Tandem Mass Spectrometry Workflows. *Anal. Chem.* **2013**, *85* (11), 5577–5585.

(204) Enjalbert, Q. Nouvel Outil de Quantification de Biomarqueur Couplant La Spectroscopie Optique et La Spectrométrie de Masse : La Photo-SRM, Claude Bernard Lyon 1, 2013.

(205) Michalski, A.; Damoc, E.; Hauschild, J.-P.; Lange, O.; Wieghaus, A.; Makarov, A.; Nagaraj, N.; Cox, J.; Mann, M.; Horning, S. Mass Spectrometry-Based Proteomics Using Q Exactive, a High-Performance Benchtop Quadrupole Orbitrap Mass Spectrometer. *Mol. Cell. Proteomics* **2011**, *10* (9), M111.011015.

(206) Paul Gates. High Performance Liquid Chromatography Mass Spectrometry (HPLC/MS) http://www.bris.ac.uk/nerclsmsf/techniques/hplcms.html.

(207) Iribarne, J. V; Thomson, B. A. On the Evaporation of Small Ions from Charged Droplets. *J. Chem. Phys.* **1976**, *64* (6), 2287–2294.

(208) Thomson, B. A.; Iribarne, J. V. Field Induced Ion Evaporation from Liquid Surfaces at Atmospheric Pressure. *J. Chem. Phys.* **1979**, *71* (11), 4451–4463.

(209) Notizen: Ein Neues Massenspektrometer Ohne Magnetfeld . *Zeitschrift für Naturforschung A* . 1953, p 448.

(210) Masse-spec.fr http://masse-spec.fr/analyseurs-quadripolaires.

(211) Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal. Chem.* **2000**, *72* (6), 1156–1162.

(212) Hardman, M.; Makarov, A. A. Interfacing the Orbitrap Mass Analyzer to an Electrospray Ion Source. *Anal. Chem.* **2003**, *75* (7), 1699–1705.

(213) Perry, R. H.; Cooks, R. G.; Noll, R. J. Orbitrap Mass Spectrometry: Instrumentation, Ion Motion and Applications. *Mass Spectrom. Rev.* **2008**, *27* (6), 661–699.

(214) Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R. The Orbitrap: A New Mass Spectrometer. *J. Mass Spectrom.* **2005**, *40* (4), 430–443.

(215) Masse-spec.fr http://masse-spec.fr/orbitrap-principe-utilisations.

(216) Makarov, A.; Denisov, E.; Kholomeev, A.; Balschun, W.; Lange, O.; Strupat, K.; Horning, S. Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *Anal. Chem.* **2006**, *78* (7), 2113–2120.

(217) Makarov, A.; Denisov, E.; Lange, O. Performance Evaluation of a High-Field Orbitrap Mass Analyzer. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (8), 1391–1396.

(218) Olsen, J. V; de Godoy, L. M. F.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M. Parts per Million Mass Accuracy on an Orbitrap Mass Spectrometer via Lock Mass Injection into a C-Trap. *Mol. Cell. Proteomics* **2005**, *4* (12), 2010–2021.

(219) Bilbao, A.; Zhang, Y.; Varesio, E.; Luban, J.; Strambio-De-Castillia, C.; Lisacek, F.; Hopfgartner, G. Ranking Fragment Ions Based on Outlier Detection for Improved Label-Free Quantification in Data-Independent Acquisition LC-MS/MS. *J. Proteome Res.* **2015**, *14* (11), 4581–4593.

(220) SABLIN, M.-P.; LE TOURNEAU, C.; LOIRAT, D.; DIÉRAS, V. Les Inhibiteurs de Tyrosine Kinase Dans Les Cancers Du Sein HER2+. *La Lett. du cancérologue* **2015**.

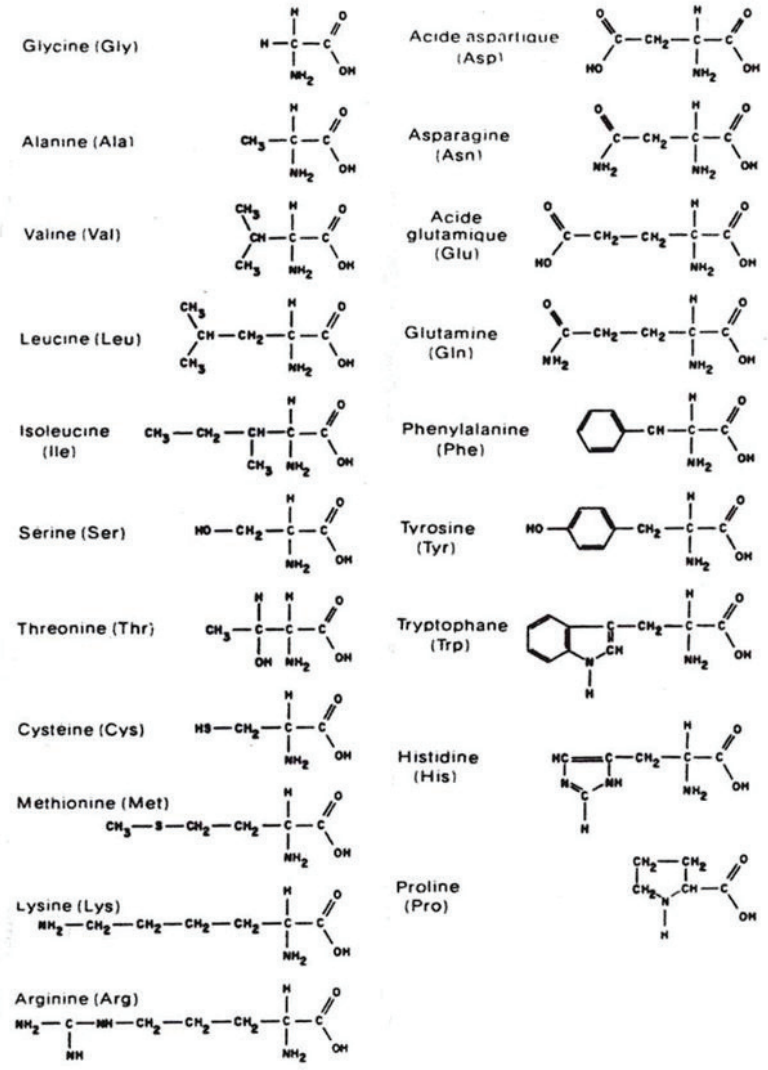
(221) Shao, C.; Zhang, Y.; Sun, W. Statistical Characterization of HCD Fragmentation Patterns of Tryptic Peptides on an LTQ Orbitrap Velos Mass Spectrometer. *J. Proteomics* **2014**, *109*, 26–37.

(222) Michalski, A.; Neuhauser, N.; Cox, J.; Mann, M. A Systematic Investigation into the Nature of Tryptic HCD Spectra. *J. Proteome Res.* **2012**, *11* (11), 5479–5491.

(223) Holden, D. D.; Makarov, A.; Schwartz, J. C.; Sanders, J. D.; Zhuk, E.; Brodbelt, J. S. Ultraviolet Photodissociation Induced by Light-Emitting Diodes in a Planar Ion Trap. *Angew. Chemie Int. Ed.* **2016**, *55* (40), 12417–12421.

Annexes

Annexe 1 : Liste des 20 acides aminés naturels constituant les protéines chez l'homme



Annexe 2 : Liste des 354 peptides à cystéine présent au sein de la bibliothèque de spectres LID à 473 nm rapporteurs de 354 protéines kinases humaines

Peptides à cystéine	Protéines kinases
AACLLDGVPVALK	Serine/threonine-protein kinase Nek7
ACANDLLVDEFLK	Mitogen-activated protein kinase kinase kinase 5
ACDALYPPGPPSNPK	Titin
ACDGRPYAGAVQK	Serine/threonine-protein kinase 19
ACGLLGYSSQDLIGQK	PAS domain-containing serine/threonine-protein kinase
ACPPEVYAIMR	High affinity nerve growth factor receptor
ADLWSIGTIVYQCLTGK	Serine/threonine-protein kinase ULK1
ADQQYECVAEIGEGAYGK	Cell division protein kinase 6
ADSDSADSACTTVPSPPR	Ephrin type-B receptor 3
AGFNGTSVEGQCK	Mitogen-activated protein kinase kinase kinase 4
AGFQTEDFSLYACASPK	B-Raf proto-oncogene serine/threonine-protein kinase
AGLECSDIEK	Kalirin
ALAIVCDFK	Serine/threonine-protein kinase MRCK alpha
ALDVWATGVTLYCFVYGK	Calcium/calmodulin-dependent protein kinase kinase 1
ALEAGGCAVVR	Serine/threonine-protein kinase/endoribonuclease IRE2
ALLEVICTLHK	Phosphorylase b kinase gamma catalytic chain, skeletal muscle isoform
APEILLGCK	Cell division protein kinase 2
APEILLGLPFCEK	Homeodomain-interacting protein kinase 4
APELFSVQSHCVIDER	Serine/threonine-protein kinase 16
APGPSTGDLTGPGPCPR	Obscurin
ASCSTAGSGSR	MAP/microtubule affinity-regulating kinase 4
ATAGDCLESSYFK	Cell division protein kinase 10
AVLVNGTECLLAR	Macrophage-stimulating protein receptor
AVPAECEGLTK	Serine/threonine-protein kinase WNK3
AVSACQEIQAIFTQK	Serine/threonine-protein kinase PINK1, mitochondrial
CAAQLASALEYIHAR	Serine/threonine-protein kinase SBK2
CAVSNDLTQQEIR	Kinase suppressor of Ras 1
CDNVFITGPTGSVK	Serine/threonine-protein kinase WNK4
CEETATGLK	Myosin light chain kinase family member 4
CENALLQGFNLK	Testis-specific serine/threonine-protein kinase 3
CENVLLSPDER	Testis-specific serine/threonine-protein kinase 6
CFEISDADTK	Serine/threonine-protein kinase PLK1
CGGPGDASPR	Wee1-like protein kinase
CGSPSDSSTTEEMEVAVSK	RAC-beta serine/threonine-protein kinase
CHAPENVELALR	Serine/threonine-protein kinase PDIK1L
CIGSDEQLGK	Eukaryotic translation initiation factor 2-alpha kinase 4
CIHATTNMEFAVK	Ribosomal protein S6 kinase alpha-6
CLDEGGYYISPR	Tyrosine-protein kinase BLK

CLIELEPVLR	Receptor-interacting serine/threonine-protein kinase 2
CLLADLPLPPELPGGDDLSK	Cell division cycle 2-like protein kinase 5
CLQAPGTEPGEQTPEGAR	Putative myosin light chain kinase 3
CMQEIPQEQIK	Mixed lineage kinase domain-like protein
CPAAGNPTPSISWLK	Fibroblast growth factor receptor 3
CPAAGNPTPTIR	Fibroblast growth factor receptor 4
CPAEAVPPAGLTK	Microtubule-associated serine/threonine-protein kinase 1
CPETPAQPDSR	Wee1-like protein kinase 2
CPSSGTPNPTLR	Basic fibroblast growth factor receptor 1
CQDASEVQR	Mitogen-activated protein kinase 4
CSESSTLLEK	Serine/threonine-protein kinase DCLK2
CSGLLDAPR	Microtubule-associated serine/threonine-protein kinase 2
CSIPFVKPNR	Cyclin-dependent kinase-like 2
CSLNPEWNETFR	Protein kinase C beta type
CSMDETHPGYGK	Phosphorylase b kinase gamma catalytic chain, testis/liver isoform
CSNILLNNSGQIK	Cell division cycle 2-related protein kinase 7
CSPSLHVDDR	Receptor-interacting serine/threonine-protein kinase 4
CTEALYAPAEGDK	Microtubule-associated serine/threonine-protein kinase 4
CVQQLGLALDFMHGR	Serine/threonine-protein kinase SBK1
CVSAHPGLLQTLAGPLQTPR	Putative uncharacterized serine/threonine-protein kinase SgK110
CVSGSTAISTYPK	Serine/threonine-protein kinase Nek11
CYEATDTETGSAYAVK	Serine/threonine-protein kinase PLK3
DALDELGCFQLELR	Atrial natriuretic peptide receptor B
DCANVNDFFMR	Serine/threonine-protein kinase RIO1
DDEYNPCQGSK	Proto-oncogene tyrosine-protein kinase FGR
DEQTCLMWAYEK	Serine/threonine-protein kinase TNNI3K
DFICHLLEK	Calcium/calmodulin-dependent protein kinase type 1G
DGAPQVCPIPPEQSK	Serine/threonine-protein kinase MRCK beta
DGVCAQIEK	CaM kinase-like vesicle-associated protein
DLECALLLR	Serine/threonine-protein kinase TAO2
DLKPSNLAVNEDCELK	Mitogen-activated protein kinase 14
DLKPSNLLLNTTCDLK	Mitogen-activated protein kinase 1
DLKPSNVAVNEDCELR	Mitogen-activated protein kinase 11
DLLADLIGCQR	Serine/threonine-protein kinase PRP4 homolog
DLQPVGSGAYGAVCSAVDGR	Mitogen-activated protein kinase 12
DLVSCTYQLAR	Fibroblast growth factor receptor 2
DMCSFDNEQLFTMK	Protein kinase C iota type
DNLLIADIELGCGNFGSVR	Tyrosine-protein kinase ZAP-70
DQADEAQYCQLVR	[3-methyl-2-oxobutanoate dehydrogenase [lipoamide]] kinase, mitochondrial
DQCIVISGESGSGK	Myosin-IIIB
DQLVVGQLIPDCYVELEK	Leucine-rich repeat serine/threonine-protein kinase 2

DSGDYECAAR	Alpha-type platelet-derived growth factor receptor
DSVESGCQTESSK	Serine/threonine-protein kinase NIM1
DWAHISCAAK	MAP kinase-interacting serine/threonine-protein kinase 2
EAPPCHLLIIHR	Kinase suppressor of Ras 2
EASDVLCTITK	Ribosomal protein S6 kinase alpha-2
EASSCAVNLVLR	STE20/SPS1-related proline-alanine-rich protein kinase
EATVPCSATGR	Tyrosine-protein kinase-like 7
ECLQALDFLHSNQVIHR	Serine/threonine-protein kinase PAK 3
ECLQALEFLHANQVIHR	Serine/threonine-protein kinase PAK 2
ECLQALEFLHSNQVIHR	Serine/threonine-protein kinase PAK 1
ECVYIIPSSK	Transient receptor potential cation channel subfamily M member 7
EEAFEIIVEFPETNCDVK	Alpha-protein kinase 1
EEWIAYICR	TRAF2 and NCK-interacting protein kinase
EGEEFTVTCTIK	Mast/stem cell growth factor receptor
EGFCLQNVDK	Ribosomal protein S6 kinase alpha-5
EGIDSECGPFIK	Tyrosine-protein kinase JAK1
EIEEFLSEAACMK	Proto-oncogene tyrosine-protein kinase MER
EIGQCAIQISDYLK	Transient receptor potential cation channel subfamily M member 6
EISIMQQCDSPHVVK	Serine/threonine-protein kinase 4
EISIMQQCDSPYVVK	Serine/threonine-protein kinase 3
ELFSACAQSVHEYLR	G protein-coupled receptor kinase 5
ELGAALQGCLLR	Non-receptor tyrosine-protein kinase TYK2
ELMQLCWSSEPK	Receptor-interacting serine/threonine-protein kinase 3
ELNDMQAQLVEECAHR	Rho-associated protein kinase 1
ELNLPTHPHCSR	Testis-expressed protein 14
ELSDSCEGAFGGGR	Serine/threonine-protein kinase WNK2
EMEQAVELCGR	Inhibitor of nuclear factor kappa-B kinase subunit beta
ENCISGPVK	Serine/threonine-protein kinase 6
ENLLLSDSPSCR	Rho-associated protein kinase 2
EPGLGQLCFK	Mitogen-activated protein kinase kinase kinase 14
EPNLCLVMEFAR	Mitogen-activated protein kinase kinase kinase 9
EPVGSGCVAQVYK	Uncharacterized aarF domain-containing protein kinase 2
EQGECLSLAVLDLAR	Tyrosine-protein kinase JAK3
EQSQPCADNAVLSSGLTAAR	Casein kinase II subunit alpha'
ESCPSVSIPSSDEHR	Serine/threonine-protein kinase Sgk3
ESDCHYLEAR	Dual specificity protein kinase CLK4
ESHSCPQQVPMPPGQEAQR	Serine/threonine-protein kinase Nek8
ETACSSPGSIER	Interleukin-1 receptor-associated kinase-like 2
ETIEDCWDQDAEAR	Bone morphogenetic protein receptor type-2
ETIGVGSYSECK	Ribosomal protein S6 kinase alpha-1
EVLDSETLCR	Serine/threonine-protein kinase 11

EVSALECEIQLLK	Mitogen-activated protein kinase kinase kinase 3
EVYDVMLGCWQR	NT-3 growth factor receptor
EVYSWGCGEYGR	Serine/threonine-protein kinase Nek9
FEACNYPLELYER	Protein kinase C gamma type
FEVALLPEPCR	Retinal guanylyl cyclase 1
FFSFIEGCLVK	Mitogen-activated protein kinase kinase kinase kinase 4
FGALSCWLGEMSTEHR	Mitogen-activated protein kinase kinase kinase kinase 1
FHCPSDEWAR	Mitogen-activated protein kinase 15
FIDFIDTCLIK	Misshapen-like kinase 1
FLLFCVHGDAR	Serine/threonine-protein kinase MARK1
FLSPQLIAEEFCLK	Serine/threonine-protein kinase Nek1
FPECGFFGMYDK	Serine/threonine-protein kinase D3
FPECGFYGLYDK	Serine/threonine-protein kinase D2
FPQCVPINLK	Serine/threonine-protein kinase MAK
FSPEFVDFTAQCLR	Dual specificity mitogen-activated protein kinase kinase 3
FTCEQALQHPWIAGDTALDK	Calcium/calmodulin-dependent protein kinase type 1
FTCIVTGYPEPEVTWYK	Alpha-protein kinase 3
FTGVASGSTGSTGSCSVGAK	Serine/threonine-protein kinase tousled-like 1
FTTEIHPSCVTR	Ephrin type-A receptor 2
FVDDVNNNYYEAPSCPR	Nik-related protein kinase
FWQILSAVEYCHDHHIVHR	Serine/threonine-protein kinase SIK1
FYAAEICIALNFLHER	Protein kinase C zeta type
FYSWEECMNLR	Serine/threonine-protein kinase ICK
GAPCTTPPSAPR	Ephrin type-B receptor 4
GAWCTVVLVR	Serine/threonine-protein kinase receptor R3
GAYCLSVSDFDNAK	Proto-oncogene tyrosine-protein kinase Src
GCEEIEFSQLIFK	Alpha-protein kinase 2
GCFQVYEQGK	Activin receptor type-1
GDVVTILEACENK	Megakaryocyte-associated tyrosine-protein kinase
GECVITPSTDVK	Serine/threonine-protein kinase/endoribonuclease IRE1
GELLGSGCFGVVHR	Non-receptor tyrosine-protein kinase TNK1
GEVCNAVLAK	Muscle, skeletal receptor tyrosine protein kinase
GFPPSASLCLLDLVK	DNA-dependent protein kinase catalytic subunit
GFSVSVESACSNYATTVQVK	LIM domain kinase 2
GGFGEVCAVQVK	G protein-coupled receptor kinase 7
GGFGEVSACQMK	Rhodopsin kinase
GGPIYIITEYCR	Beta-type platelet-derived growth factor receptor
GGQGLEAALSCPR	Fas-activated serine/threonine kinase
GHYVLCDFGSATNK	AP2-associated protein kinase 1
GIDCLAVAYGVAVNK	Mitogen-activated protein kinase kinase kinase 11
GLDTETTVEVAWCELQDR	Serine/threonine-protein kinase WNK1

GLNLECLWEK	Serine/threonine-protein kinase LMTK3
GMDQSCLSADGAGR	Ribosomal protein S6 kinase-like 1
GPLSCDTVLK	Cyclin G-associated kinase
GPNCIEHNEALLEEAK	Receptor-interacting serine/threonine-protein kinase 1
GQFCIPEHISPK	Tribbles homolog 1
GQHCIEISSLPGDLVIVEK	Tau-tubulin kinase 2
GSCGDLPGLR	Mitogen-activated protein kinase kinase kinase 12
GSCIANAEEVDVPIK	Ephrin type-B receptor 2
GSLALDCAPLLR	ALK tyrosine kinase receptor
GSLENHCTGNPR	Homeodomain-interacting protein kinase 2
GSSSSESQKPCEALR	Calcium/calmodulin-dependent protein kinase kinase 2
GSTESCNTTTEDEDLK	Calcium/calmodulin-dependent protein kinase type II gamma chain
GTCIPNAEEVDVPIK	Ephrin type-B receptor 1
GTCVSSAEEEAENAPR	Ephrin type-A receptor 7
GYPSEFATYLNFCR	Casein kinase I isoform delta
GYPSEFSTYLNFCR	Casein kinase I isoform epsilon
HACPAYVGPEILSSR	Tribbles homolog 3
HCHLLNIAHR	MAP kinase-activated protein kinase 5
HENVVNLIEICR	Cell division protein kinase 9
HIGSINPNCNVLEVIK	[Pyruvate dehydrogenase [lipoamide]] kinase isozyme 1, mitochondrial
HLSSCAAPAPLTSAER	Twinfilin-2
HPNLVQLLGVCTR	Proto-oncogene tyrosine-protein kinase ABL1
HSSLNCQPHFDLQNR	FL cytokine receptor
HYQTIQEAGDWCVPSTEPK	[Pyruvate dehydrogenase [lipoamide]] kinase isozyme 2, mitochondrial
HYTNASDGLCTR	Proto-oncogene tyrosine-protein kinase LCK
IAFCQGFDK	Mitogen-activated protein kinase 10
ICMNEDPAK	Integrin-linked protein kinase
ICQAHFAPISPGFSR	Serine/threonine-protein kinase Nek5
IDVQQLACDPYLLPHIR	Serine/threonine-protein kinase tousled-like 2
IEQCDYPPLPGEHYSEK	Serine/threonine-protein kinase Nek6
IGDFGLACTDILQK	Eukaryotic translation initiation factor 2-alpha kinase 1
IGEGGFGCVYR	Interleukin-1 receptor-associated kinase 1
IGEGSTGIVCIATVR	Serine/threonine-protein kinase PAK 4
IGPHCFELLR	Ribosomal protein S6 kinase beta-2
IGVVGPWACDSLFSK	Retinal guanylyl cyclase 2
IHCASSWINPDTR	Mitogen-activated protein kinase kinase kinase kinase 3
IIDCLEVEYYDK	cGMP-dependent protein kinase 2
IIGSGDSGEVCYGR	Ephrin type-A receptor 8
INCAAAGNPFPVR	Tyrosine-protein kinase receptor Tie-1
INCATSWIHPDTK	Mitogen-activated protein kinase kinase kinase kinase 5
INGAYFCEGR	Angiopoietin-1 receptor

INPICNDHYR	Lymphokine-activated killer T-cell-originated protein kinase
IPIEPVESMPLCYIK	Bromodomain testis-specific protein
IPWVPPECIENPK	Tyrosine-protein kinase JAK2
ISDFGCSEK	Proto-oncogene serine/threonine-protein kinase mos
ITCVETVVPEK	Tyrosine-protein kinase BTK
ITNLENCNEER	STE20-related kinase adapter protein beta
IVQTLCTVR	Uncharacterized aarF domain-containing protein kinase 4
KPCESDFETIK	Microtubule-associated serine/threonine-protein kinase 3
LAAFAGTTSGCHQGPGPLR	Tyrosine-protein kinase SgK223
LAGPPSCIVPLMR	Serine/threonine-protein kinase PAK 7
LCDFGISGQLVDSIAK	Dual specificity mitogen-activated protein kinase kinase 4
LCDFGISGR	Dual specificity mitogen-activated protein kinase kinase 7
LCDFGVSTQLVNSIAK	Dual specificity mitogen-activated protein kinase kinase 5
LCDSGELVAIK	Glycogen synthase kinase-3 beta
LCGLGLAYEVYTR	Tyrosine-protein kinase STYK1
LCHEDVEALAAIYEEK	Death-associated protein kinase 3
LCTGHEYAAK	Calcium/calmodulin-dependent protein kinase type II beta chain
LCVMFEDGVLMR	Serine/threonine-protein kinase ATR
LDASCAANNPSLK	Serine-protein kinase ATM
LDEAQEAECQALR	Serine/threonine-protein kinase TAO3
LDEAQEAECQVLK	Serine/threonine-protein kinase TAO1
LDTMIFGVIEYCER	Heat-stable enterotoxin receptor
LECSGTILAHSNLR	Serine/threonine-protein kinase Nek4
LENILLDDNCNIK	NUAK family SNF1-like kinase 1
LFQQILSAVDYCHR	5'-AMP-activated protein kinase catalytic subunit alpha-2
LFQQILSGVDYCHR	5'-AMP-activated protein kinase catalytic subunit alpha-1
LGAGDLYPLFACMLTAR	Uncharacterized aarF domain-containing protein kinase 1
LGVHCITGQK	BR serine/threonine-protein kinase 1
LHIDEILSHCWMQPK	Testis-specific serine/threonine-protein kinase 1
LIGVCFQGSER	Tyrosine-protein kinase receptor UFO
LLAQCSEVQLGK	Serine/threonine-protein kinase SMG1
LLEAPDVLCLR	Ribosomal protein S6 kinase delta-1
LLGCEQLLTAVPGR	Serine/threonine-protein kinase N3
LLLDHGACVDAQER	Ankyrin repeat and protein kinase domain-containing protein 1
LMDVCATSR	Cell division protein kinase 4
LMEFSEHCAIILEDDR	TGF-beta receptor type-2
LMTECWAHNPASR	Bone morphogenetic protein receptor type-1B
LNCSVEGMEEPDIQWVK	Tyrosine-protein kinase receptor TYRO3
LQCVGDTAPLPWHIR	Interleukin-1 receptor-associated kinase 3
LQDCDALVTMGTGTFGR	Serine/threonine-protein kinase PRKY
LSAGCVEER	Activin receptor type-2B

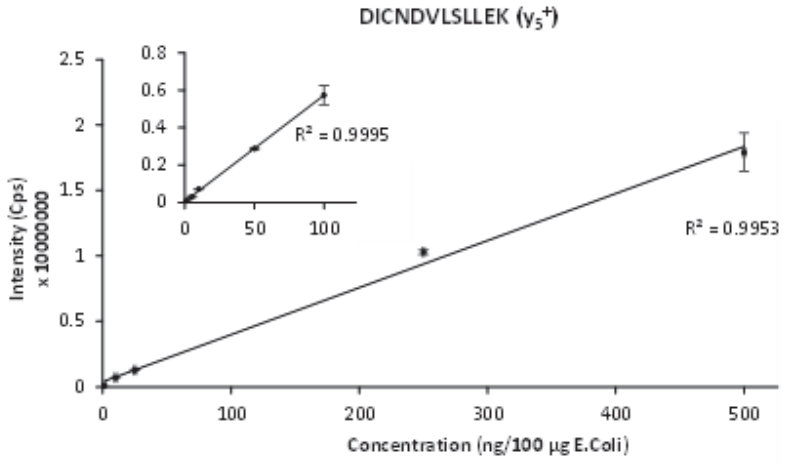
LSDFGFCAQISK	Serine/threonine-protein kinase PAK 6
LSEEAECPNPSTPSK	Serine/threonine-protein kinase 10
LSPDLSVCGQPR	Serine/threonine-protein kinase haspin
LSQNACILESVSEK	STE20-like serine/threonine-protein kinase
LTLPIPSTCPEPFAK	Mitogen-activated protein kinase kinase kinase
LTLPIPSTCPEPFAR	Mitogen-activated protein kinase kinase kinase 10
LTLVFEFCDQDLK	Cell division protein kinase 5
LTSVVGTILYSCPEVLK	Serine/threonine-protein kinase Nek10
LVQCVDAFEEK	Myosin light chain kinase, smooth muscle
LVTIEPPNPINPSTNCNGLK	Cyclin-dependent kinase-like 3
LYCGADGDWLVPLGR	Ephrin type-A receptor 6
LYGACSQDGPLLLIVEYAK	Proto-oncogene tyrosine-protein kinase receptor ret
LYICELALALEYLQR	Serine/threonine-protein kinase 32B
LYLAENYCFDSIPK	Cytoplasmic tyrosine-protein kinase BMX
MCHLPEPELNK	Protein kinase C theta type
MEQPEGCPPK	Tyrosine-protein kinase ABL2
MGGSFLICSK	Chaperone activity of bc1 complex-like, mitochondrial
MLCDQYYLSSPELK	[Pyruvate dehydrogenase [lipoamide]] kinase isozyme 4, mitochondrial
MSPDTCATILEK	Myosin IIIA
MYLVMELCEDGELK	Serine/threonine-protein kinase 33
NACLQTSSLAVR	SCY1-like protein 2
NCLVGENLLVK	BDNF/NT-3 growth factors receptor
NCLVGENNVLK	Proto-oncogene tyrosine-protein kinase FER
NCMVAEDFTVK	Insulin-like growth factor 1 receptor
NCVIDDTLQVK	Tyrosine-protein kinase RYK
NDDGVIPCQNK	Tyrosine-protein kinase Tec
NDIEQLCYVLR	Cell cycle-related kinase
NDYTQGCEPGHR	Dual specificity protein kinase CLK1
NGMCAIADLGLAVR	Activin receptor type-1B
NILVGENTLCK	Tyrosine-protein kinase 6
NILWSAENECFK	Serine/threonine-protein kinase Kist
NPGGQMEGPWCFTQNK	Tyrosine-protein kinase transmembrane receptor ROR2
NQGSELSGVISSACDK	Mitotic checkpoint serine/threonine-protein kinase BUB1
NSDCVGSYTLIPYVVTATGR	Mitogen-activated protein kinase kinase kinase 6
NTGIICTIGPASR	Pyruvate kinase PKM
NTMECSGPQDSK	Serine/threonine-protein kinase ULK4
NVLVDDGLACK	Tyrosine-protein kinase Srms
SAESCATWK	Serine/threonine-protein kinase VRK2
SAPTSPCDQEIK	Protein kinase C epsilon type
SCVPLSVQPTEPR	SPS1/STE20-related protein kinase YSK4
SDIQDSLCYK	Hormonally up-regulated neu tumor-associated kinase

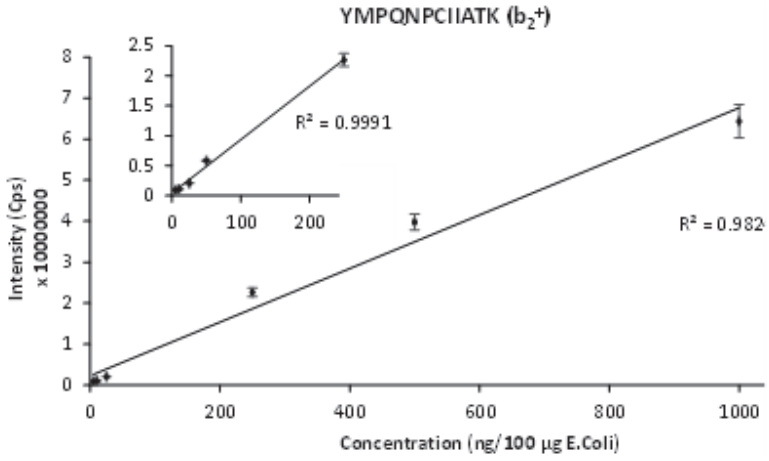
SDISIPCHYK	Tyrosine-protein kinase ITK/TSK
SDSSADCQWLDTLR	Mitogen-activated protein kinase kinase kinase MLT
SEPSVIIVSCK	Striated muscle preferentially expressed protein kinase
SEQYDLDSLCAGMEQSLR	Serine/threonine-protein kinase LATS2
SEVGQALCLTAR	Myosin light chain kinase 2, skeletal/cardiac muscle
SFACENNDFLVTDGK	Proto-oncogene tyrosine-protein kinase ROS
SFPGGAGPLGCK	Ephrin type-A receptor 1
SLAYIHSQGVCHR	Glycogen synthase kinase-3 alpha
SLDAFPEDFCR	N-terminal kinase-like protein
SLECLRPPQPQELSPEK	Leukocyte tyrosine kinase receptor
SLPELSASAELCIEDRPMPK	Mitotic checkpoint serine/threonine-protein kinase BUB1 beta
SNCITENPAGPYGQR	Mitogen-activated protein kinase kinase kinase 15
SNYQGLPVPCVK	Serine/threonine-protein kinase SRPK3
SPCVEEEQSR	Tau-tubulin kinase 1
SQACGGNLGSIEELR	Serine/threonine-protein kinase 24
SQCLSTLVRPVFGELK	Serine/threonine-protein kinase 25
SQQLSNDHICYFLYQILR	Mitogen-activated protein kinase 3
SSCHGLPPTPK	Mitogen-activated protein kinase kinase kinase kinase 2
STVIENPGALCSPQSQR	Tyrosine-protein kinase FRK
TACTNFMMTPYVVTR	Mitogen-activated protein kinase 9
TATIIEELADALTYCHDK	Serine/threonine-protein kinase 13
TCEIVEDTESSR	Serine/threonine-protein kinase 40
TCVFFEAPGVR	Anti-Muellerian hormone type-2 receptor
TDDLSNDVCAVLK	Serine/threonine-protein kinase N2
TEPFQDGYSLCPGR	Serine/threonine-protein kinase 17A
TFSFCGTIEYMAPEIIR	Ribosomal protein S6 kinase alpha-4
TFSPHFHHFVEQCLQR	STE20-related kinase adapter protein alpha
TICGTPNYLSPEVLNK	Serine/threonine-protein kinase PLK2
TLCGTPEYIAPEVLLR	Serine/threonine-protein kinase H2
TLGIGAFGEVCLAR	Serine/threonine-protein kinase LATS1
TLGSGACGEVK	Serine/threonine-protein kinase Chk2
TNCPATVINGQFVER	Insulin receptor
TNQTGVEPCYGDK	Activin receptor type-2A
TQLGDCTVATQASGLLSNK	Homeodomain-interacting protein kinase 1
TTCGTPEYIAPEVLVR	Serine/threonine-protein kinase H1
TTSVGPSNSGGSLCAMSGR	Serine/threonine-protein kinase ULK2
TVAACDLLQSLLHK	Serine/threonine-protein kinase Sgk2
VCYGLGMEHLR	Receptor tyrosine-protein kinase erbB-2
VDVFSFGIVLCEIIGR	LIM domain kinase 1
VELYGCLWR	Epithelial discoidin domain-containing receptor 1
VEPSSCVFQER	Homeodomain-interacting protein kinase 3

VGTTVHCDYLNIPHK	Transcription initiation factor TFIID 210 kDa subunit
VHSENNACINFK	Serine/threonine-protein kinase PFTAIRE-1
VIHNSMCIAEIDLIPR	TGF-beta receptor type-1
VIPTECAIK	Serine/threonine-protein kinase RIO3
VIYICSSPDAFR	Atrial natriuretic peptide receptor A
VLDANNCDYEQR	MAP/microtubule affinity-regulating kinase 3
VLDANSCQSELHEK	Serine/threonine-protein kinase MARK2
VMCDSNDLPK	Protein kinase-like protein SgK196
VPFAWCAPESLK	Activated CDC42 kinase 1
VPLCINPQSAAFK	Serine/threonine protein kinase NLK
VPTMCVDWSNIR	RAF proto-oncogene serine/threonine-protein kinase
VPTVCVDMSTNR	A-Raf proto-oncogene serine/threonine-protein kinase
VPVNLLNSPDCDVK	Dual specificity protein kinase TTK
VSCLSEELIASR	Protein-associating with the carboxyl-terminal domain of ezrin
VSSDFLDLIQSLLCGQK	Citron Rho-interacting kinase
VSSECQHLIR	Proto-oncogene serine/threonine-protein kinase Pim-1
VVECLDHAR	Dual specificity protein kinase CLK3
VVGAGEFGEVCSGR	Ephrin type-A receptor 3
WYAPECINYYK	Tyrosine-protein kinase SYK
YCHDLDIVHR	Testis-specific serine/threonine-protein kinase 2
YCSVTTQADGR	Dual specificity tyrosine-phosphorylation-regulated kinase 3
YEELECLYAAVGK	Transformation/transcription domain-associated protein
YEGSDFQCK	Bone morphogenetic protein receptor type-1A
YGACSEPLAGK	Testis-specific serine/threonine-protein kinase 4
YLCLMGASVEALTTDGK	Death-associated protein kinase 1
YLTDFEPIQCLGR	Eukaryotic translation initiation factor 2-alpha kinase 3
YLTEQCLNHPTFQTQR	Cyclin-dependent kinase-like 5
YPVGLMGSCLPATAGFSYEK	Tyrosine-protein kinase TXK
YSFGATCVK	Epidermal growth factor receptor

Annexe 3 : Aires, CV et courbes de calibration obtenus pour les peptides à cystéine dérivés DIC-NDVLSLLEK et YMPQNPCIIATK rapporteurs des protéines 14-3-3 protein zeta/delta et Histone-binding protein RBBP4

	DICNDVLSLLEK (y_5^+)		YMPQNPCIIATK (b_2^+)	
Concentration (ng/100 µg)	**Aire**	**CV (%)**	**Aire**	**CV (%)**
1	83063	10		
5	297523	20	906961	20
10	711574	5	1063083	16
25	1280579	7	2072903	3
50	2873765	4	5817272	4
100	5735534	19	22635695	10
250	10287102	4	39762479	10
500	17906080	16	64354244	13





Annexe 4 : Liste des 47 peptides à cystéine présent au sein de la bibliothèque de spectres LID à 473 nm rapporteurs de 33 protéines humaines

Peptides à cystéine	**Protéines humaines**
AAIGCGIVESILNWVK	DNA topoisomerase 2-alpha
AGKPVICATQMLESMIK	Pyruvate kinase PKM
ALGFPESLVIQAYFACEK	UV excision repair protein RAD23 homolog A
CDISLQFFLPFSLGK	X-ray repair cross-complementing protein 5
CLAALASLR	Interleukin enhancer-binding factor 3
CMDVYTPK	Casein kinase II subunit beta
CNEIINWLDK	Heat shock cognate 71 kDa protein
CNNVLYIR	Small nuclear ribonucleoprotein F
CQGITAPIEAQVR	RNA-binding protein 5
CTAGTLHNLSHHR	Catenin beta-1
DGFYEAELCPDR	Transcription factor p65
DGVADVSIEDSVISLSGDHCIIGR	Superoxide dismutase [Cu-Zn]
DICNDVLSLLEK	14-3-3 protein zeta/delta
EDITEPAICALR	Catenin beta-1
EEAPDILCLQETK	DNA-(apurinic or apyrimidinic site) lyase
EQEMQECTTQIQYLK	Pre-mRNA-splicing regulator WTAP
FEDWLCNK	RNA-binding protein 5
GIFPVLCK	Pyruvate kinase PKM
GLCAIAQAESLR	40S ribosomal protein S3
GWPLELLCEK	Interleukin enhancer-binding factor 3
ICSWNVDGLR	DNA-(apurinic or apyrimidinic site) lyase
IEAELQDICNDVLELLDK	14-3-3 protein beta/alpha
ITFCTGIR	26S proteasome non-ATPase regulatory subunit 4
LCFQVTVR	Transcription factor p65
LDYFLLSHSLLPALCDSK	DNA-(apurinic or apyrimidinic site) lyase
LFQCLLHR	X-ray repair cross-complementing protein 5
LQAQQDAVNIVCHSK	26S proteasome non-ATPase regulatory subunit 4
LQGINCGPDFTPSFANLGR	Eukaryotic translation initiation factor 4 gamma 1
MADCGGLPQISQPAK	Protein NDRG1
MCLFAGFQR	Heterogeneous nuclear ribonucleoprotein U
NILEESLCELVAK	DNA-dependent protein kinase catalytic subunit
NLALCPANHAPLR	Catenin beta-1
NNTQVLINCR	Small nuclear ribonucleoprotein Sm D2
NSASWCVEWFNTK	E3 ubiquitin/ISG15 ligase TRIM25
NTCTSVYTK	DNA-dependent protein kinase catalytic subunit
NTGIICTIGPASR	Pyruvate kinase PKM
RPPSAFFLFCSEYHPK	Putative high mobility group protein B1-like 1
SFGSTCQLSEK	DNA topoisomerase 2-alpha

SICEVLDLER	Protein DEK
SLCVPYFVLELDQTEDGR	Thioredoxin reductase 1, cytoplasmic
SYLYQILQGIVFCHSR	Cyclin-dependent kinase 1
TAAQCLEHYEFLLDK	Cell division cycle 5-like protein
TCSPASLSQASADLEATLR	E3 ubiquitin/ISG15 ligase TRIM25
TQHLSNDHICYFLYQILR	Mitogen-activated protein kinase 1
VCNPIITK	Heat shock cognate 71 kDa protein
VNQIGSVTESLQACK	Alpha-enolase
WCSWSLSQAR	4F2 cell-surface antigen heavy chain
YMPQNPCIIATK	Histone-binding protein RBBP4