



HAL
open science

Développement d'une approche analytique non ciblée pour l'étude des protéines dans les milieux complexes, environnementaux et biologiques

Balkis Eddhif

► **To cite this version:**

Balkis Eddhif. Développement d'une approche analytique non ciblée pour l'étude des protéines dans les milieux complexes, environnementaux et biologiques. Chimie analytique. Université de Poitiers, 2017. Français. NNT : 2017POIT2291 . tel-01951522

HAL Id: tel-01951522

<https://theses.hal.science/tel-01951522>

Submitted on 11 Dec 2018

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

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

THESE

Pour l'obtention du Grade de

DOCTEUR DE L'UNIVERSITE DE POITIERS

(Faculté des Sciences Fondamentales et Appliquées)
(Diplôme National - Arrêté du 25 mai 2016)

Ecole Doctorale : Sciences pour l'environnement - Gay Lussac

Secteur de Recherche : Chimie théorique, physique, analytique.
Présentée par :

Balkis EDDHIF

Développement d'une approche analytique non ciblée pour l'étude des protéines dans les milieux complexes, environnementaux et biologiques

Directeur(s) de Thèse :

Nathalie KARPEL VEL LEITNER, Pauline POINOT et Claude GEFFROY-RODIER

Soutenue le 4 Décembre 2017

devant la Commission d'Examen

JURY

Karen GAUDIN	Professeur, Université de Bordeaux, Rapporteur
Emilie DESTANDAU	Maître de conférences, Université d'Orléans, Rapporteur
Eric MARCHIONI	Professeur, Université de Strasbourg, Examineur
Jean-Christophe GARRIGUES	Ingénieur de Recherche, CNRS, Université Paul Sabatier, Examineur
Pascal CARATO	Professeur, Université de Poitiers, Examineur
Nathalie KARPEL VEL LEITNER	Directrice de Recherche CNRS, Université de Poitiers, Directeur de thèse
Pauline POINOT	Maître de conférences, Université de Poitiers, Co-directeur de thèse
Claude GEFFROY-RODIER	Maître de conférences, Université de Poitiers, Co-directeur de thèse

Remerciements

*Du profond de mon cœur je dédie cette thèse à la mémoire de ma chère grand-mère
« Dadaya » disparue il y a peu*

Mes vifs remerciements s'adressent aux membres du jury. Merci au Pr. Karen Gaudin et au Dr. Emilie Destandau d'avoir accepté d'évaluer ce travail en qualité de rapporteurs. Merci au Pr. Eric Marchioni, au Dr. Jean-Christophe Garrigues et au Pr. Pascal Carato d'avoir accepté de participer au jury de cette thèse et d'évaluer ce travail.

Ces travaux ont été réalisés au sein de l'équipe Eaux, Biomarqueurs, Contaminants Organiques, Milieux (E-BiCOM) du laboratoire de l'IC2MP (UMR CNRS 7285) dirigé par Madame Sabine Petit que je remercie pour m'avoir accueillie dans cette structure.

Il me plaît particulièrement d'exprimer ma gratitude et toute ma reconnaissance envers mes co-directrices de thèse : Pauline Pointot et Claude Geffroy-Rodier pour leur dévouement, leur grande disponibilité et leur veille perpétuelle au bon déroulement de la thèse. Merci d'avoir été présentes pour moi aussi bien sur le plan académique que personnel. Je vous remercie également pour toutes les discussions fructueuses, pour les ondes positives que vous m'avez transmises et qui m'ont poussé au-delà de mes limites. Merci !!!!!

Un grand merci au Dr. Nathalie Karpel Vel Leitner d'avoir accepté de diriger ce travail de thèse.

Cette thèse ne serait probablement jamais accomplie sans l'aide de collaborateurs qui par leur expertise se sont engagés dans la réussite de ce travail. Un grand merci à Yan Batonneau pour sa précieuse contribution sur le traitement des résultats Raman. Tu m'as permis d'acquérir un savoir concernant cette technique qui m'était inconnue. Je tiens à remercier Nadia Guignard pour toutes les manipulations Raman. Merci pour ta gentillesse et ta disponibilité. Je suis également très reconnaissante envers Sébastien Papot pour m'avoir fait profiter de son expertise en chimie organique. Merci pour ton investissement pour l'aboutissement de ma publication. Mes remerciements les plus sincères vont à Jérôme Labanowski et Leslie Mondamert pour m'avoir fourni les échantillons de biofilm. Merci à toi Jérôme pour tes conseils et ton aide plus qu'enrichissante au niveau de la spectrofluorimétrie. Merci à vous deux d'avoir été également présents dans les moments de « pannes » techniques de l'orbitrap !

Je tiens à remercier Jonathan Clarhaut de m'avoir fourni les échantillons biologiques de souris, merci également pour ta contribution dans l'accomplissement de la publication. Un grand merci à Benoit Teychene de m'avoir fait découvrir l'AFM. Je remercie également Justin Lange pour les heures passées sur *Skyline* !! On y ait arrivé finalement ☺

Mes remerciements les plus chaleureux sont adressés aux différentes personnes travaillant dans le laboratoire : Merci à Laurent Lemée pour tes conseils, ce fut un plaisir de travailler avec toi. Merci à Laurent Grasset pour ta bonne humeur. Je remercie du fond du cœur Joëlle Laduranty pour sa disponibilité tout au long de ma thèse, pour sa gentillesse et sa bienveillance. Ma reconnaissance est également adressée à Jean François Chollet, Cécile Marivingt-Mounir, Emilie Dugaro, André Amblés et à Lillian Ripault.

Un grand merci à Sylvie Liu de m'avoir formée à l'utilisation de l'orbitrap. Merci pour ta générosité, ta sympathie et tes précieux conseil--s.

Un immense merci à Audrey Allavena, Bertrand Gombert, Marie Deborde, Thibaut Le Guet, Cristina Soïca, Elodie Laurent pour leurs conseils et leurs encouragements incessants.

Je tiens particulièrement à remercier Claude Fontaine pour toutes les manipulations et interprétations des résultats d'adsorption des protéines sur les argiles. Merci pour tes conseils et pour ta bonne humeur.

Je remercie chaleureusement tous mes collègues de travail : Sophie, Pascaline, Hanxiang Wu, Khaled, Thomas, Justin, Tiffany, Virginie, Florence, Kossounou, Yoan, pour leur sympathie, leur bonne humeur et pour les bons moments partagés. Merci également à tous les stagiaires ayant grandement contribué à la réussite de ce travail : Guillaume, Kevin et Alain.

Je remercie vivement tout le personnel administratif du laboratoire : Claire Guichard, Annie-Claude Granger, Nicole Mahoungou, Nathalie KOTRI.

Enfin, je tiens à remercier toutes les personnes qui ont été à mes côtés, qui m'ont soutenu et supporté tout au long de cette période. C'est à ma chère famille et mes amis que je m'adresse pour manifester mon amour et ma reconnaissance.

Je remercie d'abord mes très chers parents : ma maman Radhia et mon papa Lassaad, mes très chers frères Mongi et Moslem et ma belle-sœur Alia. Votre soutien émotionnel et moral m'ont permis de surmonter toutes les difficultés. Que Dieu vous bénisse.

Je remercie chaleureusement ma belle-mère Omria et mon beau-frère Jalloul. Merci d'avoir toujours cru en moi.

Un grand merci à toute ma famille : cousins et cousines (Lobna, Imen, Noussa...), oncles et tantes pour votre soutien quotidien.

Un immense merci à ma sœur de cœur Rachida Mersni pour tous les moments passés ensemble et surtout pour les heures de téléphone à parler de la thèse et de la mode surtout ☺ notre complicité et notre amour ne cessent de grandir avec le temps. Un grand merci à mon cher ami Oussama Achour pour sa bonne humeur et ses précieux conseils. Que Dieu protège votre petite famille ☺

Un gros merci également à mes amis Jojo, Bouziane, Mouheb, Ghizlaine, Amira, Ahmed, Lobna, Imen, Mohammed, pour leur soutien moral et les fous rires que nous avons partagés. Je remercie particulièrement ma meilleure amie Hajoura. Tu étais loin des yeux, mais tout le temps proche de mon cœur. Je te remercie d'avoir toujours cru en moi.

Enfin, un énorme merci à mon mari chéri Amine. Merci pour ton soutien, ta tendresse et ton amour !!

Table des matières

Liste des figures	10
Liste des tableaux	17
Liste des abréviations	18
Introduction générale	19
Introduction bibliographique	22
1. Les protéines : structures et fonctions	24
1.1. Structure des protéines.....	24
1.1.1. Les acides aminés.....	24
1.1.2. La liaison peptidique	26
1.1.3. Conformation des protéines	27
1.1.3.1. Structure secondaire.....	28
1.1.3.2. Structure tertiaire	34
1.1.3.3. Structure quaternaire	35
1.1.4. Stabilité conformationnelle des protéines	35
1.1.4.1. Stabilité de la structure native de la protéine	35
1.1.4.2. Influence du milieu sur l'arrangement des protéines	37
1.1.4.3. Dénaturation des protéines.....	38
1.2. Fonctions des protéines	39
1.2.1. Fonction de défense.....	42
1.2.2. Fonction catalytique	43
1.2.3. Fonction structurale.....	44
2. Stratégie générale en protéomique	45
2.1. La protéomique.....	45
2.2. Recherche des biomarqueurs protéiques	47
2.2.1. Intérêts scientifiques et défis associés	47
2.2.2. Dans le domaine de la santé	49
2.2.3. Dans le domaine de l'environnement.....	52
Chapitre II : De la préparation d'échantillon à l'analyse des protéines	55
1. Préparation de l'échantillon	58
1.1. Extraction des protéines.....	58
1.1.1. Extraction solide/liquide	58
1.1.1.1. Principe général	59
1.1.1.2. Applications.....	59
1.1.1.3. Verrous	62

1.1.2.	Extraction par immunoextraction.....	64
1.1.2.1.	Principe.....	64
1.1.2.2.	Applications.....	65
1.1.2.3.	Verrous	66
1.2.	Purification des protéines	66
1.2.1.	Précipitation, lavage et solubilisation	66
1.2.2.	Extraction sur phase solide (SPE).....	70
1.3.	Hydrolyse enzymatique des protéines	72
1.4.	Séparation des protéines	73
1.4.1.	Electrophorèse bidimensionnelle (2D).....	73
1.4.2.	Techniques séparatives « hors gel » : la chromatographie liquide.....	76
1.4.2.1.	La chromatographie Liquide à Très Haute Pression (UHPLC)	79
1.4.2.2.	La nanochromatographie	79
1.4.2.3.	La Chromatographie multidimensionnelle (MDLC : Multidimensional Liquid Chromatography).....	80
2.	Analyse des protéines par spectrométrie de masse	85
2.1.	Principe et description	85
2.2.	Les sources d'ionisation.....	85
2.2.1.	Ionisation électrospray (ESI)	86
2.2.2.	Ionisation désorption laser assistée par matrice (Matrix Assisted Laser Desorption Ionisation, MALDI).....	90
2.3.	Les analyseurs de masses	93
2.3.1.	Les analyseurs en tandem.....	94
2.3.1.1.	Le triple quadripôle.....	95
2.3.1.2.	Le Q-TOF.....	96
2.3.1.3.	Le Q-exactive.....	97
2.4.	Interprétation des analyses de protéines par spectrométrie en tandem	102
2.5.	Stratégies de quantification.....	103
3.	Autres méthodes d'analyse des protéines.....	105
3.1.	Méthodes immunologiques.....	105
3.2.	Détection des protéines en microscopie à fluorescence	106
3.3.	Etude de la structure des protéines	107
3.3.1.	La cristallographie aux rayons X	107
3.3.2.	Spectroscopie RMN	108
3.3.3.	Spectrofluorométrie	108
3.3.3.1.	Fluorescence intrinsèque	108

3.3.3.2.	Fluorescence extrinsèque	110
3.3.4.	Spectroscopie Raman.....	111
Objectifs et stratégies expérimentales		115
Partie expérimentale		121
1.	Matériels.....	122
1.1.	Protéines, peptides et acides aminés standards.....	122
1.2.	Matrices environnementales et biologiques étudiées	122
1.3.	Tampons et solvants utilisés pour la préparation d'échantillon	123
1.4.	Réactifs et tampons utilisés pour la digestion trypsique et l'analyse UPLC-MS/MS 124	
1.5.	Réactifs et tampons utilisés pour la spectrofluorométrie.....	124
2.	Méthodes	124
2.1.	Analyse UPLC-MS/MS.....	124
2.1.1.	Paramètres UPLC-MS/MS	124
2.1.1.1.	Mode d'analyse direct.....	124
2.1.1.2.	Système « <i>dual trap</i> »	125
2.1.2.	Analyse par spectrométrie de masse haute résolution (UPLC-HRMS/MS)	126
2.1.3.	Analyses bio-informatiques.....	127
2.2.	Optimisation de la digestion trypsique	127
2.3.	Mise au point d'un protocole de purification des protéines	128
2.3.1.	Quantification des protéines par dosage de Bradford et électrophorèse monodimensionnelle (1D)	128
2.3.2.	Quantification des protéines par UPLC-HRMS/MS.....	129
2.3.3.	Analyses statistiques	129
2.3.4.	Spectrofluorométrie.....	129
2.3.5.	Micro-spectroscopie Raman.....	130
2.3.6.	La microscopie à force atomique (AFM).....	131
2.4.	Comparaison des protocoles de purification sur matrices réelles.....	131
2.4.1.	Extraction et purification des protéines du sol.....	132
2.4.2.	Extraction et purification des protéines de biofilm	133
2.4.3.	Extraction et purification des protéines du foie de souris	134
2.5.	Analyse quantitative	135
2.5.1.	Étalonnage externe	135
2.5.2.	Quantification relative par le logiciel Skyline des protéines issues d'échantillons complexes	135
2.6.	Etudes des paramètres influençant l'extraction des protéines	135

2.6.1.	Extraction sur une matrice simple : sable de Fontainebleau dopé	136
2.6.1.1.	Effet de la filtration	136
2.6.1.2.	Effet du solvant	136
2.6.2.	Extraction sur une matrice complexe : Sol du site hydrogéologique du Deffend dopé	137
2.6.2.1.	Effet de la température d'extraction	137
2.6.2.2.	Effet du temps d'agitation.....	138
2.6.2.3.	Effet de la méthode d'extraction.....	138
2.6.2.4.	Effet de la soude.....	139
2.6.2.5.	Effet de nouveaux agents chimiques et biochimiques sur l'extraction des protéines	139
2.6.2.6.	Extraction séquentielle : MgCl ₂ - acides aminés/DMSO - TMAH/DMSO	140
2.6.3.	Extraction à partir d'une matrice biologique : xénogreffe de tumeur	140
Résultats et discussions		142
Chapitre I : Développement analytique		143
1.	Optimisation des paramètres chromatographiques : mise en place d'un système de préconcentration en ligne	145
2.	Identification des protéines cibles par spectrométrie de masse.....	153
3.	Quantification des protéines.....	155
3.1.	Quantification absolue	155
3.2.	Quantification relative	158
4.	Optimisation de la digestion trypsique.....	161
5.	Conclusion.....	164
Chapitre II : Développement d'une approche simple pour la purification des protéines		165
1.	Étude préliminaire	167
2.	Compréhension moléculaire de l'effet de l'éthanol/HCl sur la déstabilisation du <i>molten globule</i>	173
2.1.	Spectrofluorométrie	174
2.2.	Spectroscopie Raman	175
2.3.	Microscopie à force atomique (AFM)	183
3.	Etude de l'efficacité du protocole de purification sur matrices réelles	186
3.1.	Etude sur des protéines standards déposées sur matrice.....	186
3.2.	Analyse des protéines endogènes aux matrices	189
3.3.	Quantification des protéines endogènes <i>via Skyline</i>	200
4.	Conclusion.....	202

Chapitre III : Etude des paramètres influençant l'extraction des protéines	203
1.Essais sur matrices dopées	205
1.1. Développement sur matrice modèle : sable	205
1.1.1. Optimisation de l'étape de filtration post-extraction.....	205
1.1.2. Optimisation du solvant d'extraction	207
1.2. Extraction à partir d'une matrice environnementale : d'un échantillon de sol dopé (site hydrogéologique du Deffend).....	210
1.2.1. Études des paramètres influençant l'extraction des protéines.....	211
1.2.1.1. Paramètres physiques et mécaniques	211
1.2.1.2. Paramètres chimiques	214
1.2.1.3. Extraction séquentielle : Effets combinés.....	218
2.Extraction des protéines endogènes dans des matrices environnementales et biologiques	220
2.1. Matrice environnementale : sol du Deffend	220
2.2. Matrice biologique : tumeur de souris	220
3.Conclusion.....	222
Conclusion générale et perspectives	223
Annexes	226
Annexe 1: Détermination de la quantité totale des protéines par électrophorèse 1D et dosage Bradford	227
Annexe 2: Spectres d'émission de la HSA, lysozyme et Rnase après précipitation au TCA et lavage à l'éthanol/HCl.....	228
Annexe 3: Liste des peptides endogènes du biofilm quantifiés par Skyline avec les transitions PRM correspondantes	230
Annexe 4: Extraction des protéines cibles à partir d'une matrice sable : effet d'un mélange de DMSO/Tris-HCl.....	234
Annexe 5: Effet de la température sur l'extraction des protéines	235
Annexe 6: Effet du temps sur l'extraction des protéines	236
Annexe 7: Effet des ultrasons sur l'extraction des protéines	237
Annexe 8: Adsorption des protéines sur les minéraux argileux.....	238
1.Analyse par diffraction aux rayons X (DRX)	238
1.2. Conditions expérimentales	238
1.3. Résultats et discussions	239
Références bibliographiques	243
Contributions scientifiques	269
Résumés	272

Liste des figures

Figure 1. Représentation schématique de la biosynthèse des protéines au sein d'une cellule eucaryote	24
Figure 2. Structure d'un acide aminé. Le carbone central C _α est lié à un groupement amine NH ₂ , à un groupement carboxyle C'OOH, à un atome d'hydrogène et à une chaîne latérale R	25
Figure 3. Diagramme de Venn représentant les acides aminés classés en fonction de leurs propriétés physicochimiques (d'après (Faure et al., 2008))	26
Figure 4. Liaison peptidique entre deux acides aminés (dipeptide)	27
Figure 5. Représentation des quatre niveaux structuraux des protéines	28
Figure 6. Trajectoire de l'hélice α illustrant l'orientation des dipôles peptidiques ; l'extrémité positive du dipôle est localisée au niveau de l'extrémité N-terminale et l'extrémité négative se trouve à l'extrémité C-terminale	29
Figure 7. Structure de la myoglobine (la molécule d'oxygène se fixe à la protéine grâce à l'hème) (PDB, 1A6K)	30
Figure 8. a. Structure de la ferritine isolée de la rate de cheval (Hempstead et al., 1997) (PDB, 1AEW) ; b. Structure de la HSA complexée à l'ibuprofène (PDB, 2BXG)	31
Figure 9. a. Structure des feuillet β (Lowik et al., 2010) ; b. Structure de la ribonucléase bovine, constituée d'une prédominance de feuillet β (PDB, 8RAT)	31
Figure 10. Représentation de la structure de la thiorédoxine d'Escherichia coli (Noguera et al., 2017)	32
Figure 11. Structure de la protéine de transport FhuA d'Escherichia coli composée entièrement de feuillet β complexée à un dérivé de la rifamycine (antibiotique) (Ferguson et al., 2001) .	32
Figure 12. Exemple de représentation de boucles (cercles bleus) au sein de structures secondaires répétitives de la nucléase de Staphylococcus aureus (PDB 1EQV)	33
Figure 13. À gauche : représentation schématique d'un anticorps IgG ; à droite : Illustration du pli immunoglobuline avec les boucles à la surface (en rouge), correspondant aux sites de fixation de l'antigène. L'illustration de l'anticorps a été réalisée par Irving Geis/Geis Archives Trust (d'après Voet and Voet, 2016)	34
Figure 14. Exemple de structure quaternaire de la ribulose-1,5-bisphosphate carboxylase/oxygénase (Spinacia oleracea (plante)) (PDB, 1RCX)	35
Figure 15. Illustration du dépôt des plaques amyloïdes dans le cerveau humain, mettant en évidence la formation de feuillet β (la structure secondaire du β-amyloïde est tirée de : (Xiao et al., 2015))	36
Figure 16. Représentation schématique des interactions établies entre une protéine (régions hydrophiles/hydrophobes) et son environnement (hydrophile/hydrophobe) (d'après Marchin and Berrie, 2003)	38
Figure 17 Représentation de la structure du dodécylsulfate de sodium (SDS)	39
Figure 18. Représentation schématique de quelques fonctions des protéines	40
Figure 19. Représentation schématique des étapes de polymérisation des microtubules (d'après Pampaloni and Florin, 2008)	41

Figure 20. Micrographie du fuseau mitotique d'une cellule animale après marquage par des fluorescents spécifiques. Les tubulines sont colorées en vert et l'ADN en bleu (d'après O'Connell and Khodjakov, 2007)	41
Figure 21. Représentation schématique d'une interaction anticorps-antigène (Ac-Ag). Les lymphocytes B sont des cellules participant à la réponse immunitaire, elles produisent des anticorps qui vont servir à la reconnaissance spécifique (Ac-Ag)	42
Figure 22. Structure de la laccase de <i>Trametes versicolor</i> . Le site actif est formé d'atomes de cuivre (PDB, 1GYC)	43
Figure 23. Effet des substances chimiques sur l'activité enzymatique de la ribonucléase bovine (d'après Berg et al., 2012).....	44
Figure 24. Représentation schématique du déplacement de la myosine le long des filaments d'actine (d'après Kneussel and Wagner, 2013)	45
Figure 25. De la génomique vers la protéomique (Les différents organites de la cellule eucaryote ne sont pas illustrés)	46
Figure 26. Illustration graphique de l'abondance des différentes protéines du plasma : les points jaunes représentent les biomarqueurs potentiels et sont situés essentiellement dans la gamme relative aux faibles concentrations (d'après Schiess et al., 2009).....	49
Figure 27. Représentation de la stratégie « top-down » et « bottom-up »	57
Figure 28. Approche ascendante « bottom-up » pour l'identification des protéines par spectrométrie haute résolution (HRMS/MS).....	58
Figure 29. Extraction des protéines en présence de détergents anioniques	59
Figure 30. Exemples de méthodes d'extraction des protéines par solvant.....	61
Figure 31. Détection sur gel d'électrophorèse bidimensionnelle des protéines les plus abondantes dans le plasma humain (Jiang et al., 2004)	63
Figure 32. Complexation des enzymes extracellulaires aux composés de sols (Nannipieri et al., 1996).....	64
Figure 33. Représentation schématique du principe de l'immunoextraction sur billes magnétiques et sur support (Płotka-Wasyłka et al., 2015).....	65
Figure 34. Purification des protéines à partir de matrices environnementales ou biologiques	69
Figure 35. Préparation et application du support Ti^{4+} -IMAC ; a. Représentation schématique des étapes d'assemblage du dispositif Ti^{4+} -IMAC ; b. Centrifugation assistée par une extraction sur phase solide au moyen d'un Ti^{4+} -IMAC (Yao et al., 2017)	71
Figure 36. Digestion typique d'une protéine.....	72
Figure 37. Schéma d'une analyse protéomique basée sur l'électrophorèse bidimensionnelle (Chevalier, 2010; Rabilloud and Lelong, 2011).....	74
Figure 38. Séparation des protéines d' <i>E.coli</i> par électrophorèse bidimensionnelle. 1100 protéines ont été séparées. Les cellules bactériennes ont été lysées par sonication puis traitées avec de la Dnase et de la Rnase avant d'être solubilisées dans un tampon de lyse. 25 μ l d'échantillons ont été déposés sur le gel d'électrophorèse (O'Farrell, 1975).....	75
Figure 39. a. Représentation schématique des interactions établies entre les différents sites des peptides tryptiques et les phases stationnaires utilisées en chromatographie liquide : RP : phase inverse ; AX : échange d'anions ; CX : échange de cations ; n-HILIC : interaction hydrophile-neutre ; ERLIC : répulsion électrostatique-interaction hydrophile ; z-HILIC : interaction hydrophile-zwitterionique. b. Représentation schématique de la sélectivité des phases	

stationnaires vis-à-vis des charges positives, négatives, groupements apolaires et polaires. Le cercle interne/gris clair : faible sélectivité ; les cercles externes/gris : forte sélectivité (Di Palma et al., 2012).....	77
Figure 40. Représentation d'une phase en carbone graphité poreux (Hypercarb) poreux et d'une phase monolithique « Chromolith® » commercialisée par Merck	79
Figure 41. Principe d'une chromatographie liquide 2D (Nägele et al., 2004)	81
Figure 42. Représentation schématique du système multidimensionnel développé par Wang et al. (2010)	82
Figure 43. Représentation générale d'un spectromètre de masse	85
Figure 44. Représentation schématique de la source à l'électrospray en mode positif (Cech and Enke, 2001)	87
Figure 45. Représentation schématique du mécanisme d'ionisation-désolvatation ; a. modèle de l'évaporation ionique, b. modèle de la charge résiduelle (Wilm, 2011).	88
Figure 46. a. Chromatogramme du peptide 'SQTECDIYPLR' issu de la digestion trypsique de la β -glucuronidase; b. spectre de masse (MS) du peptide ($[M+2H]^{2+}$, m/z= 691,32367 et du $[M+3H]^{3+}$, m/z= 461,21829)	89
Figure 47. Représentation schématique du principe de la DESI (Ifa et al., 2010).....	90
Figure 48. Principe de l'ionisation MALDI (Di Girolamo et al., 2013)	91
Figure 49. Spectre de masse de la protéine (entérotoxine B staphylococcique) obtenu par MALDI-TOF/MS (85 fmol ou 2,5 ng). Les pics correspondent à l'ion moléculaire de la protéine mono et di-chargée (Schlosser et al., 2007)	92
Figure 50. Représentation schématique du processus du SELDI (Di Girolamo et al., 2013)..	93
Figure 51. Représentation schématique de la résolution d'un analyseur de masse.....	93
Figure 52. Analyse en tandem MS/MS	94
Figure 53. Représentation schématique du triple quadripôle (Mode SRM) (d'après Aebersold and Mann, 2016)	96
Figure 54. Représentation schématique du spectromètre hybride Q-TOF (Lacorte and Fernandez-Alba, 2006).....	97
Figure 55. Représentation du Q-exactive (d'après Thermo Fisher).....	98
Figure 56. Représentation schématique de l'analyseur orbitrap (Zubarev and Makarov, 2013)	99
Figure 57. Principe du mode PRM (Aebersold and Mann, 2016).....	100
Figure 58. Les différents modes d'acquisition des données protéomiques. L'analyse des données est réalisée au moyen de logiciels dédiés à chaque mode d'acquisition (d'après Aebersold and Mann, 2016).....	101
Figure 59. Nomenclature des ions peptidiques générés par fragmentation MS/MS.....	102
Figure 60. Exemple de spectre de fragmentation (Rinalducci et al., 2009)	103
Figure 61. Représentation schématique de l'effet matrice sur l'ionisation des analytes par électrospray (Panuwet et al., 2016)	105
Figure 62. Micrographie d'une cellule animale après marquage par des fluorescents spécifiques. Les filaments d'actine et les microtubules sont colorés respectivement en rouge et en vert (Risler, 2009). Source: Mark Shipman, James Blyth and Louise Cramer, MRC-Laboratory Molecular Cell Biology and Cell Biology Unit, UCL, London UK	106

Figure 63. Principe de la cristallographie aux rayons X, la source de rayons X produit un faisceau incident qui sera diffracté par le cristal. Le spectre de diffraction est ensuite recueilli par un détecteur	107
Figure 64. Représentation schématique des différentes zones de fluorophores; a : zone caractéristique des protéines (acides aminés aromatiques) ; b : zone fluorophore caractéristique des acides nucléiques ; c : zone fluorophore caractéristique des sous-produits protéiques microbiens ; d : zone fluorophore caractéristique des acides humiques (Chen et al., 2003) .	109
Figure 65. Spectre de fluorescence des acides aminés aromatiques dilués dans un tampon phosphate à pH 7, λ_{ex} (nm) = 270 nm (Albani JR, 2001).....	110
Figure 66. Spectre de fluorescence de l'ANS dans différents environnements : polaires et apolaires. L'intensité du signal de l'ANS augmente en fonction de l'hydrophobicité du milieu	111
Figure 67. Représentation schématique de la diffusion de la lumière d'un échantillon éclairé par une source laser. Lorsqu'une partie des photons est diffusée avec une fréquence inférieure à la fréquence incidente, on parle de diffusion Stokes. A l'inverse, si les photons sont diffusés avec une fréquence supérieure à la fréquence incidente, la diffusion Raman est de type anti-Stokes	112
Figure 68. Spectre Raman des bandes amides de deux protéines solubilisées dans le H ₂ O et le D ₂ O ; L : Lysozyme ; C : Chymotrypsine. L'attribution des différentes bandes de vibration est comme suit : B1h (940 cm ⁻¹) : squelette N-C α -C (Mode d'étirement relatif à l'hélice α), AmIIIb'-f (960 cm ⁻¹) : Amide III' feuillet- β de l'échange isotopique, AmIIIa'-f (980 cm ⁻¹) : Amide III' feuillet- β de l'échange isotopique, AmIIIa-f (1240 cm ⁻¹) : Amide III feuillet- β , AmIII-h (1340 cm ⁻¹) : Amide III hélice- α , AmI (1650-1660 cm ⁻¹) : Amide I (Hedoux et al., 2013).....	113
Figure 69. Caractéristiques physicochimiques des cibles protéiques	118
Figure 70. Représentation de la carte de localisation de Poitiers et carte géologique du site	122
Figure 71. Traitements successifs appliqués à la protéine sous forme native.....	130
Figure 72. Traitements successifs appliqués à la protéine native	131
Figure 73. Comparaison de la purification des protéines à partir de matrices complexes.....	132
Figure 74. Photo du sonificateur utilisé lors de cette étude (sonde de diamètre de 2 mm)....	138
Figure 75. Stratégie expérimentale.....	145
Figure 76. Représentation schématique du système chromatographique ; A. Injection, B. Préconcentration trapping, C. Elution	146
Figure 77. Analyse en Composantes Principales (ACP) basée sur le rendement de rétention (analytique) des peptides (ratio d'aire du peptide après préconcentration par rapport à l'aire du peptide sans préconcentration (injection directe)) ; a. Cercle des corrélations ; b. Graphique des observations. Codes des différentes configurations : nom de la colonne-débit de chargement-temps de chargement-temps du backflush. Le terme « series » désigne l'arrangement en série des colonnes de préconcentration, Hypersil Gold aQ (1) et Hypercarb (2) connectées à la colonne analytique Hypersil Gold Q.....	151
Figure 78. Courant ionique total (m/z 300-2000) d'une solution de peptides standards analysés par UPLC-MS/MS en mode « Full scan » avec une résolution égale à 70 000. La concentration injectée dans les deux systèmes est de 10 ng.mL ⁻¹ . 1. gly-tyr ; 2. val-tyr-val ; 3. met-enképhaline ; 4. Angiotensine II ; 5. leu-enképhaline.....	153

Figure 79. Exemple de spectre de masse MS et MS/MS pour peptide HSA_599-609.....	155
Figure 80. Exemple de quantification relative du lysozyme à partir de matrices complexes, obtenue par le logiciel Skyline. a. Quantification sur l'ion précurseur NTDGSTDYGILQINSR ($m/z=877,4212$ $[MH]^{2+}$) et sur ses ions fragments y8, y6 et y5 ; b. Quantification sur le précurseur GTDVQAWIR ($m/z=523,2749$ $[MH]^{2+}$) et sur ses ions fragments y5, y4 et y3 .	160
Figure 81. Effet du tampon de digestion sur l'efficacité de la digestion de l'angiotensine II. Le ratio trypsine/substrat est fixé à 1/6.....	162
Figure 82. Effet du ratio E/S sur l'efficacité de la digestion de l'angiotensine II.....	163
Figure 83. Stratégie expérimentale.....	167
Figure 84. Electrophorèse 1D d'une solution de protéines standards (HSA, lysozyme, Rnase et exoglucanase-1) précipitée au TCA et lavage avec les différents solvants (éthanol, acétone et éthanol/HCl). La révélation des bandes protéiques a été réalisée par coloration au nitrate d'argent. Les protéines ont été également quantifiées par la méthode de Bradford	170
Figure 85. Stratégie d'analyse	171
Figure 86. Comparaison des rendements de récupération des protéines après précipitation au TCA et lavage avec les différents solvants. * pourcentage de protéines insolubles =100 % - % protéines dans le surnageant - % protéines dans le tampon d'analyse.....	172
Figure 87. Représentation du spectre Raman du lysozyme à différentes étapes de purification (entre 1200-1800 cm^{-1}). a. Lysozyme natif (bleu); b. Lysozyme précipité au TCA (rouge); c. Lysozyme lavé à l'éthanol/HCl (noir); Lysozyme dans le tampon Tris-bicarbonate d'ammonium (vert)	176
Figure 88. 1. Analyse Raman de la bande amide I (bandes de 1-4) à différentes étapes de purification. a. Lysozyme natif (bleu); b. Lysozyme précipité au TCA (rouge); c. Lysozyme lavé à l'éthanol/HCl (noir) ; d. Lysozyme dans le tampon Tris-bicarbonate d'ammonium (vert); 2. Distribution des différentes contributions (bandes de 1-4) à différentes étapes de purification	177
Figure 89. Représentation du spectre Raman du HSA à différentes étapes de purification (entre 1200-1800 cm^{-1}). a. HSA native (bleu); b. HSA précipitée au TCA (rouge); c. HSA lavée à l'éthanol/HCl (noir) ; HSA dans le tampon Tris-bicarbonate d'ammonium (vert).....	179
Figure 90. 1. Analyse Raman de la bande amide I (bandes de 6-9) à différentes étapes de purification. a. HSA native (bleu); b. HSA précipitée au TCA (rouge); c. HSA lavée à l'éthanol/HCl (noir);d. HSA dans le tampon Tris-bicarbonate d'ammonium (vert). 2. Distribution des différentes contributions (bandes de 6-9) à différentes étapes de purification	180
Figure 91. Représentation du spectre Raman la Rnase à différentes étapes de purification (entre 1200-1800 cm^{-1}). a. Rnase native (noir); b. Rnase précipitée au TCA (rouge); c. Rnase lavée à l'éthanol/HCl (bleu) ;d. Rnase dans le tampon Tris-bicarbonate d'ammonium (vert)	181
Figure 92. 1. Analyse Raman de la bande amide I (bandes de 10-13) à différentes étapes de purification. a. Rnase native (bleu); b. Rnase précipitée au TCA (rouge); c. Rnase lavée à l'éthanol/HCl (noir); d. Rnase dans le tampon Tris-bicarbonate d'ammonium (vert). 2. Distribution des différentes contributions (bandes de 10-13) à différentes étapes de purification	182

Figure 93. Représentation du spectre Raman (région de l'amide I). 1.a Lysozyme précipitée au TCA (rouge); b. Lysozyme lavée à l'acétone (gris). 2.a HSA précipitée au TCA (rouge); b. HSA lavée à l'acétone (gris)	183
Figure 94. Image AFM de la HSA à différentes étapes de purification.....	185
Figure 95. Récupération des protéines dopées en fonction des deux approches de purification : purification à l'éthanol/HCl et purification selon le protocole de la littérature. *P<0.1, **P<0.05 and ***P<0.01; analyse de variance (ANOVA) avec analyse de la petite différence significative (test LSD, Least Significant Difference) ($\alpha < 5\%$)	187
Figure 96. Récupération des protéines dopées en fonction du nombre de lavages à l'éthanol/HCl	188
Figure 97. Représentation schématique du pourcentage des protéines communes identifiées dans le biofilm et le foie de souris après purification soit avec l'éthanol/HCl soit avec le protocole publié dans la littérature	196
Figure 98. Répartition des protéines endogènes dans les différents compartiments cellulaire (noyau, membrane, mitochondrie, etc.), en fonction du protocole développé (EtOH/HCl : éthanol/HCl) et du protocole publié dans la littérature. a. Répartition cellulaire dans la matrice biofilm marin ; b. Répartition cellulaire dans la matrice de foie de souris	198
Figure 99. Quantification relative des protéines endogènes en fonction du protocole développé (EtOH/HCl) et du protocole publié dans la littérature. Les aires de pics ont été extraites des spectres MS2 correspondants aux ions fragments puis sommées pour générer les aires des ions précurseurs. La contribution de chaque ion peptide est reportée dans l'histogramme par différentes couleurs. Les protéines marquées par des astérisques (*) représentent les membranaires. La liste des transitions (couple ion précurseur/ions fragments) générée par Skyline est détaillée dans les tableaux 16 et 17 (cf. annexe 3), correspondants respectivement à la matrice biofilm, foie de souris.....	201
Figure 100. Rôle de l'éthanol/HCl dans la déstabilisation de l'état molten globule.....	202
Figure 101. Stratégie expérimentale.....	204
Figure 102. Effet de la filtration sur la récupération des protéines cibles.....	206
Figure 103. Effet du solvant sur la récupération des protéines pour une extraction séquentielle	209
Figure 104. Développement d'une bulle de cavitation à la surface d'un solide	212
Figure 105. Effet de la soude sur le rendement de récupération des protéines cibles (agitation 2h00, à 25 °C).....	214
Figure 106. Photos de trois tubes eppendorfs après extraction des protéines cibles à partir d'un	215
Figure 107. Effet des différents agents chimiques (TMAH, MgCl ₂) et biochimiques (acides aminés basiques) sur le rendement d'extraction des protéines	217
Figure 108.Extraction séquentielle : effet du TMAH	218
Figure 109. Extraction séquentielle : MgCl ₂ -Acides aminés/DMSO-TMAH-DMSO	219
Figure 110. Diagramme de Venn représentant le nombre de protéines communes et différentes identifiées dans chaque extrait	221
Figure 111. Electrophorèse 1D d'une solution de protéines standards (HSA, lysozyme, Rnase et exoglucanase-1) précipitée au TCA et lavée avec les différents solvants (éthanol/HCl 1,25 M, acétone/HCl 0,06 M et acétone/HCl 1,25 M). La révélation des bandes protéiques a été	

réalisée au nitrate d'argent. La quantité totale de protéines a également été déterminée par un dosage Bradford	227
Figure 112. Spectre d'émission du lysozyme-ANS ($\lambda_{exc}= 400$ nm) à différentes étapes de purification	228
Figure 113. Spectre d'émission du HSA-ANS ($\lambda_{exc}= 400$ nm) à différentes étapes de purification	228
Figure 114. Spectre d'émission de la Rnase-ANS ($\lambda_{exc}= 400$ nm) à différentes étapes de purification	229
Figure 115. Extraction des protéines avec un mélange DMSO/Tris-HCl.....	234
Figure 116. Effet de la température d'extraction (le temps d'extraction a été fixé à 5h00) ..	235
Figure 117. Effet du temps d'extraction (la température d'extraction a été fixée à 25 °C) ...	236
Figure 118. Effet d'une extraction assistée par ultrasons sur la récupération des protéines. Les deux histogrammes représentent les mêmes résultats, a. histogramme empilé ; b. histogramme mettant en évidence les barres d'erreurs pour chaque méthode.....	237
Figure 119. Diffractogrammes de l'échantillon Kaolinite préparé sur lame orientée séchée à l'air ; (1) témoin : après contact à l'eau ultrapure ; (2) en présence d'une solution de HSA à 1 mg.mL ⁻¹ ; (3) en présence d'une solution de lysozyme à 1 mg.mL ⁻¹ . Minéraux : (A) Anatase, (K) Kaolinite, (M) Mica	241
Figure 120. Diffractogrammes de l'échantillon montmorillonite préparé sur lame orientée séchée à l'air ; (1) témoin : après contact à l'eau ultrapure ; (2) en présence d'une solution de HSA à 1 mg.mL ⁻¹ ; (3) en présence d'une solution de lysozyme à 1 mg.mL ⁻¹ . Minéraux : (F) Feldspath, (Q) Quartz, (M) Mica, (S) Smectite.....	241

Liste des tableaux

Tableau 1. Exemples de biomarqueurs protéiques essentiels pour le diagnostic médical	50
Tableau 2. Exemples de biomarqueurs protéiques essentiels pour le diagnostic environnemental	53
Tableau 3. Comparaison des caractéristiques techniques de l'UPLC, de la NanoLC par rapport à l'HPLC (Gumustas et al., 2013; Wilson et al., 2015).....	80
Tableau 4. Exemples de séparation des peptides par chromatographie liquide (1D et 2D).....	83
Tableau 5. Performances techniques des principaux analyseurs de masse (Holčapek et al., 2012)	95
Tableau 6. Données spectrales du tryptophane, tyrosine et phénylalanine dans une solution d'eau à pH neutre (Ghisaidoobe and Chung, 2014)	109
Tableau 7. Conditions opératoires.....	131
Tableau 8. Caractéristiques des cibles peptidiques	148
Tableau 9. Analyse multiple de la variance (MANOVA - n = 3 - niveau de confiance à 95%) en fonction sur le rendement analytique des oligopeptides dans les 21 configurations LC...	148
Tableau 10. Liste des peptides identifiés après analyse UPLC-HRMS/MS après digestion trypsique des quatre protéines cibles. Nomenclature : LYSO= lysozyme ; RNASE= Ribonucléase ; HSA= Albumine humaine sérique ; EXO = exoglucanase-1. t _r (min) = temps de rétention en min.....	156
Tableau 11. Évaluation de la quantité de HSA présente dans le surnageant. Les valeurs n'ayant pas la même lettre sont significativement différentes ; NI : non identifiée.....	168
Tableau 12. Propriétés spectroscopiques de la HSA, lysozyme et Rnase après précipitation au TCA et lavage à l'éthanol/HCl.....	174
Tableau 13. Nombre de peptides et de protéines identifiés dans les matrices complexes	190
Tableau 14. Liste des protéines endogènes communes identifiées dans chaque matrice après purification selon le protocole développé (éthanol/HCl) et le protocole publié dans la littérature	190
Tableau 15. Caractérisation du sol du Deffend (Célérier, 2008)	210
Tableau 16. Liste des peptides endogènes du biofilm quantifiés par Skyline avec les transitions PRM correspondantes	230
Tableau 17. Liste des peptides endogènes du foie de souris quantifiés par Skyline avec les transitions PRM correspondantes.....	231

Liste des abréviations

UPLC-HRMS/MS : Chromatographe liquide ultra performance-spectrométrie de masse en tandem haute résolution

RPLC : Chromatographie en phase inverse

HILIC : Chromatographie à interaction hydrophile

Nano-LC : Nano-chromatographie liquide

ESI : Ionisation par électrospray

HSA : Albumine humaine sérique

Rnase : Ribonucléase

PDB : Protein Data Bank

SDS : Dodécylsulfate de sodium

CHAPS : 3-[(3-cholamidopropyl)diméthylammonio]-1-propanesulfonate

ARN : Acide ribonucléique

ADN : Acide désoxyribonucléique

TCA : Acide trichloroacétique

NaCl : Chlorure de sodium

DTT : Dithiothréitol

IAA : Iodoacétamide

ANS : 1-anilino-8-naphthalène sulfonate

DMSO : Diméthylsulfoxyde

Tris-HCl : Trishydroxyméthyl aminométhane-acide chlorhydrique

NaOH : Hydroxyde de sodium

MgCl₂ : Chlorure de magnésium

TMAH : Hydroxyde de tetraméthylammonium

ppm : Partie par million

Da : Dalton

mL: Millilitre

g : Gramme

s : seconde

min : minute

h : heure

t_r : temps de rétention

Introduction générale

Introduction générale

Les protéines sont responsables du maintien de l'équilibre vital des organismes vivants uni ou pluricellulaires. En effet, elles représentent à la fois les briques élémentaires et les machines moléculaires des cellules. Elles s'organisent pour créer la structure des cellules (le collagène des os et des tendons ou la kératine des cheveux), assurent des fonctions très variées et aussi remarquables qui se déclinent en fonction du rôle biochimique accompli (catalyse, défense, etc.). Cette remarquable diversité de fonctions découle principalement de la structure de la protéine, qui est déterminée à son tour par la séquence d'acides aminés. Tout dysfonctionnement au niveau de la structure et donc de la fonction des protéines est susceptible de causer de graves pathologies.

L'intérêt grandissant pour l'étude des protéines a donné lieu à une nouvelle science : la protéomique qui étudie l'ensemble des protéines, exprimées par une cellule ou un tissu, ou présentes dans un milieu à un moment donné et dans des conditions données.

Depuis une vingtaine d'années, l'analyse protéomique a marqué indéniablement un tournant dans de nombreux domaines de recherches tels que la santé (thérapie, diagnostique, pronostique, etc.) ou l'environnement (écologie, agriculture, etc.), notamment grâce aux avancées prometteuses de techniques ultra-sensibles compatibles avec l'analyse des molécules de hauts poids moléculaires, comme la spectrométrie de masse. Une telle analyse permet à la fois de caractériser, d'appréhender la complexité du vivant dans son ensemble et d'évaluer l'état de réponse d'un organisme soumis à un stress. Actuellement, la découverte de nouveaux biomarqueurs pour le dépistage des maladies, ou l'évaluation de l'exposition à une pollution anthropique génère un intérêt accru pour cette science. Cependant ces molécules très complexes sont difficiles à analyser, en raison de leurs faibles quantités.

Ainsi, l'étude des protéines fait encore aujourd'hui face à de nombreux verrous, liés notamment à la complexité de l'échantillon. Un des principaux enjeux scientifiques de la recherche protéomique réside dans la présence d'une grande variété de protéines aux fonctions, natures et concentrations très diverses. Devant une telle complexité, le développement d'un processus analytique adapté, simple et efficace est crucial, afin de purifier au mieux la matrice et ainsi augmenter la sensibilité de détection des protéines d'intérêts. Dans ce contexte, de nombreux protocoles analytiques ont été développés afin de détecter, caractériser et comprendre l'activité et la fonctionnalité biologique de biomarqueurs protéiques issus d'échantillons très divers et complexes. Cependant, dans leur ensemble, les protocoles publiés à l'heure actuelle sont relativement longs et multi-étapes et le plus souvent ciblés pour une

Introduction générale

protéine donnée. On estime ainsi que le taux de récupération des protéines présentes dans les milieux environnementaux est inférieur à 7 %.

Les progrès techniques autour de la spectrométrie de masse ont donc masqué des questions essentielles liées principalement à la qualité de préparation de l'échantillon.

L'objectif principal de ce travail de thèse consiste à développer une approche analytique non ciblée, basée sur une simplification des protocoles expérimentaux existants afin de limiter les pertes en protéines et ainsi amplifier le signal de ces biomolécules au cours d'une analyse par UPLC-HRMS/MS (analyse « hors gel »). En vue d'une optimisation de l'approche expérimentale, chaque étape de l'analyse protéomique « *bottom-up* » a été décortiquée et optimisée, afin de réduire au maximum le nombre d'étapes.

La première partie consiste en une synthèse bibliographique portant sur la structure et les fonctions des protéines et l'intérêt scientifique de l'analyse protéomique. La description des différentes stratégies de préparation de l'échantillon, ainsi que des méthodes d'analyse par LC-HR-MS/MS et par spectroscopies (Spectrofluorométrie, Raman, etc.) sont ensuite présentées.

A l'issue de cette étude bibliographique, les objectifs et la stratégie expérimentale sont exposés. La première partie de ce travail a pour objectif d'optimiser les paramètres analytiques pour la détection des protéines par UPLC-HRMS/MS. La deuxième partie, est consacrée au développement d'une approche de purification simplifiée pour l'analyse des protéines dans les milieux complexes (biologiques et environnementaux). Une étude de la structure moléculaire des protéines a été réalisée à chacune des étapes de purification et nous a permis au final de valider notre approche expérimentale. La troisième partie de cette étude, consiste à étudier les paramètres expérimentaux influençant l'extraction des protéines afin de proposer à terme un protocole d'extraction à la carte compatible avec une analyse directe par UPLC-HRMS/MS.

Introduction bibliographique

Chapitre I : Généralités sur l'étude des protéines

Introduction bibliographique

1. Les protéines : structures et fonctions

Les protéines sont des macromolécules complexes résultant de l'expression des gènes portés par l'acide désoxyribonucléique (ADN) (figure 1). Elles ont en outre un rôle prépondérant dans de nombreux processus biologiques et remplissent des fonctions étroitement liées à leur séquence en acides aminés (aminoacides), leur repliement et à leur stabilité dans un milieu donné (Dobson et al., 1998; Berg et al., 2012).

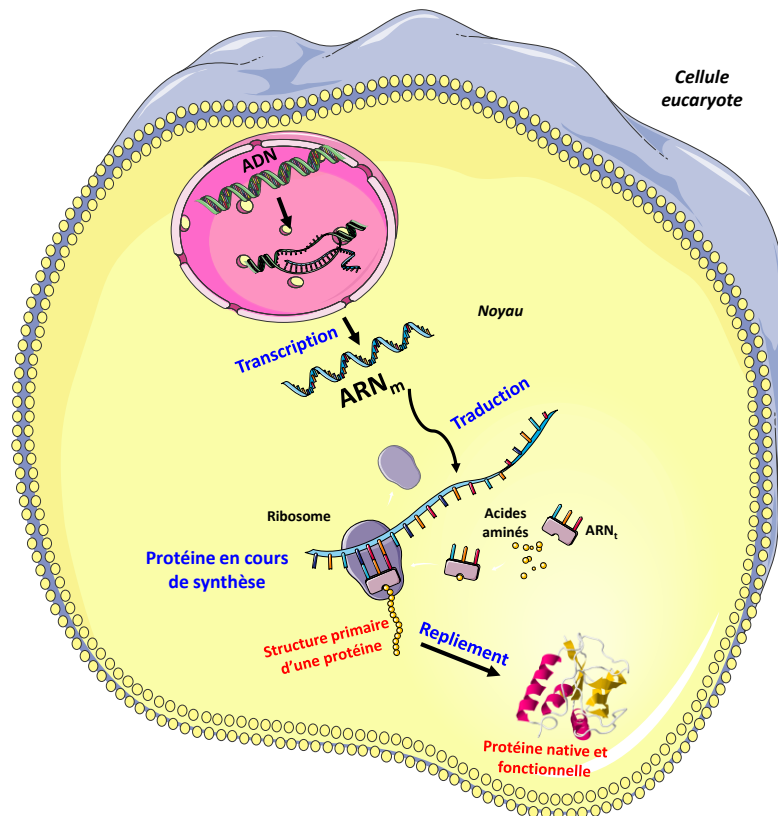


Figure 1. Représentation schématique de la biosynthèse des protéines au sein d'une cellule eucaryote

1.1. Structure des protéines

1.1.1. Les acides aminés

Les acides aminés constituent les briques élémentaires des protéines. Un acide aminé est formé d'un atome de carbone central tétraédrique (α), lié à quatre groupes différents : une fonction amine (NH_2), une fonction acide carboxylique (COOH), un atome d'hydrogène et un groupement variable (R) appelé chaîne latérale (Lodish et al., 2000) (figure 2). La masse des acides aminés est comprise entre 75 et 205 $\text{g}\cdot\text{mol}^{-1}$. Le poids moléculaire moyen des protéines

Introduction bibliographique

est compris entre 5500 et 220 000 g.mol⁻¹ en fonction de la longueur de la chaîne peptidique (Berg et al., 2012).

Les acides aminés sont des molécules chirales. Ils existent donc sous deux formes non superposables, les formes L (lévogyre) et D (dextrogyre). L'énantiomère prépondérant dans les systèmes vivants sur terre correspond à la configuration L (Fuchs et al., 2005). Les acides aminés de la série D ont été en effet détectés à de très faibles concentrations dans certains tissus et sont souvent associés à des maladies liées à l'âge (Parkinson, cataracte) (os, rétine, peau, tissus musculaires cardiaques, tissus nerveux, etc.) (Fujii et al., 2011). Ils ont été également détectés au niveau des parois bactériennes (peptidoglycane) (Friedman, 1999).

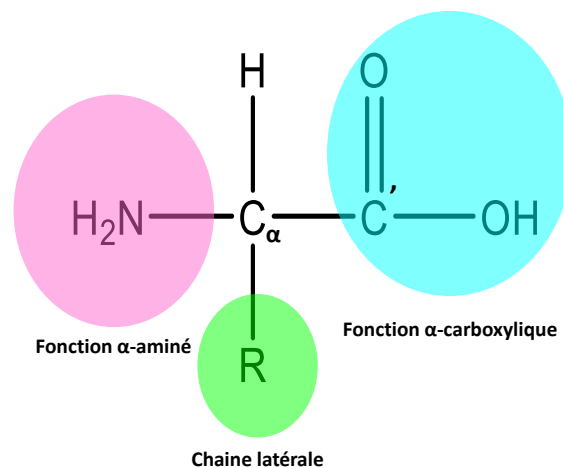


Figure 2. Structure d'un acide aminé. Le carbone central C_{α} est lié à un groupement amine, à un groupement carboxyle, à un atome d'hydrogène et à une chaîne latérale R

Il existe environ 500 acides aminés mais seuls 20 acides aminés sont protéiques et spécifiés dans le code génétique (universel) (Walsh et al., 2013) (différents selon la réactivité chimique de leurs chaînes latérales, l'encombrement stérique, la polarité, la charge et l'aromaticité) (figure 3). À pH dit physiologique (pH physiologique du corps humain : pH 7,4), les aminoacides se trouvent essentiellement sous forme zwitterionique, où la fonction amine est protonée (NH_3^+) et la fonction acide carboxylique est déprotonée (COO^-) (Berg et al., 2012). L'état d'ionisation d'un acide aminé dépend du pH de la solution et de son pKa.

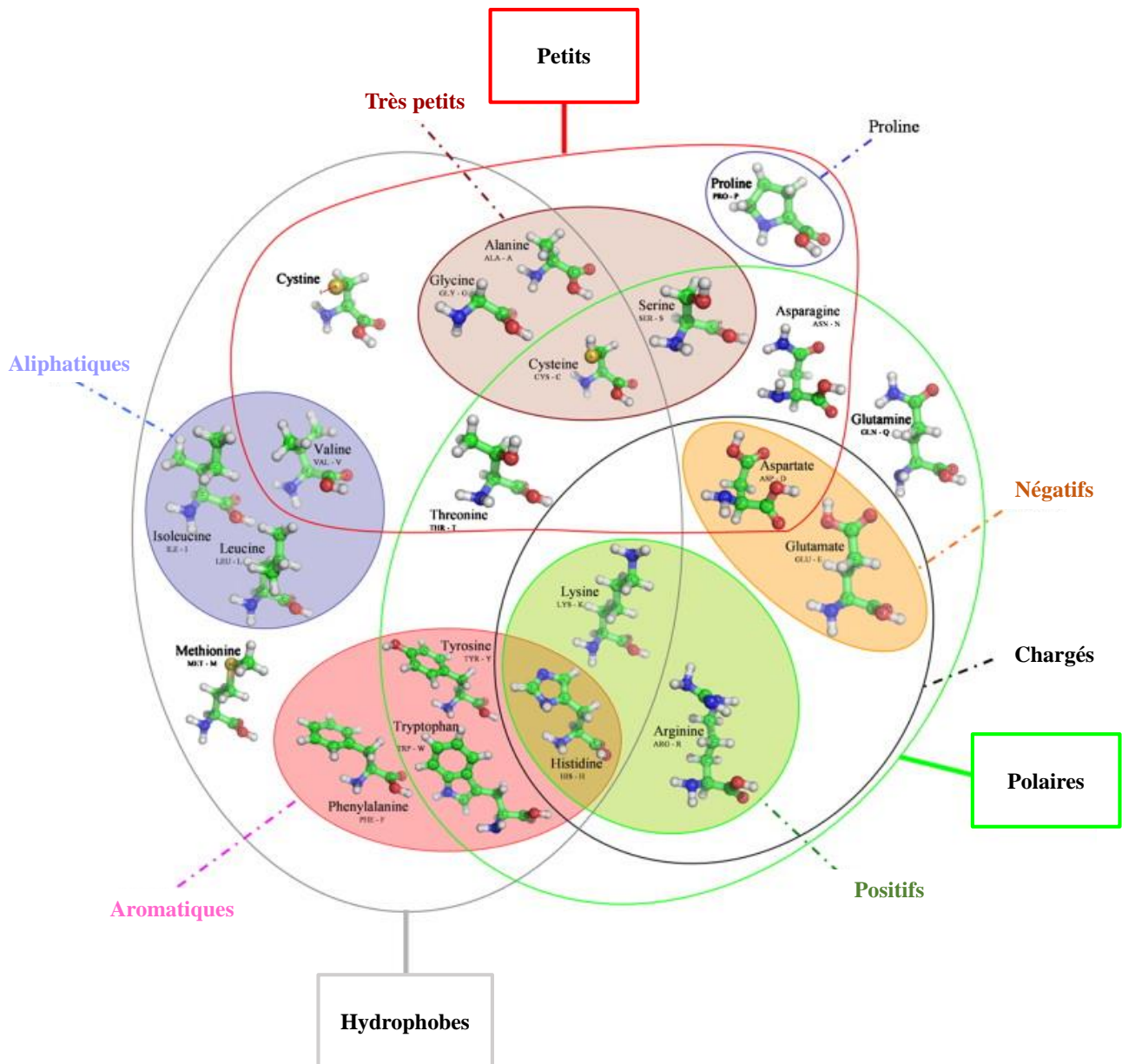


Figure 3. Diagramme de Venn représentant les acides aminés classés en fonction de leurs propriétés physicochimiques (d'après (Faure et al., 2008))

1.1.2. La liaison peptidique

Au sein des protéines, les acides aminés sont reliés entre eux par une liaison amide, appelée aussi liaison peptidique (figure 4). Cette liaison covalente résulte de la réaction entre le groupe α -carboxyle d'un acide aminé avec le groupe α -amine d'un autre acide aminé. La formation d'un dipeptide est accompagnée par la perte d'une molécule d'eau (Berg et al., 2012; Lodish et al., 2000).

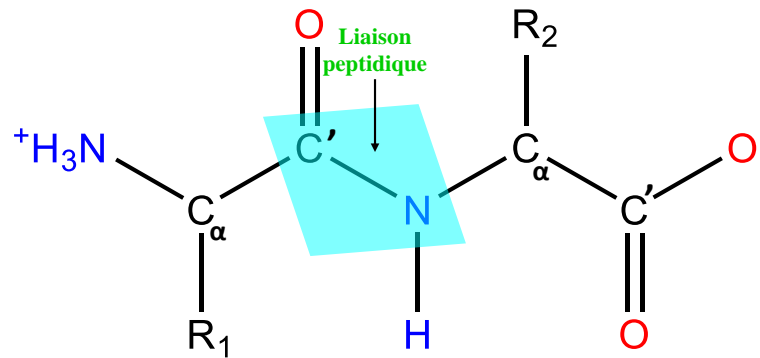


Figure 4. Liaison peptidique entre deux acides aminés (dipeptide)

Les chaînes polypeptidiques peuvent contenir jusqu'à plusieurs milliers de résidus (un résidu fait référence à un aminoacide engagé dans une liaison peptidique au sein de la protéine). Les groupements amines et carboxyles restés libres aux extrémités de la chaîne polypeptidique, sont respectivement appelés, extrémités N-terminale et C-terminale. La désignation "*protéine*" est attribuée aux chaînes polypeptidiques dont le nombre de résidus aminoacides se situe entre 50 et 2000.

1.1.3. Conformation des protéines

Le repliement des protéines est un phénomène complexe. On distingue ainsi quatre niveaux structuraux distincts : une structure primaire qui se réfère à la succession linéaire d'acides aminés ; une structure secondaire qui décrit un deuxième niveau de repliement de la chaîne polypeptidique dans des régions localisées ; une structure tertiaire qui correspond à un repliement plus élaboré de la chaîne polypeptidique ; et enfin une structure quaternaire désignant l'assemblage de plusieurs chaînes polypeptidiques formant chacune une sous-unité protéique (figure 5) (Jez, 2017).

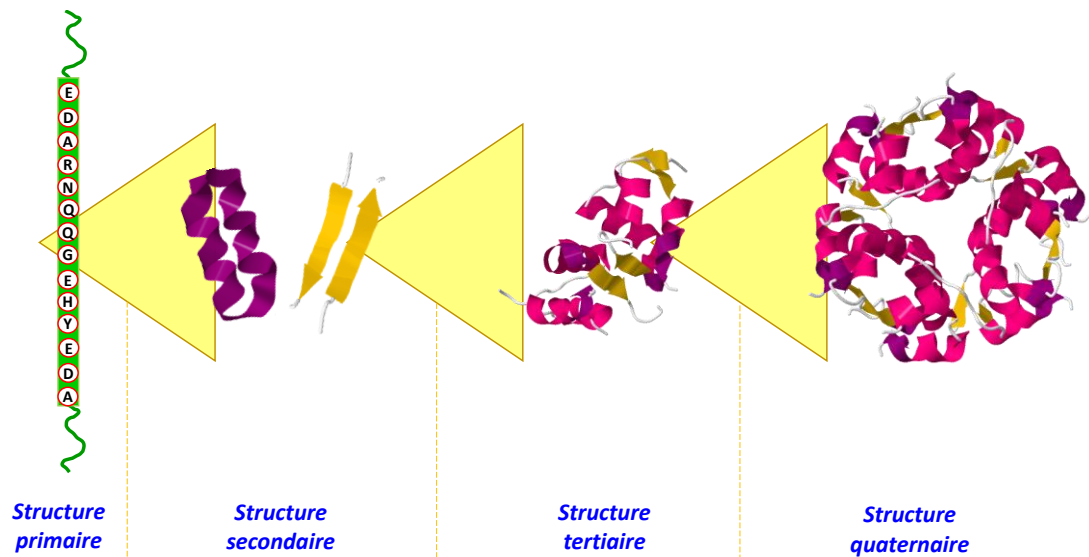


Figure 5. Représentation des quatre niveaux structuraux des protéines

Seules les structures secondaires, tertiaires et quaternaires seront décrites ci-après ; la structure primaire d'une protéine correspondant à la succession linéaire des acides aminés, formée après la traduction de l'ARN messager (figure 1).

1.1.3.1. Structure secondaire

La structure secondaire d'une protéine résulte du repliement de la structure primaire en une structure régulière, qui suit des conformations spécifiques appelées, hélices, feuillets et coudes.

En 1951 Linus Pauling, Robert Corey et Herman Branson (Pauling et al., 1951) ont proposé deux conformations répétitives appelées hélice α et feuillet β . Par la suite, d'autres motifs structuraux non répétitifs ont été identifiés et ont été désignés comme, coude β et boucle Ω . Ces éléments de structures secondaires sont stabilisés par des liaisons hydrogènes établies entre le groupement N-H et le groupement C=O de résidus d'acides aminés proches au sein de la chaîne polypeptidique (Makarov and Plaxco, 2003).

1.1.3.1.1. Hélice α

Les hélices furent découvertes en 1951 par Linus Pauling (Pauling et al., 1951). La première image structurale des motifs secondaires a été obtenue au cours des travaux de Kendrew sur la myoglobine (Kendrew et al., 1958).

Introduction bibliographique

L'hélice α se caractérise sous la forme d'un ruban enroulé en spirale. Elle est stabilisée par des liaisons hydrogènes entre le groupement carbonyle C=O de l'acide aminé n et le groupement amine N-H de l'acide aminé situé quatre résidus plus loin ($n+4$) (figure 6). De cette façon, les chaînes latérales des acides aminés sont orientées vers l'extérieur de la chaîne peptidique.

Comme on peut le remarquer sur la figure 6, chaque unité peptidique est caractérisée par un moment dipolaire généré par la polarité différente des groupements N-H et C=O. Ces moments dipolaires pointent dans la même direction parallèlement à l'axe de l'hélice. Il en résulte un macrodipôle, caractérisé par une charge partielle positive à l'extrémité N-terminale et une charge partielle négative à l'extrémité C-terminale (Hol, 1985). Dans la plupart des cas, le dipôle de l'hélice contribue de façon remarquable à la fixation de petites molécules chargées (exemple : des ligands contenant des groupements phosphates tels que les phospholipides (cardiolipine ou glycérol biphosphatidyle)), peuvent se lier aux extrémités N-terminale de l'hélice α (Brändén and Tooze, 1999; Ruprecht et al., 2014).

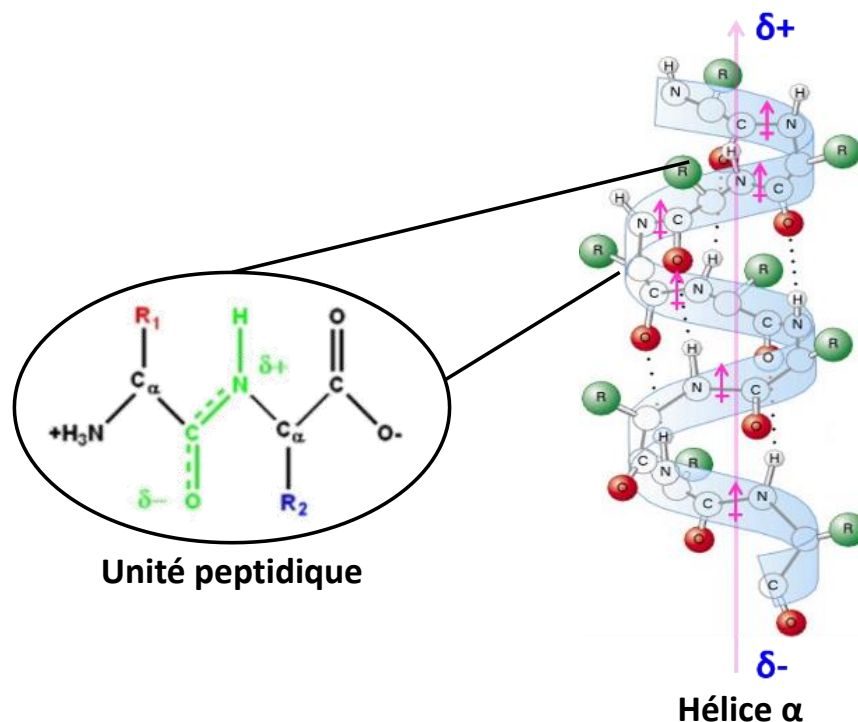


Figure 6. Trajectoire de l'hélice α illustrant l'orientation des dipôles peptidiques ; l'extrémité positive du dipôle est localisée au niveau de l'extrémité N-terminale et l'extrémité négative se trouve à l'extrémité C-terminale

Introduction bibliographique

Des structures secondaires « supérieures » de type α peuvent se former à partir d'un faisceau d'hélices assemblées pour former un cœur hydrophobe. Cette structure est notamment caractéristique des protéines globulaires, telles que la myoglobine. Cette protéine est constituée d'une seule chaîne polypeptidique formée de 8 hélices α qui se replie pour former une cage apolaire fixant un noyau d'hème (groupement prosthétique, caractérisé par une structure aromatique contenant au centre un atome de fer) (Berg et al., 2012; Brändén and Tooze, 1999) (figure 7).

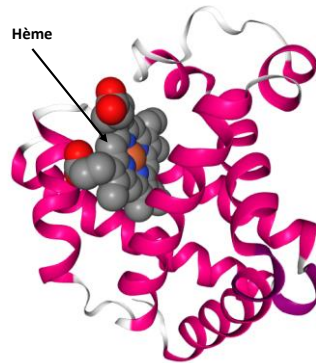


Figure 7. Structure de la myoglobine (la molécule d'oxygène se fixe à la protéine grâce à l'hème) (PDB, 1A6K)

Les proportions en hélices varient considérablement d'une protéine à une autre, de 0 à presque 100 % d'hélices α . À titre d'exemple, la ferritine est une protéine α -hélicoïdale présente chez de nombreux organismes. Environ 75 % de ses résidus d'acides aminés sont engagés dans des structures type hélices α où le fer peut être stocké. Un deuxième exemple, l'albumine humaine sérique (HSA), protéine la plus abondante du plasma sanguin humain, est formée exclusivement d'hélices α stabilisées par des ponts disulfures. Elle est produite par le foie et présente des propriétés de transport d'hormones, de médicaments (diazépam, ibuprofène), d'acides gras etc.(Ghuman et al., 2005) (figure 8).

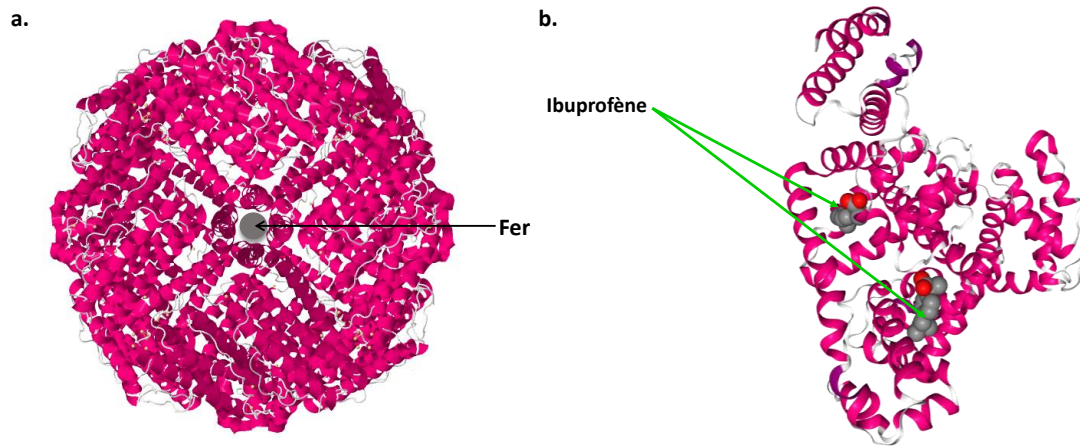


Figure 8. a. Structure de la ferritine isolée de la rate de cheval (Hempstead et al., 1997) (PDB, 1AEW) ; b. Structure de la HSA complexée à l'ibuprofène (PDB, 2BXG)

1.1.3.1.2. Feuillet β

Les feuillets β résultent de la formation de liaisons hydrogènes entre des régions distantes de la chaîne polypeptidique. Deux arrangements de feuillets β sont observés : les feuillets β sont dits parallèles lorsque les brins sont orientés dans le même sens ; les feuillets β antiparallèles lorsque les deux brins sont orientés en sens inverse (figure 9). Les brins de feuillets β parallèles antiparallèles, peuvent se combiner pour former des feuillets mixtes.

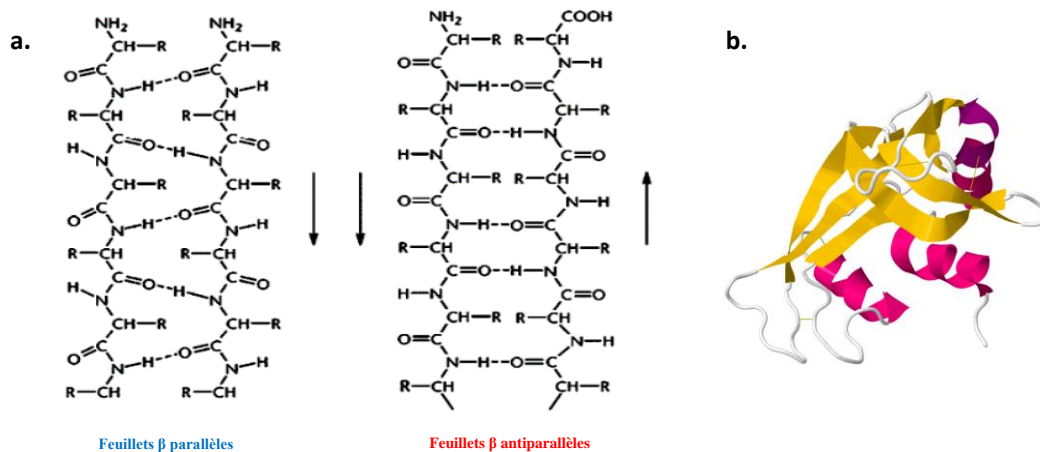


Figure 9. a. Structure des feuillets β (Lowik et al., 2010) ; b. Structure de la ribonucléase bovine, constituée d'une prédominance de feuillets β (PDB, 8RAT)

Parmi les structures connues chez les organismes vivants, seulement 20 % des protéines possèdent à la fois des chaînes β parallèles et des chaînes β antiparallèles. La structure

Introduction bibliographique

secondaire de la thiorédoxine (oxydoréductase) qui possède des feuillets β mixtes est représentée sur la figure 10 (Nikkanen and Rintamäki, 2014).



Figure 10. Représentation de la structure de la thiorédoxine d'*Escherichia coli* (Noguera et al., 2017)

Les feuillets β peuvent également se recourber en un cylindre fermé pour adopter une structure en « tonneau β ». Cette caractéristique est couramment rencontrée dans les protéines de transport responsables de la diffusion des ions à travers les membranes cellulaires (figure 11), (exemple : protéine membranaire FhuA présente chez les bactéries gram négatives du sol (*E.coli*, *Salmonella enterica* et *Pseudomonas aeruginosa*) où elle possède également une fonction de récepteur aux bactériophages (Chaturongakul and Ounjai, 2014)).

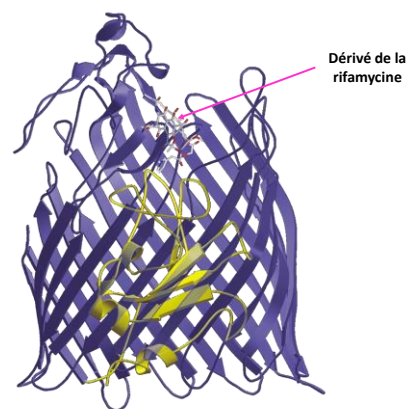


Figure 11. Structure de la protéine de transport FhuA d'*Escherichia coli* composée entièrement de feuillets β complexée à un dérivé de la rifamycine (antibiotique) (Ferguson et al., 2001)

1.1.3.1.3. Les coudes et les boucles

Introduction bibliographique

L'ensemble des éléments de structures secondaires, hélices α et feuillets β , sont reliés entre eux par des motifs structuraux appelés boucles et coudes (figure 12). Ces structures non répétitives sont caractérisées par des longueurs différentes et des formes irrégulières, impliquant un changement dans la direction du squelette polypeptidique (Brändén and Tooze, 1999).

Les boucles sont caractérisées par des segments longs (plus de quatre acides aminés) qui jouent le rôle de jonctions entre deux structures secondaires (hélices α et/ou feuillets β). Elles sont généralement exposées vers l'extérieur du milieu après repliement de la protéine. Au niveau de ces régions, les groupements chimiques C=O et N-H sont dirigés vers le milieu, ce qui leur permet d'établir des liaisons hydrogènes avec les molécules d'eau environnantes (Berg et al., 2012; Brändén and Tooze, 1999; Lodish et al., 2000).

Les coudes sont caractérisés par des segments plus courts comportant généralement deux à quatre acides aminés. Ils permettent de lier deux brins β antiparallèles adjacents. La glycine et la proline sont fréquemment trouvées dans les structures de type coude (Lodish et al., 2000). Une liaison hydrogène établie entre le groupe C=O d'un résidu n et le groupe N-H d'un résidu $n+3$ d'une chaîne polypeptidique, permet de stabiliser cette structure irrégulière.

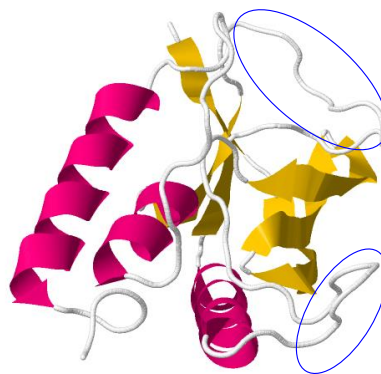


Figure 12. Exemple de représentation de boucles (cercles bleus) au sein de structures secondaires répétitives de la nucléase de *Staphylococcus aureus* (PDB 1EQV)

De tels motifs irréguliers sont fondamentaux, puisqu'ils fonctionnent comme des unités de jonctions des éléments de structures secondaires et participent de la sorte au maintien de la conformation protéique et à la reconnaissance des composés spécifiques. C'est par exemple le cas des immunoglobulines (Ig) (anticorps), toutes formées de deux chaînes lourdes et de deux chaînes légères reliées entre elles par des ponts disulfures et à l'extrémité desquelles se trouvent des boucles (domaine variable. Celles-ci constituent le site de reconnaissance de l'antigène

Introduction bibliographique

(région appelée CDR ou *Complementarity determining regions*). La séquence en acides aminés ou le repliement de ces boucles est différent d'une Ig à l'autre ce qui lui permet de reconnaître un antigène spécifique (figure 13). Domaine variable V 3 variables hypervariables (CDR).

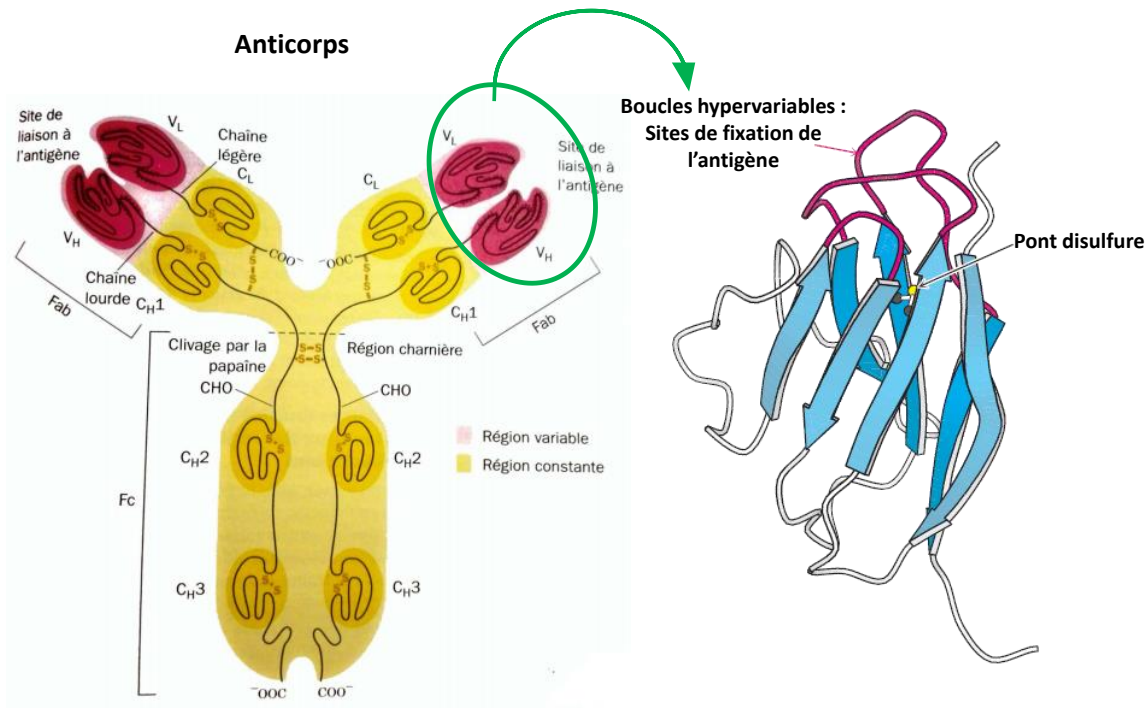


Figure 13. À gauche : représentation schématique d'un anticorps IgG ; à droite : Illustration du pli immunoglobuline avec les boucles à la surface (en rouge), correspondant aux sites de fixation de l'antigène. L'illustration de l'anticorps a été réalisée par Irving Geis/Geis

Archives Trust (d'après Voet and Voet, 2016)

1.1.3.2. Structure tertiaire

La fonction biologique des protéines est liée intrinsèquement à leur structure tertiaire (Lodish et al., 2000; Ponnuswamy et al., 1982). Elle correspond à la conformation dite native de la protéine et constitue à elle seule une entité unique et complexe (Dobson et al., 1998). La structure tertiaire résulte du repliement des motifs secondaires (hélices α , feuillets β et coudes) grâce à des liaisons de faible énergie (les liaisons hydrogènes, électrostatiques, Van Der Waals, interactions hydrophobes), qui ont un rôle décisif dans le maintien d'une structure stable (Baker and Hubbard, 1984; Dill, 1990; Pace, 2009; Ponnuswamy and Gromiha, 1994; Nick Pace et al., 2014). Des liaisons covalentes participent également au maintien de cette structure tertiaire. Il s'agit de ponts disulfures qui se forment après oxydation du groupement thiol (SH) de deux cystéines (Lodish et al., 2000; Voet and Voet, 2016).

1.1.3.3. Structure quaternaire

La structure quaternaire décrit l'arrangement spatial de structures tertiaires *via* des liaisons diverses (hydrogènes, hydrophobes, etc.). Ainsi, plusieurs chaînes polypeptidiques peuvent s'assembler pour former des oligomères contenant de deux jusqu'à dix ou même d'avantage de monomères. On parle alors de dimères, trimères, tétramères pentamères, etc (figure 14). Les structures quaternaires sont très courantes, comme les fibrilles de collagène. Cette association est également retrouvée dans les enzymes multimériques, ce qui leur permet d'augmenter leur efficacité catalytique, puisque chaque sous-unité protéique sera dotée d'un site actif (Voet and Voet, 2016).

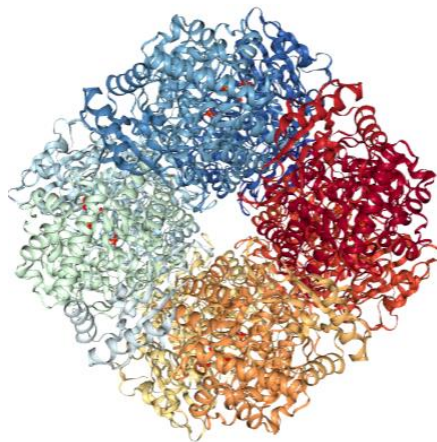


Figure 14. Exemple de structure quaternaire de la ribulose-1,5-bisphosphate carboxylase/oxygénase (*Spinacia oleracea* (plante)) (PDB, 1RCX)

L'analyse de la structure d'une protéine permet de mieux appréhender les différents paramètres influençant la stabilité de sa conformation native (conditions extrêmes de pH et de températures, etc.). Inversement, la connaissance de ces conformations permet de mieux cibler les techniques d'analyse (techniques spectroscopiques).

1.1.4. Stabilité conformationnelle des protéines

1.1.4.1. Stabilité de la structure native de la protéine

Tous les processus biologiques sont rendus possibles grâce à la stabilité structurale, à la réactivité chimique, mais également à la diversité des protéines. En 2001, Lesk décrit la structure primaire d'une protéine comme des lettres, la structure secondaire comme des mots, la structure tertiaire comme une phrase et enfin sa structure globale comme un texte (M. Lesk,

Introduction bibliographique

2000). Autrement dit, une déstabilisation de la structure tertiaire, secondaire ou primaire d'une protéine entraînera de ce fait une modification de sa fonction biologique initiale (Elkordy et al., 2008).

Un exemple significatif correspond aux maladies neurologiques, telles que la maladie d'Alzheimer, la maladie du prion (encéphalopathie spongiforme transmissible ou maladie de Creutzfeldt–Jakob) ou la maladie de Parkinson. En effet, ces pathologies sont liées à un dépôt d'agrégats de protéines, qui bien qu'elles soient initialement solubles, convertissent une partie de leurs hélices α en feuillets β entraînant alors leur agrégation sous forme de protéines amyloïdes insolubles, stabilisées par des liaisons hydrogènes (Apostol et al., 2013; Brundin et al., 2010; Cobb et al., 2007; Eisenberg and Jucker, 2012; Goedert, 2015; Gross, 2000; Kourouski et al., 2015). Il en résulte des plaques amyloïdes (ou plaques séniles) qui induisent une toxicité pour les cellules nerveuses (figure 15).

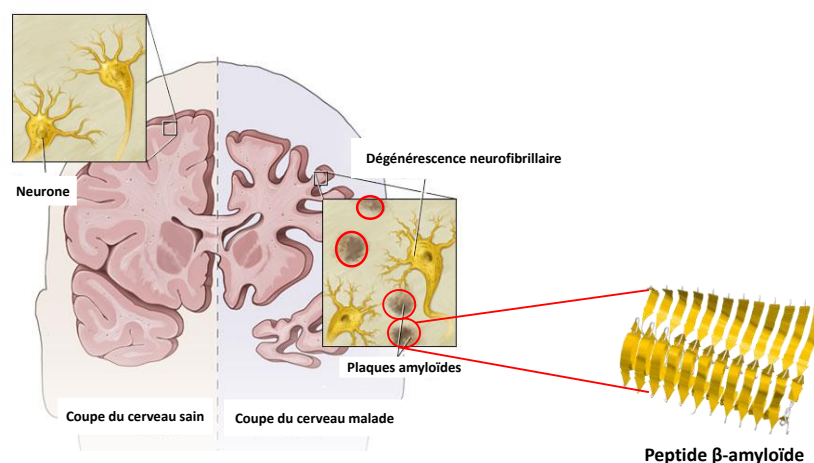


Figure 15. Illustration du dépôt des plaques amyloïdes dans le cerveau humain, mettant en évidence la formation de feuillets β (la structure secondaire du β -amyloïde est tirée de : (Xiao et al., 2015))

Ces structures particulières de type amyloïde ont également été observées dans de nombreuses protéines modèles lorsque celles-ci étaient soumises à un changement de température ou de pH. L' α -lactalbumine, par exemple, forme des fibres amyloïdes à pH acide (pH 2) et adopte une conformation intermédiaire appelée *molten globule*, riche en feuillets β (Dolgikh et al., 1981; Goers et al., 2002; Kourouski et al., 2015; Semisotnov et al., 1991). De façon générale, l'état *molten globule* est caractérisé par une structure compacte réversible, partiellement dénaturée, qui préserve les mêmes éléments de structures secondaires que l'état

Introduction bibliographique

natif mais dans des proportions variables. En outre, elle perd l'intégrité de sa structure tertiaire (Dolgikh et al., 1981; Ptitsyn, 1995). Cet état peut être différent pour une même protéine en fonction des conditions environnantes, mais également différent d'une protéine à une autre, pour des conditions identiques.

1.1.4.2. Influence du milieu sur l'arrangement des protéines

Les liaisons hydrogènes contribuent significativement à la stabilisation des protéines (Rose and Wolfenden, 1993; Myers and Pace, 1996; Pace, 2009). Ces liaisons s'établissent essentiellement avec les molécules d'eau, entre les chaînes latérales polaires (ou chargées) ou encore au niveau du squelette peptidique (formation des hélices α et/ou feuillets β).

En milieu aqueux, les molécules d'eau forment des liaisons hydrogènes avec les chaînes latérales polaires des protéines constituant ainsi une « cage » de solvation qui entoure les protéines (Dill, 1990; Kauzmann, 1959; Rose and Wolfenden, 1993).

L'ensemble des groupements hydrophobes de la protéine est alors enfoui au cœur de la protéine lui conférant ainsi une organisation plus compacte, maintenue par des interactions hydrophobes internes (Lesser and Rose, 1990). De ce fait, les interactions entre les molécules d'eau et les résidus hydrophobes sont moins favorables.

En présence d'un solvant organique, les chaînes apolaires des acides aminés vont être dirigées vers l'extérieur de la protéine alors que les chaînes polaires ont tendance à réagir entre elles au cœur de la molécule (Dwyer and Bradley, 2000; Khabiri et al., 2013; Shao et al., 2012) (figure 16).

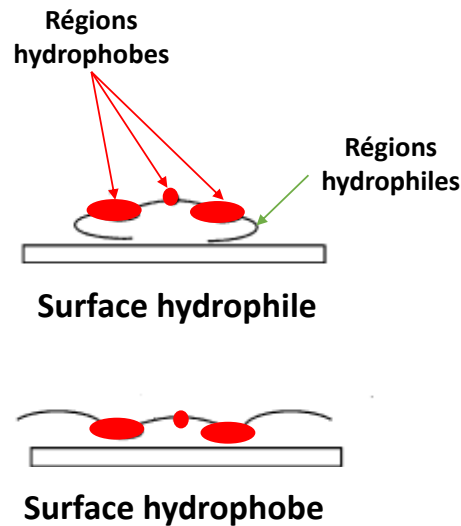


Figure 16. Représentation schématique des interactions établies entre une protéine (régions hydrophiles/hydrophobes) et son environnement (hydrophile/hydrophobe) (d'après Marchin and Berrie, 2003)

1.1.4.3. Dénaturation des protéines

Les protéines sont stables à pH neutre et à température ambiante. À l'inverse, elles peuvent être dénaturées par n'importe quel traitement susceptible de provoquer une rupture des liaisons de faible énergie stabilisant leurs structures secondaires et tertiaires. Parmi ceux-ci, on peut citer une augmentation ou une diminution de la température ou du pH, la présence de fortes concentrations en solvant organique ou de détergents (exemple : le dodécylsulfate de sodium ou SDS) (Voet and Voet, 2016). Ces derniers font partie des agents chimiques les plus utilisés pour dénaturer les protéines. Ils s'associent *via* des interactions hydrophobes avec les résidus d'acides aminés apolaires et déstabilisent par conséquent les interactions hydrophobes impliquées dans la stabilité de la structure native de la protéine. Celle-ci aura donc tendance à exposer ses motifs apolaires vers l'extérieur et sera ainsi « dépliée ». Le SDS peut jouer ce rôle de dénaturant. C'est en effet un tensioactif anionique constitué d'une chaîne hydrocarbonée hydrophobe et une extrémité sulfate chargée négativement lui conférant des propriétés amphiphiles (figure 17).

Introduction bibliographique

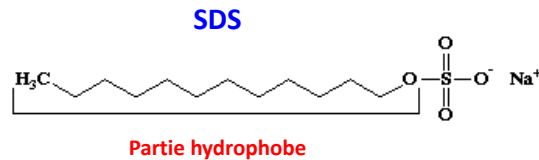


Figure 17 Représentation de la structure du dodécylsulfate de sodium (SDS)

Le SDS se lie aux protéines avec un ratio de 1,4 g SDS/g protéine (w/w) (Reynolds and Tanford, 1970). Il établit des interactions hydrophobes avec les chaînes latérales des résidus d'acides aminés apolaires *via* sa partie hydrocarbonée, entraînant ainsi une dénaturation de la structure tertiaire de la protéine (Tejaswi Naidu and Prakash Prabhu, 2011).

D'autres agents dénaturants ou chaotropes tels que l'urée déstabilisent la structure protéique en formant des liaisons hydrogènes avec les protéines (groupements NH ou CO) et ont donc tendance à remplacer les molécules d'eau, ce qui favorise leur solubilité dans ce type de milieu (Fennema, 1996; Lim et al., 2009).

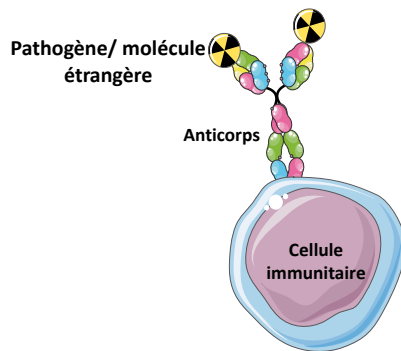
Cette brève présentation montre la complexité des protéines, leurs fortes interactions chimiques avec le milieu et pointe donc la difficulté d'analyse inhérente à cette famille. Cependant, les dernières avancées technologiques permettent aujourd'hui d'aller plus avant dans l'étude de ces macromolécules. Leurs fonctions fondamentales dans les organismes vivants, et leur forte dépendance au milieu qui les contient, en font d'eux des marqueurs dans des domaines aussi divers que la médecine et l'environnement.

Dans la partie suivante seront discutés quelques exemples de ces fonctions principales.

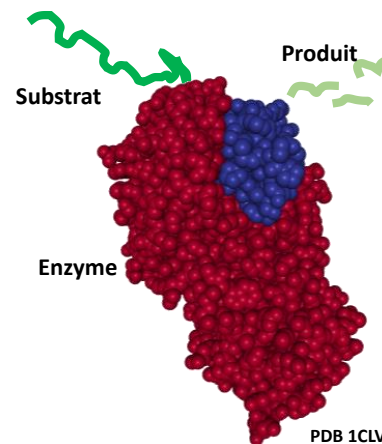
1.2.Fonctions des protéines

Les protéines jouent des rôles cruciaux puisqu'elles assurent des fonctions fondamentales à la survie et la multiplication cellulaire. Parmi ces fonctions, on distingue la fixation et la reconnaissance spécifique d'une molécule (liaison anticorps-antigène), le transport (diffusion des molécules au travers des membranes biologiques, etc), la catalyse de réactions biochimiques, le maintien de la structure et de la solidité cellulaire (Uversky and Dunker, 2010). La figure 18 présentée ci-dessous, récapitule certaines de ces fonctions.

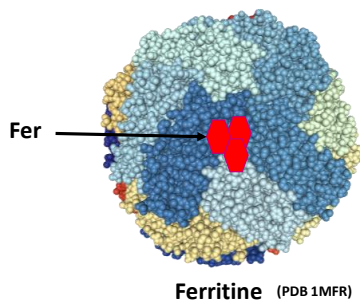
Défense



Catalyse enzymatique



Stockage



Structure et mouvements cellulaire

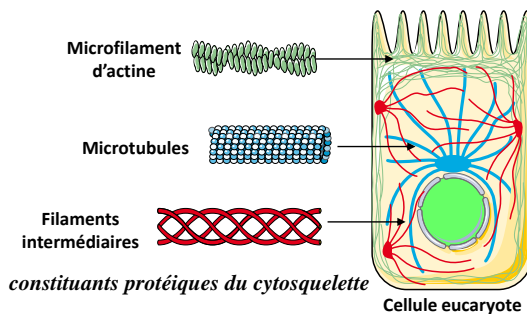


Figure 18. Représentation schématique de quelques fonctions des protéines

Certaines protéines peuvent exercer une seule et unique fonction, tandis que d'autres vont avoir un rôle dans divers processus biologiques. Parmi ces dernières, nous pouvons citer la tubuline qui intervient dans différents processus cellulaires. Il s'agit d'une protéine hétérodimérique, formée de deux sous-unités polypeptidiques : α -tubuline et β -tubuline, qui s'associent pour former des protofilaments (figure 19). La polymérisation de ces protofilaments aboutit à la formation de fibres tubulaires appelées aussi microtubules, constituants majeurs du cytosquelette des cellules (figure 18). Sous cette forme, les microtubules interviennent dans les mouvements intracellulaires, dans les déplacements des cellules sur un support ou encore dans la contraction des fibres musculaires. Les microtubules sont présents aussi bien dans les cellules animales que végétales.

Introduction bibliographique

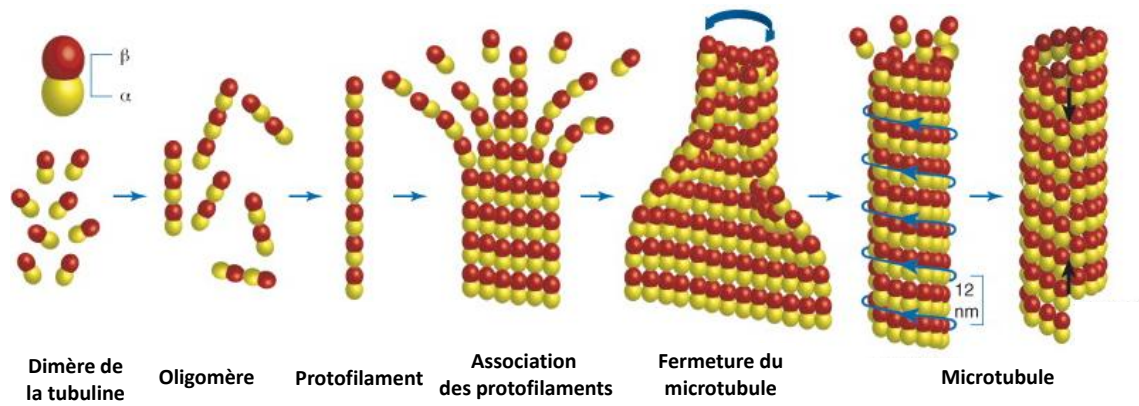


Figure 19. Représentation schématique des étapes de polymérisation des microtubules (d'après Pampaloni and Florin, 2008)

La tubuline accomplit une fonction essentielle à la division cellulaire (mitose), où elle forme le fuseau mitotique qui aboutit à la séparation des chromosomes dans les deux cellules filles (figure 20).

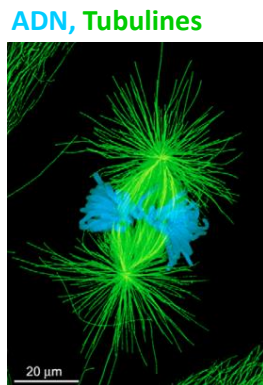


Figure 20. Micrographie du fuseau mitotique d'une cellule animale après marquage par des fluorescents spécifiques. Les tubulines sont colorées en vert et l'ADN en bleu (d'après O'Connell and Khodjakov, 2007)

Étant donné son rôle majeur lors de la division cellulaire, la tubuline fait l'objet d'une attention particulière dans le cadre du développement de nombreux agents anticancéreux. En effet, elle possède des sites de fixation spécifiques à certains agents, tels que le Taxol®, qui lorsqu'il se lie à la β -tubuline, entraîne un changement de conformation au niveau d'une région spécifique de cette sous-unité. Telle qu'énoncée précédemment, la modification de la structure de cette protéine entraîne une modification de sa fonction initiale qui se traduit en l'occurrence par une inhibition de la dissociation des microtubules et par conséquent un arrêt de la

Introduction bibliographique

prolifération des cellules cancéreuses (Alushin et al., 2014; Arnal and Wade, 1995; Mukhtar et al., 2014).

1.2.1. Fonction de défense

La fonction de défense repose principalement sur leurs propriétés structurales qui leur permettent d'acquérir une complémentarité de forme avec un ligand spécifique (ici l'antigène) qui est en outre maintenue par des liaisons faibles, telles que des liaisons hydrogènes (Berg et al., 2012).

C'est cette complémentarité de forme, qui permet notamment aux anticorps de reconnaître de manière spécifique des unités de fixation de l'antigène, nommées déterminant antigénique ou encore épitope, et d'assurer ensuite leur fonction de défense au sein du système immunitaire (figure 21).

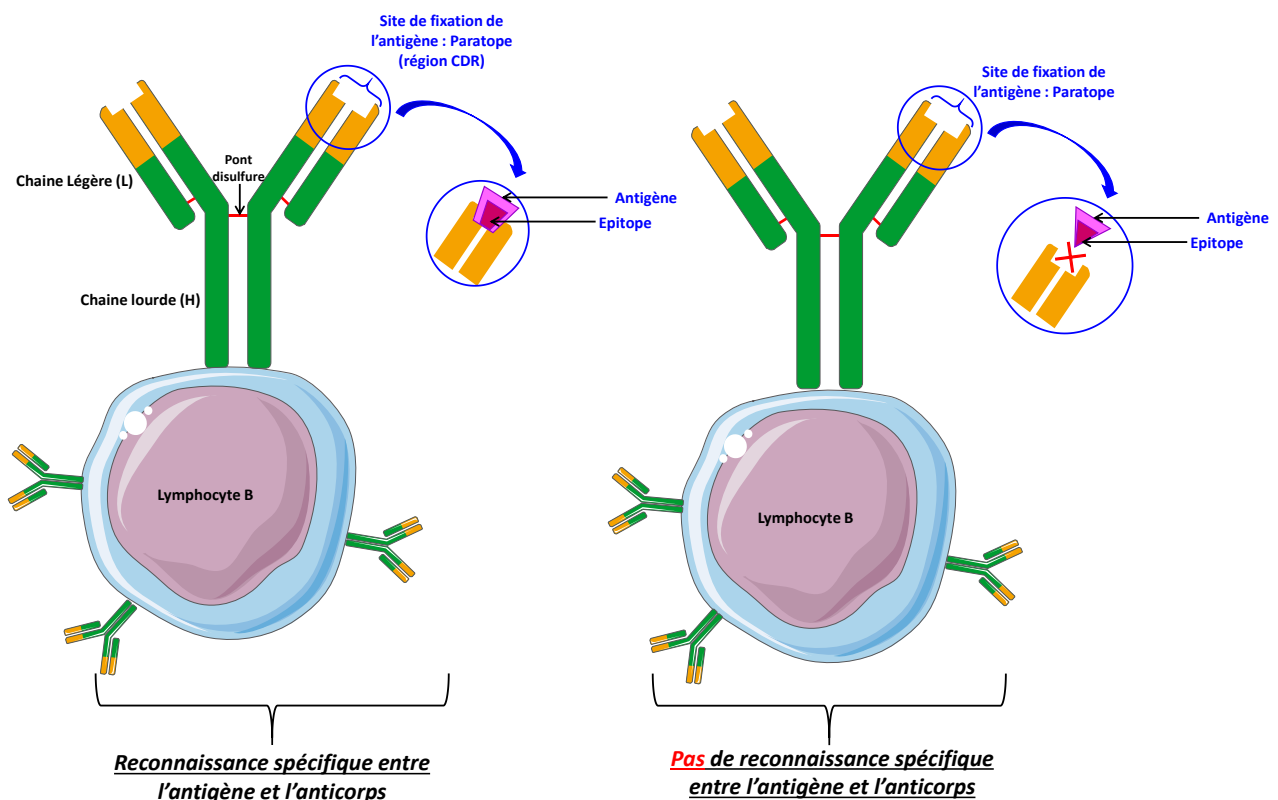


Figure 21. Représentation schématisée d'une interaction anticorps-antigène (Ac-Ag). Les lymphocytes B sont des cellules participant à la réponse immunitaire, elles produisent des anticorps qui vont servir à la reconnaissance spécifique (Ac-Ag)

Introduction bibliographique

1.2.2. Fonction catalytique

Cette fonction est assurée par des protéines appelées enzymes. Ces catalyseurs biologiques sont de véritables machines moléculaires qui accélèrent de manière remarquable la vitesse de la réaction chimique. À titre d'exemple, dans les globules rouges, la transformation du CO₂ en acide carbonique en présence d'eau par l'anhydrase carbonique, est effectuée à une vitesse 10⁷ fois supérieure à la vitesse de la réaction spontanée (Radzicka and Wolfenden, 1995).

Les enzymes accomplissent plusieurs rôles dans divers domaines de recherche, y compris le domaine de l'environnement. Les enzymes d'origine microbienne (cellulase, lipase, uréase, phosphatase, phénol oxydase) qui sont présentes dans les milieux naturels (sol, milieu aquatique, sédiment, etc.) sont impliquées dans les cycles biogéochimiques (carbone, azote, etc.) mais également dans la dégradation de nombreux polluants organiques (Bastida et al., 2015).

La laccase (phénol oxydase) est une enzyme cuivre-dépendant qui catalyse l'oxydation d'un large spectre de composés phénoliques en présence d'oxygène. Les laccases d'origine fongiques (*Trametes versicolor*) participent à l'élimination des micropolluants présents dans les eaux usées et dans les sols contaminés par les hydrocarbures aromatiques polycycliques (HAP) (Arca-Ramos et al., 2016; Y. Wu et al., 2008).

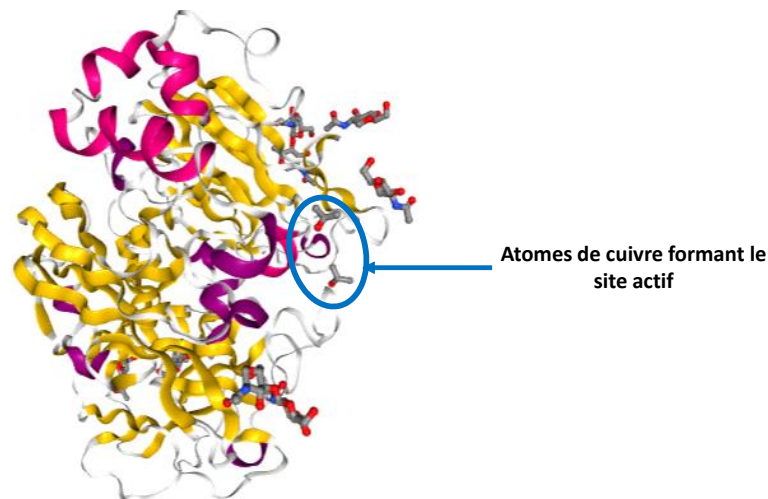


Figure 22. Structure de la laccase de *Trametes versicolor*. Le site actif est formé d'atomes de cuivre (PDB, 1GYC)

La capacité de l'enzyme à former un complexe avec son substrat repose sur la structure du site actif. C'est l'affinité de l'enzyme pour son substrat qui va permettre le déroulement d'une

Introduction bibliographique

réaction enzymatique optimale. De la même façon, une modification de la structure (secondaire ou tertiaire) du site actif modifie son activité. Ainsi, l'activité enzymatique de la ribonucléase bovine (Rnase) diminue après ajout d'un agent réducteur des ponts disulfures. Une fois, l'agent réducteur éliminé, les groupes sulfhydryles de l'enzyme dénaturée sont réoxydés par l'air et l'enzyme retrouve son activité catalytique après repliement spontané (figure 23) (R. T. Raines, 1998)). Cette protéine sera également prise pour modèle dans notre étude.

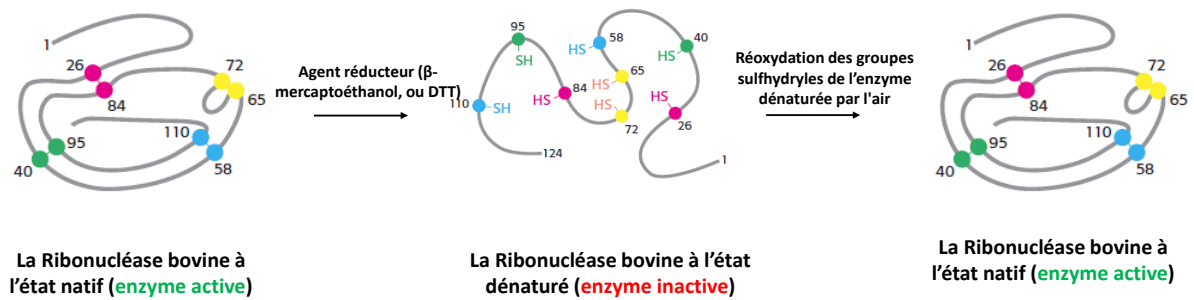
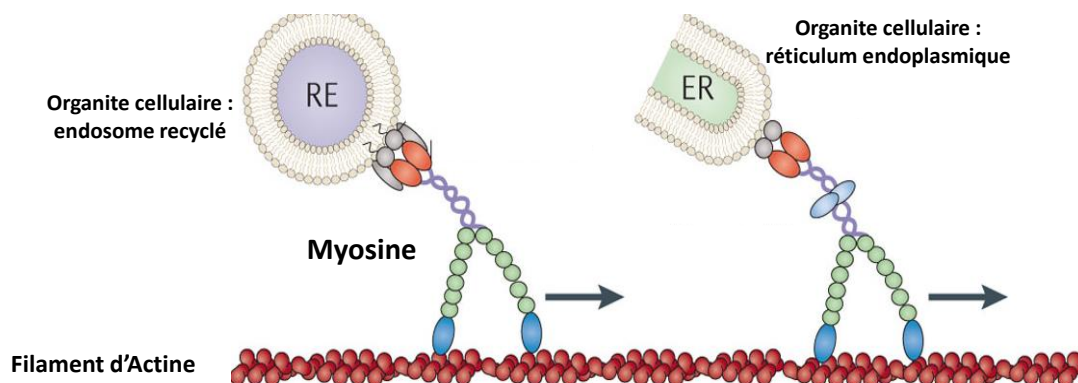


Figure 23. Effet des substances chimiques sur l'activité enzymatique de la ribonucléase bovine (d'après Berg et al., 2012)

1.2.3. Fonction structurale

Au niveau microscopique, une cellule est formée de protéines structurales particulières qui définissent la forme, la résistance et la flexibilité de ses éléments qui la composent. Parmi ces protéines, les filaments d'actine (protéine d'assemblage qui aboutit aux microfilaments) constituent des structures dynamiques, qui forment un réseau sur lequel peuvent se déplacer des moteurs moléculaires appelés myosines en utilisant l'énergie libérée par l'hydrolyse d'une molécule d'adénosine triphosphate (ATP). L'interaction entre ces deux protéines est directement associée à la contraction musculaire (figure 24) (dos Remedios et al., 2003).



Introduction bibliographique

Figure 24. Représentation schématique du déplacement de la myosine le long des filaments d'actine (d'après Kneussel and Wagner, 2013)

Les extensines sont un autre exemple de protéines structurales, largement répandues dans le règne végétal. Ces glycoprotéines riches en hydroxyprolines (HRGPs) sont localisées au niveau de la paroi des végétaux supérieurs et jouent un rôle dans la défense des plantes contre les pathogènes (Hall and Cannon, 2002).

Suivant l'ensemble des exemples présentés ci-dessus, il apparaît que les protéines exercent un nombre considérable de fonctions qui sont intimement liées à leur structure native déterminée à leur tour par la séquence primaire en aminoacides, leurs structures secondaires et tertiaires et leur stabilité dans un environnement donné (Wright and Dyson, 1999). Toute modification de la structure d'une protéine entraînera une altération de sa fonction, ce qui est susceptible de faire dysfonctionner toute la chaîne métabolique dans laquelle elle est impliquée.

2. Stratégie générale en protéomique

2.1.La protéomique

La protéomique désigne la science qui étudie l'ensemble des protéines, exprimées par une cellule ou un tissu, ou présentes dans un milieu à un moment donné et dans des conditions données (Anderson and Anderson, 1998; Blackstock and Weir, 1999; Wilkins et al., 1996). Le terme « protéomique » fut introduit pour la première fois en 1997 par analogie à la « génomique » qui étudie le génome d'un organisme et la « transcriptomique », qui s'intéresse à l'ensemble des transcrits (ARNs messagers) (James, 1997) (figure 25).

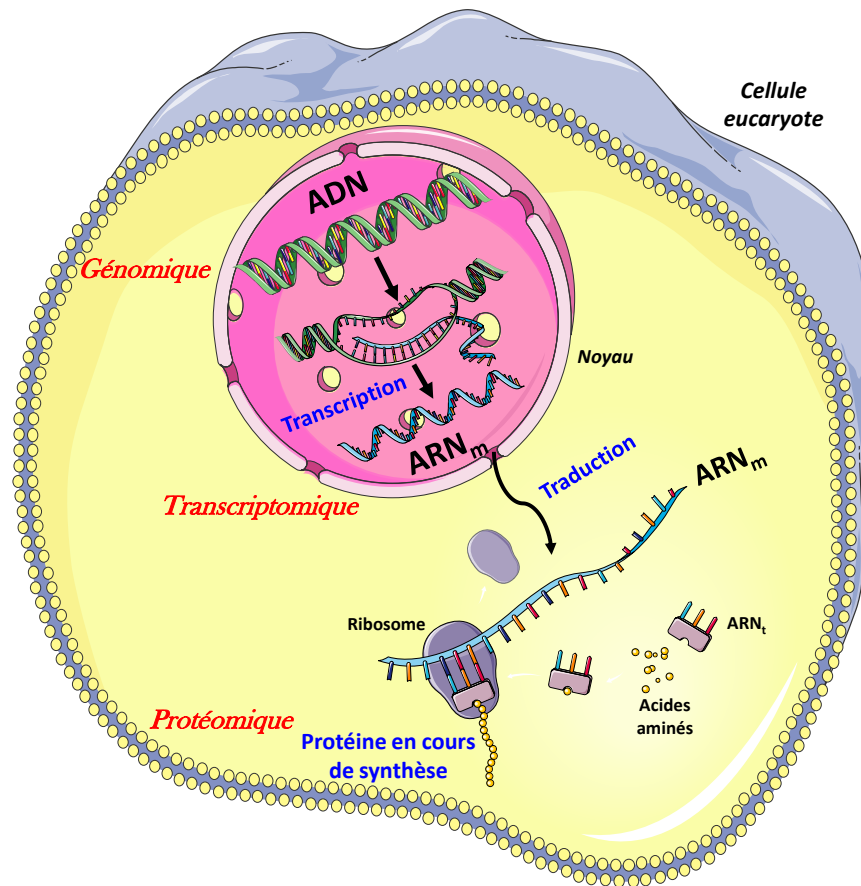


Figure 25. De la génomique vers la protéomique (Les différents organites de la cellule eucaryote ne sont pas illustrés)

L'ensemble des processus biologiques d'un organisme donné sont codés par son génome (c'est-à-dire l'ensemble des gènes). À titre d'exemple, le génome de *Vibrio Chlorae* (bactérie responsable de la maladie du choléra) comporte une multitude de gènes codant des protéines aux fonctions diverses. Toutefois, le simple catalogue des gènes d'un organisme donné ne rend pas compte du fonctionnement de ces gènes. En effet, certains gènes sont fortement exprimés tandis que d'autres le sont rarement (Voet and Voet, 2016). La transcriptomique fournit quant à elle une vision fidèle de l'expression des gènes. Ainsi, l'identification des transcrits d'un type cellulaire donne une indication sur l'implication des gènes dans les processus biologiques. Alors qu'elle est l'intérêt de la protéomique ?

Il existe une faible corrélation entre les niveaux d'abondance (ou la quantité) d'un ARN messager donné et de la protéine fonctionnelle qu'il code (la diversité du protéome est bien supérieure à la diversité des ARNs messagers). De même, les protéines codées par un même gène peuvent avoir une fonction différente suivant les modifications post-traductionnelles dont elles font l'objet. Ces modifications ne sont pas visibles à l'échelle de l'ARN messager

Introduction bibliographique

(Bannister and Kouzarides, 2011; Smith and Workman, 2009). En outre, les processus dynamiques de maturation et de dégradation subis par une protéine interviennent dans son niveau d'abondance à un temps donné et dans un compartiment cellulaire donné (cytoplasme, noyau, mitochondrie, etc.).

La protéomique a marqué indéniablement un tournant dans de nombreux domaines de recherche tels que la santé (thérapie, diagnostique, pronostique, etc.) ou l'environnement (écologie, agriculture, etc.), grâce aux progrès méthodologiques, instrumentaux et techniques.

L'analyse du protéome d'une cellule représente un véritable challenge pour de nombreux scientifiques. En effet, si l'on considère qu'entre 20 000 et 40 000 gènes s'expriment dans les cellules des organismes supérieurs, que chaque gène code au minimum pour trois protéines, et que chacune peut subir différentes modifications post-traductionnelles, le nombre de protéines différentes exprimées chez cet organisme pourrait varier entre 50 000 jusqu'à un demi-million (Hochstrasser et al., 2002). En 2014, l'équipe de Kim et al. (2014), a identifié environ 30 000 protéines constitutives du protéome humain et fœtal. Cette étude fut la première à établir une carte exhaustive des protéines à partir de différents tissus (poumons, cœur, foie, prostate, ovaires, etc.) (Kim et al., 2014).

2.2. Recherche des biomarqueurs protéiques

L'une des approches suivies en analyse protéomique correspond à la protéomique différentielle. Son objectif est de déterminer les protéines dont les taux d'expression sont significativement différents entre deux échantillons (Monteoliva and Albar, 2004). Cette différence de niveau d'expression protéique peut être qualitative (présence ou absence d'un ensemble de protéines) ou quantitative (variation de la quantité de protéines). Cette approche est notamment suivie afin d'analyser les variations d'expression de protéines par un organisme soumis à un stress (élévation de la température ; (Shi et al., 2017), ou une maladie ; (Xiao et al., 2016) ou après un traitement (Prieto et al., 2008).

Cette approche est aujourd'hui utilisée pour mettre en évidence des protéines dites biomarqueurs, dont l'expression est fortement modifiée en présence d'un stimulus donné.

2.2.1. Intérêts scientifiques et défis associés

La définition de biomarqueur varie suivant le domaine de recherche scientifique. Dans le domaine de la santé, un biomarqueur est défini comme étant « une caractéristique qui est

Introduction bibliographique

mesurée et évaluée objectivement comme un indicateur d'un processus biologique normal, pathologique ou d'une réponse pharmacologique à une intervention thérapeutique » (Biomarkers Definitions Working Group, 2001).

Dans le domaine de l'environnement, un biomarqueur permet de détecter la présence de polluants ou de suivre leur évolution au cours du temps. Il a été défini par Depledge et Fossi (Depledge and Fossi, 1994) comme étant « une réponse biologique à un agent chimique mesurant l'exposition ou l'effet toxique ».

En géochimie organique, le terme biomarqueur fait référence à un composé traceur, d'origine biologique, qui dérive de précurseurs biochimiques *via* des processus d'altération réducteurs et oxydatifs, et qui permet de renseigner sur l'origine, mais aussi le devenir d'une matière organique (Simoneit, 2002).

Dans ce contexte, la protéomique permet de décrire et quantifier des biomarqueurs protéiques exprimés par un organisme donné (uni ou pluricellulaire) et permet d'avoir une vision globale de sa réponse à certains stimuli ou perturbations extérieures (Hirsch et al., 2004).

Néanmoins, étant donné la présence d'une grande variété de protéines aux concentrations, localisations et fonctions très diverses (figure 26) (Frantzi et al., 2014; Hüttenhain et al., 2009; Schiess et al., 2009), la détection de tels biomarqueurs nécessitent de développer des protocoles d'extraction et de purification efficaces, et de les associer à des méthodes d'analyses ultrasensibles afin d'identifier et/ou de quantifier sans ambiguïtés une ou plusieurs protéines (Bastida et al., 2009; Schiess et al., 2009; Keiblinger et al., 2016; Frantzi et al., 2014; Keiblinger et al., 2012).

Introduction bibliographique

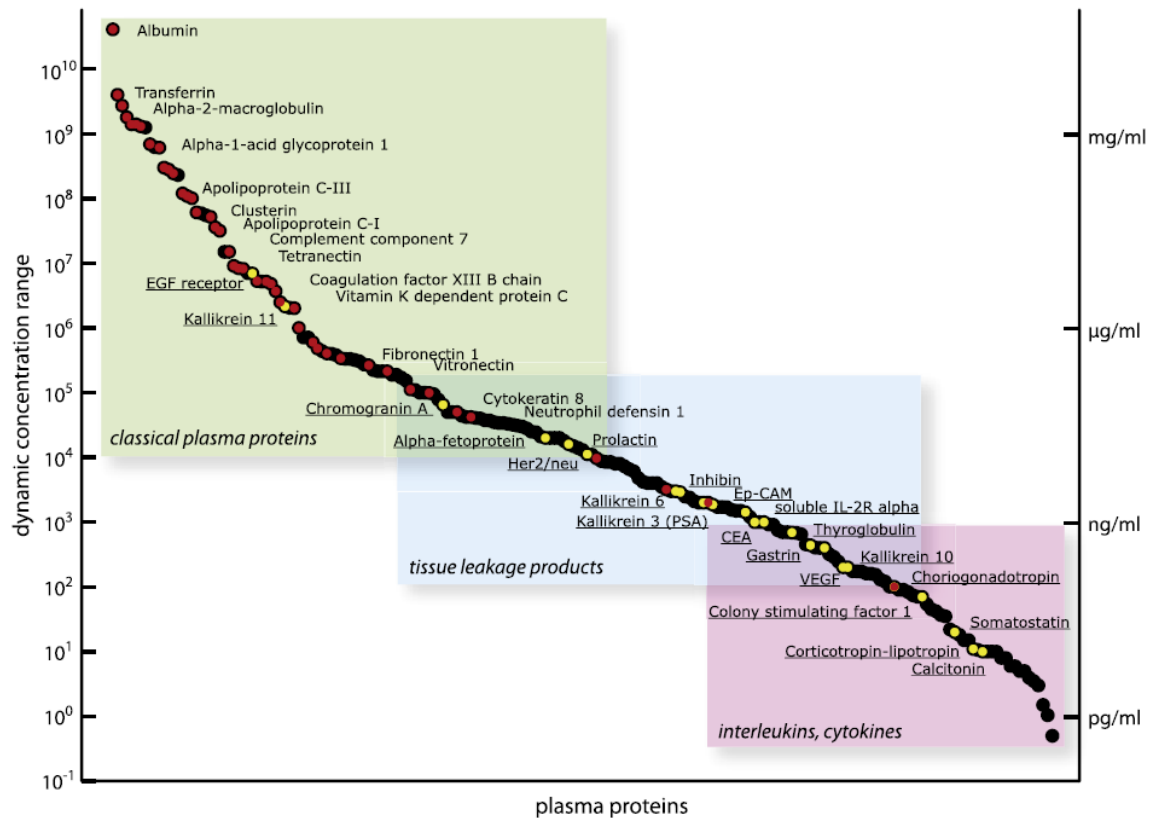


Figure 26. Illustration graphique de l'abondance des différentes protéines du plasma : les points jaunes représentent les biomarqueurs potentiels et sont situés essentiellement dans la gamme relative aux faibles concentrations (d'après Schiess et al., 2009)

2.2.2. Dans le domaine de la santé

La recherche de biomarqueurs est effectuée dans le cadre médical afin de (a) détecter précocement une maladie (tableau 1) (b) de suivre l'évolution d'une maladie (c) de prédire la réponse d'un organisme à un traitement thérapeutique (Biomarkers Definitions Working Group, 2001; Frantzi et al., 2014; Hüttenhain et al., 2009; Keshishian et al., 2007; Schiess et al., 2009).

Depuis plus d'une vingtaine d'années, de nombreuses protéines ont été identifiées et sont aujourd'hui suivies couramment dans le cadre de diagnostic médical. Une des plus connues correspond à l'hormone gonadotrophine humaine (β -HCG) produite au début de la grossesse. Le tableau ci-après présente une liste non exhaustive, de quelques biomarqueurs protéiques utilisés dans le cadre de diagnostics et suivis médicaux (maladies aiguës, chroniques, dopage, etc.)

Introduction bibliographique

Tableau 1. Exemples de biomarqueurs protéiques essentiels pour le diagnostic médical

Application	Matrice	Analyse effectuée pour détecter les cibles protéiques	Complexité de l'analyse, exprimée en nombre de protéines identifiées	Biomarqueurs potentiels	Fonction	Références
Cancer du foie	Cellules	2D-SDS-PAGE-MS/MS	11 protéines identifiées associées au AFP chez les malades	α -fœtoprotéine (AFP)	Au sein des cellules cancéreuses : régulation de l'apoptose, transport des acides gras, bilirubine, œstrogènes	(Yokoo et al., 2004)
Cancer de la prostate	Plasma	LC-MS/MS	-	Antigène prostatique (PSA)	Activité enzymatique, Liquéfaction du sperme suite à son émission	(Fortin et al., 2009)
Cancer de la thyroïde	Sérum	LC-MS/MS	-	Thyroglobuline (Tg)	Précurseur de la synthèse des hormones thyroïdiennes	(Hoofnagle et al., 2008)
Cancer de la rétine (Rétinoblastome)	Tissu rétinien	2D-DIGE-MS/MS	39 protéines identifiées et exprimées uniquement chez les malades	Protéine de liaison du rétinol (RBP3) Protéine glio-fibrillaire acide (GFAP)	RBP3 : protéine de transport GFAP : protéine de structure dans les cellules nerveuses (cellules gliales)	(Naru et al., 2017)

Introduction bibliographique

Grossesse, Cancer trophoblastiques gestationnelles, cancers testiculaires, Dopage	Urine, Plasma	LC-MS/MS	-	Hormone gonadotrophine humaine (β -HCG)	Maintien du corps jaune et développement du placenta, Stéroïde anabolisant chez les sportifs, Stimule la sécrétion du testostérone	(Lund et al., 2012)
Alzheimer Parkinson	Tissus cérébraux	2D-DIGE	-	α -synucléine Amyloïde β ($A\beta_{42}$)	Mort neuronal engendrée par l'agrégation des plaques amyloïdes	(Fernandez- Gomez et al., 2014)
Alzheimer	Liquide céphalo- rachidien	LC-MS/MS	-	Amyloïde β ($A\beta_{42}$) Protéine Tau	Mort neuronal engendrée par l'agrégation des plaques amyloïdes	(Pottiez et al., 2017)
Maladie des artères coronaires	Plasma	LC-MS/MS	Entre 122 et 159 protéines identifiées dont 18 exprimées uniquement chez les malades	Apolipoprotéine (A-I ; A-IV; C-I)	Transport des molécules hydrophobes (cholestérol)	(Basak et al., 2016)

2.2.3. Dans le domaine de l'environnement

Parmi les protéines présentes dans les matrices environnementales, peuvent être distinguées les protéines de structure (95 %) (Stevenson, 1994) et les protéines dites fonctionnelles (5 %) : les enzymes (Rillig, 2004). Les enzymes du sol jouent un rôle capital dans la dégradation des biopolymères de la litière. Elles interviennent aussi, dans la libération des nutriments au profit des plantes et dans le processus d'humification (Gobat et al., 2010). Les microorganismes, champignons et bactéries, sont les principaux producteurs des protéines du sol. Ainsi, de nombreuses activités enzymatiques ont pu être détectées dans les sols, telles que les hydrolases, les oxydoréductases, les transférases ou encore les lyases (Dick and Tabatabai, 1993).

Dans le domaine de l'environnement, les communautés microbiennes assurent le maintien de la biodiversité, *via* leurs activités dans les cycles biogéochimiques (carbone, azote et phosphore). Afin d'étudier ces populations, des approches combinant la méta-génomique, la méta-transcriptomique et la méta-protéomique ont été utilisées afin d'apporter des réponses quant à la diversité et au fonctionnement du consortium microbien présent dans ces milieux. Parmi celles-ci, la méta-protéomique a été définie comme « la science qui étudie l'ensemble des protéines exprimées par les organismes vivants au sein d'un écosystème, à un moment donné, et dans des conditions données » (Wilmes and Bond, 2004).

À partir des années 2000, ce champ disciplinaire a commencé à être appliqué à diverses matrices environnementales dans le but d'apporter des informations pertinentes sur le fonctionnement des communautés bactériennes, ainsi que sur leurs interactions au sein de cet écosystème (Benndorf et al., 2007). Depuis, la diversité et les fonctions de certains microorganismes présents dans les milieux naturels tels que des sols (Bastida et al., 2015; Hultman et al., 2015; Singleton et al., 2003; Wang et al., 2011; Williams et al., 2010; Zampieri et al., 2016; Benndorf et al., 2007), des biofilms présents sur des sédiments (Benndorf et al., 2009), de milieux marins (Dong et al., 2010; Kleiner et al., 2012; Morris et al., 2010; Wang et al., 2014) et des eaux douces (Russo et al., 2016), ont pu être déterminées (tableau 2).

Introduction bibliographique

Tableau 2. Exemples de biomarqueurs protéiques essentiels pour le diagnostic environnemental

Application	Matrice	Analyse effectuée pour détecter les cibles protéiques	Complexité de l'analyse, exprimée en nombre de protéines identifiées	Biomarqueurs potentiels	Fonction	Références
Contamination aux produits chlorés	Sol Eau souterraine	2D-PAGE-LC-MS/MS	26 protéines identifiées	Chlorocatécol dioxygénase, Acétate 2,4-dichlorophénoxy dioxygénase, Chlorobenzène dioxygénase	Dégradation des contaminants organiques par un consortium microbien	(Benndorf et al., 2007)
Contamination aux métaux lourds	Sol	SDS-PAGE	-	Les métallothionéines	Régulation des teneurs en métaux essentiels (Cuivre, zinc) et détoxification des métaux potentiellement toxiques (cadmium, mercure)	(Singleton et al., 2003)
Amendements organiques : toluène, glucose	Sol	SDS-PAGE-MS/MS	47 protéines identifiées	Protéines de choc au froid (CSP) de <i>Pseudomonas syringae pv</i> , Superoxyde dismutase de <i>Pseudomonas fluorescens</i>	Résistance aux stress thermiques, Réponse aux stress oxydatifs	(Williams et al., 2010)

Introduction bibliographique

Amendements organiques : compost et boues de STEP	Sol	LC-MS/MS	10818 protéines identifiées	Cellulases, Lipases, Uréases, Phosphatase, Phénol oxydases	Protéines microbiennes impliquées dans les cycles biogéochimiques	(Bastida et al., 2015)
Contamination aux BTEX	Sédiment	2D-PAGE-LC-MS/MS	240 protéines identifiées	Enoyl-CoA hydrolase d' <i>Azoarcus evansii</i>	Dégradation des xénobiotiques	(Benndorf et al., 2009)
Impact d'un milieu oligotrophe et eutrophe	Eau douce	2D-LC-MS/MS	1048 protéines identifiées	Enzyme photosystème II, Ribulose 1,5-biphosphate, Carboxylase/oxygénase	Photosynthèse Fixation du carbone	(Russo et al., 2016)
Pollution aux métaux	Ecrevisse de Louisiane (<i>Procambarus clarkii</i>)	2D-DIGE-MS	2136 protéines identifiées	Protéine de choc thermique (HSP 70)	Résistance au stress thermique	(Fernández-Cisnal et al., 2017)
Pollution au cadmium	Eaux usées	2D-SDS-PAGE-MS/MS	100 protéines identifiées	ATPases, Oxydoréductases, Catalase, Superoxyde dismutase	Détoxification du cadmium	(Lacerda and Reardon, 2009)

***Chapitre II : De la préparation
d'échantillon à l'analyse des protéines***

Introduction bibliographique

L'étude du protéome d'un échantillon donné repose sur une méthodologie couplant une technique de séparation à une technique d'identification des protéines, en général la spectrométrie de masse.

Il existe deux stratégies principales : la stratégie dite « *top-down* » et la stratégie dite « *bottom-up* » (figure 27). En fonction de l'une ou l'autre des stratégies, la préparation de l'échantillon sera bien différente.

La stratégie « *top-down* » consiste à analyser les protéines intactes sans hydrolyse chimique ou enzymatique au préalable. Celles-ci sont donc injectées directement dans un spectromètre de masse couplé ou non à une chromatographie liquide, ses ions moléculaires comme ceux issus de sa fragmentation MS/MS sont alors analysés (Gillet et al., 2016). Un des intérêts majeurs de cette stratégie est la caractérisation des protéoformes. Ces entités désignent les différentes formes moléculaires d'une protéine caractérisées des modifications post-traductionnelles (Toby et al., 2016). En outre, l'analyse directe des protéines « intactes » permet de couvrir la totalité de la séquence protéique permettant ainsi un gain d'informations sur l'identité de la molécule.

Cette approche présente toutefois quelques inconvénients, tels que la complexité des mélanges protéiques comparés aux mélanges peptidiques (obtenus par l'approche « *bottom-up* »). Elle nécessite donc des techniques et des méthodologies de pointe pour la séparation et l'analyse des protéines par spectrométrie de masse (Catherman et al., 2014).

L'approche « *bottom-up* » repose quant à elle sur l'analyse de mélanges peptidiques issus de l'hydrolyse enzymatique des protéines (souvent en présence de trypsine). Cette approche est la plus répandue pour l'analyse protéomique par spectrométrie de masse (Aebersold and Mann, 2016; Pragya et al., 2017). L'identification par masse de la protéine repose donc sur une comparaison des spectres expérimentaux des peptides obtenus après leur hydrolyse, aux spectres théoriques de ces peptides, puis d'une interrogation des bases de données. Des logiciels tels que *Mascot*, *X!tandem* ou encore *Sequest* sont d'ailleurs très utilisés pour l'identification des protéines en se basant sur les données de fragmentation MS/MS (Cottrell, 2011).

Néanmoins, l'identification des protéines peut parfois s'avérer être difficile, car elle dépend indéniablement de la qualité et de la reproductibilité de la digestion enzymatique. Ainsi, la présence de peptides communs à plusieurs protéines peut induire des fausses identifications.

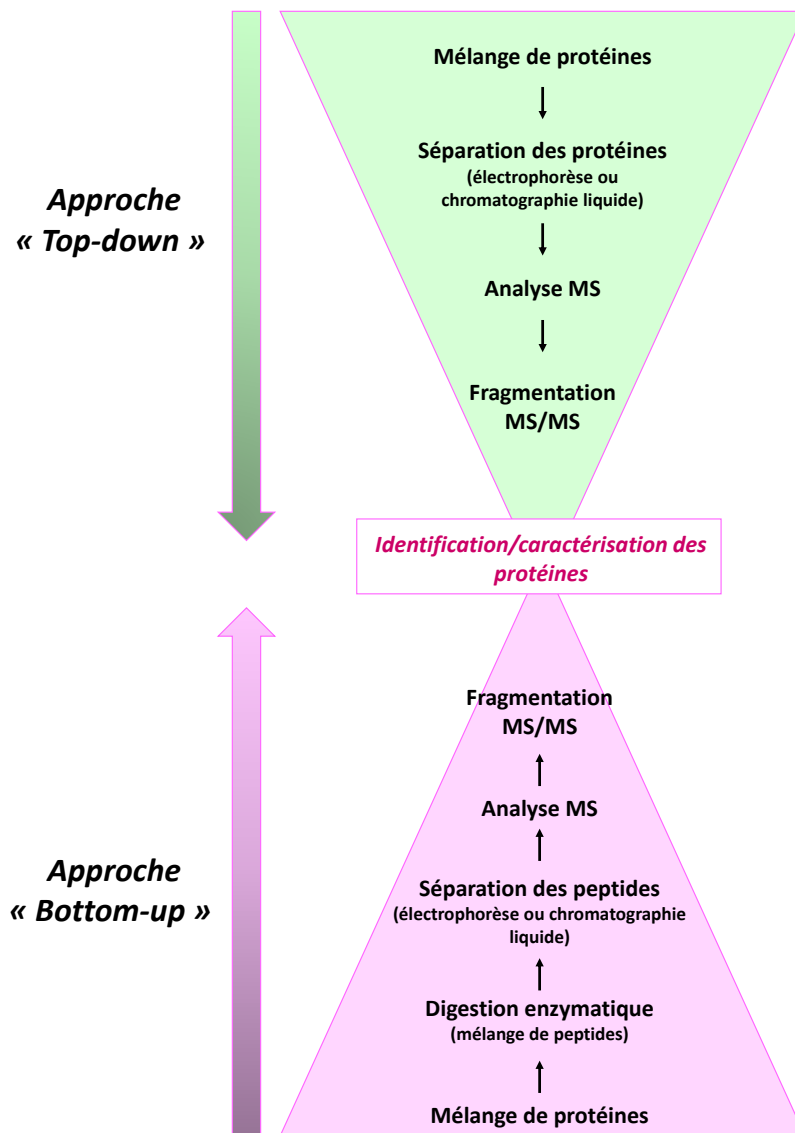


Figure 27. Représentation de la stratégie « *top-down* » et « *bottom-up* »

L'approche « *bottom-up* » qui est très largement utilisée, se décompose en plusieurs étapes (figures 27 et 28) : (a) une extraction et purification des protéines à partir d'une matrice donnée (milieu biologique ou environnemental) ; (b) une hydrolyse des protéines en fragments peptidiques ; (c) une séparation des peptides par électrophorèse et/ou chromatographie liquide (d) et une analyse de ces peptides par spectrométrie de masse.

La préparation de l'échantillon en protéomique est le plus souvent adaptée à la matrice étudiée en vue d'obtenir une meilleure détection des protéines d'intérêt.

Introduction bibliographique

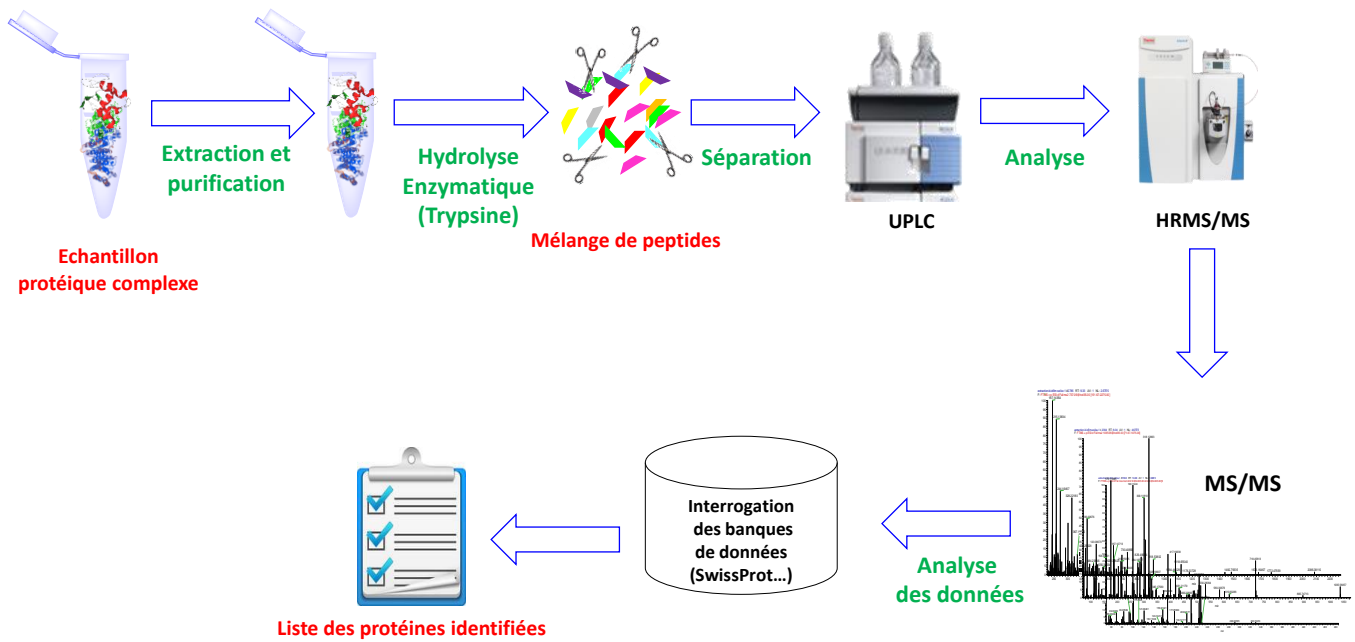


Figure 28. Approche ascendante « *bottom-up* » pour l'identification des protéines par spectrométrie haute résolution (HRMS/MS)

1. Préparation de l'échantillon

1.1. Extraction des protéines

L'étape d'extraction a pour objectif de séparer les protéines de la matrice dans laquelle elles se trouvent ce qui nécessite leur solubilisation dans un milieu extractant le plus souvent liquide. Pour cela, il faudra rompre les interactions fortes et/ou faibles entre les protéines et les composés non protéiques présents dans le milieu. Les conditions d'extraction doivent être strictement contrôlées et optimisées afin de minimiser non seulement les pertes en protéines, mais aussi de limiter les contaminations exogènes et endogènes du milieu (composés non protéiques qui interfèrent avec l'analyse par LC-MS/MS).

1.1.1. Extraction solide/liquide

L'extraction solide/liquide des protéines, bien que reposant sur des principes similaires aux extractions dédiées aux autres familles (sucres, lipides, pigments, etc.), présente quelques particularités de par la polarité et le haut poids moléculaire de ces biomolécules.

1.1.1.1. Principe général

De manière générale, l'extraction du protéome à partir d'un tissu biologique ou d'une matrice donnée repose sur l'utilisation de solutions tampons associant des détergents tels que le CHAPS (3-diméthylammonio-1-propanesulfonate), le triton X-100, le Tween 20, le SDS (dodécylsulfate de sodium), etc., à des agents chaotropes tels que l'urée et thiourée à des concentrations élevées (7 M et 2 M respectivement), ou encore à des agents réducteurs comme le DTT (dithiothréitol) ou le β -mercaptoéthanol (réducteurs des ponts disulfures) (Chen et al., 2009; Chevallet et al., 1998; Chourey et al., 2010; Gumber et al., 2007; Rabilloud et al., 2007; Richardson et al., 2008; Song et al., 2016). Lorsque, les protéines recherchées sont présentes dans le cytosol (fraction liquide du cytoplasme), l'ajout de solvant est associé à des méthodes mécaniques ou chimiques de lyse cellulaire telles que l'hydrolyse enzymatique, les ultrasons, la presse de French (cylindre creux dans lequel est enfoncé un piston pour lyser les cellules sous pression), le choc osmotique ou thermique, le broyage mécanique ou encore l'homogénéisation (Feist and Hummon, 2015; Cañas et al., 2007).

L'extraction au moyen de solutions contenant des détergents et/ou dénaturants déstabilisent les interactions faibles énergies au sein des protéines de telle sorte qu'elles perdent leur conformation native (dénaturation) et à les rendre plus solubles dans la solution d'extraction (figure 29).

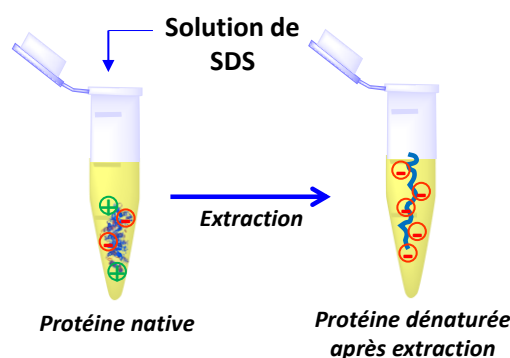


Figure 29. Extraction des protéines en présence de détergents anioniques

1.1.1.2. Applications

Ces méthodes d'extraction ont notamment été utilisées pour la mise en évidence de biomarqueurs protéiques, tels que l' α -foetoprotéine (AFP) marqueur du cancer du foie (Yokoo

Introduction bibliographique

et al., 2004) ou encore la protéine de choc thermique (*heat shock protein* (HSP)) et la phénol-oxydase présents dans un sol soumis à un stress environnemental (Zampieri et al., 2016).

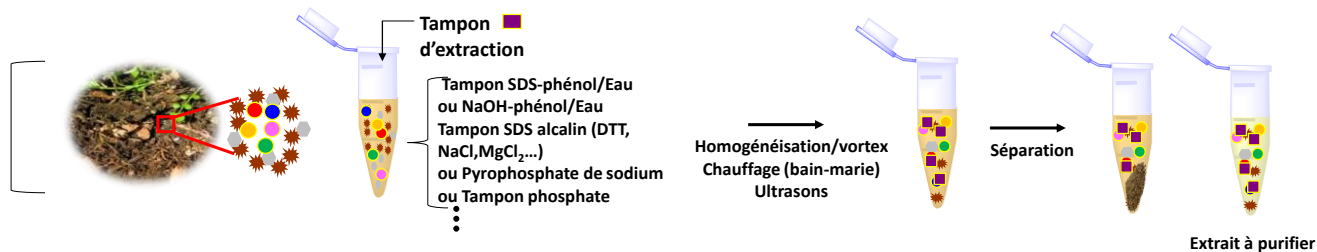
Bien qu'ils soient moins complexes en termes de nombre et de variabilité protéique, les fluides biologiques représentent également un obstacle majeur à l'analyse protéomique. Des étapes de prétraitement de l'échantillon sont nécessaires afin de dépléter des protéines majoritaires, notamment de l'albumine, l'immunoglobuline G (IgG) et l'haptoglobine, pour accroître la réponse des protéines susceptibles d'être des biomarqueurs potentiels (Kuhn et al., 2004). On peut ainsi citer le développement analytique réalisé pour l'extraction de la C-réactive (CRP) dans le plasma de personnes atteintes de polyarthrite rhumatoïde.

À l'inverse, la complexité protéique des matrices environnementales est principalement liée aux faibles concentrations de protéines (0,2 mg protéines/ g de sol) et à la présence d'interférents. La figure 30 illustrée ci-dessus récapitule quelques exemples de protocoles d'extraction par solvant utilisés dans le domaine de l'environnement et dans la biologie.

Introduction bibliographique

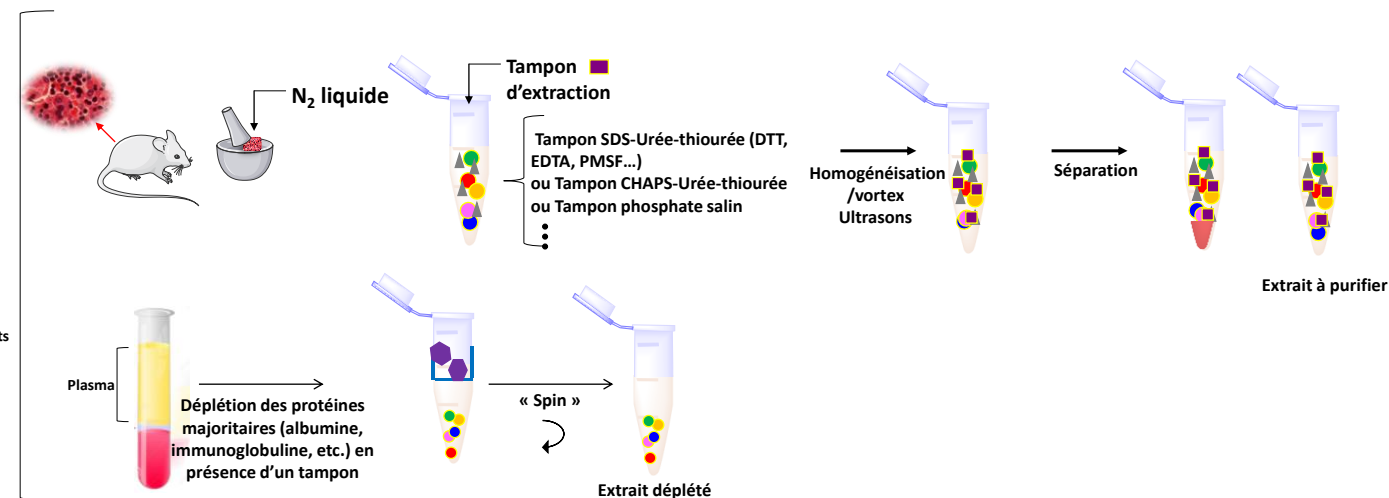
Matrice environnementale

Protéines d'intérêt
argile
Substances humiques



Matrice biologique

Protéines d'intérêt
Protéines majoritaires
interférents



Exemples de biomarqueurs environnementaux

- Phénols-oxydases, protéine de choc thermique (HSP); 638 protéines identifiées (Zampieri et al., 2016)
- Enzymes qui dégradent les herbicides, 26 protéines identifiées (Benndorf et al., 2007)

Exemples de biomarqueurs biologiques

- α -fœtoprotéine (AFP); 11 protéines identifiées (Yokoo et al., 2004)
- La protéine C-réactive (CRP), immunodépletion de l'albumine, et Ig (Kuhn et al., 2004)

Figure 30. Exemples de méthodes d'extraction des protéines par solvant

Introduction bibliographique

1.1.1.3. Verrous

Du fait de l'hétérogénéité des matrices biologiques et environnementales, l'extraction des protéines par solvant reste une étape limitante, surtout lors de l'étude des milieux hétérogènes tels que les matrices biologiques ou environnementales. De manière générale, dans ce type de milieu les protéines se trouvent dans leur état natif, associées à d'autres molécules de même nature ou ancrées aux membranes cellulaires. Or, une cellule (procaryote ou eucaryote) contient une multitude de composés endogènes souvent à des concentrations bien plus importantes que les protéines recherchées. En outre, un très grand nombre de protéines sont hydrophobes, insolubles dans les tampons d'extraction et donc difficiles à extraire (Cañas et al., 2007).

Le principal verrou lié à l'étude du protéome dans des échantillons biologiques correspond à la présence de protéines très abondantes qui peuvent interagir avec les marqueurs protéiques présents à de très faibles concentrations. Telle que démontrée par Granger et al. (2005), après déplétion de l'albumine, la concentration des protéines d'intérêt contenues dans un échantillon de plasma humain a diminué drastiquement, mettant en évidence la perte des protéines faiblement abondantes (Granger et al., 2005). La figure 31 représente un gel d'électrophorèse bidimensionnel d'un échantillon de plasma humain. Les résultats montrent bien la complexité de ce fluide biologique avant déplétion des protéines les plus abondantes.

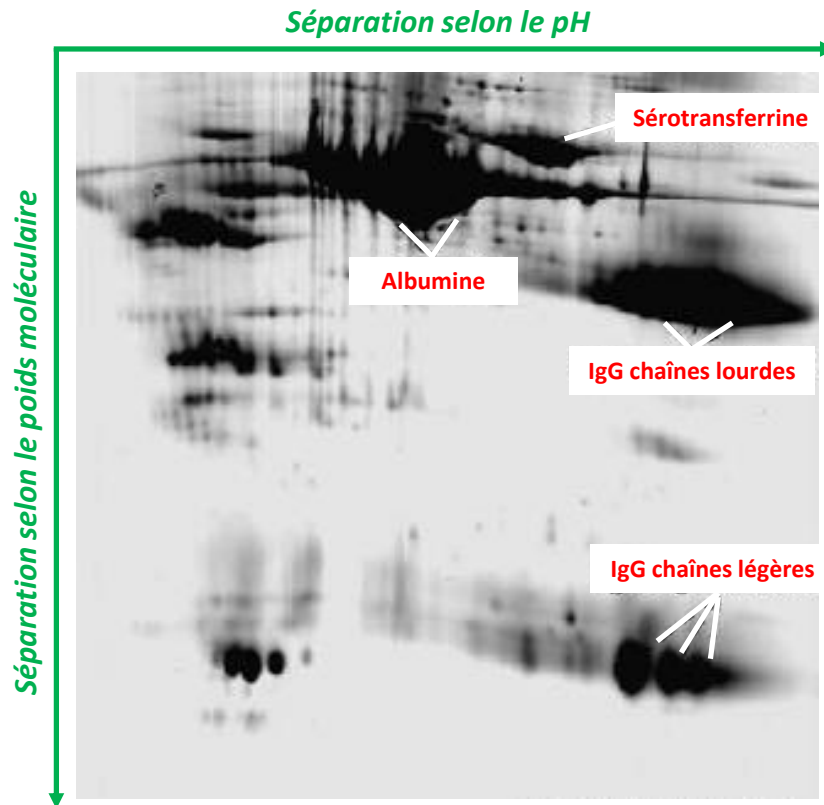


Figure 31. Détection sur gel d'électrophorèse bidimensionnelle des protéines les plus abondantes dans le plasma humain (Jiang et al., 2004)

Outre les problèmes liés aux protéines majoritaires et autres constituants, la préparation de l'échantillon doit également permettre de détecter des protéines à l'état de traces. Cette difficulté concerne notamment les matrices environnementales telles que les sols où les protéines représentent entre 70 et 80 % de l'azote organique total ($0,40 - 20,4 \text{ g. kg}^{-1}$), soit environ 0,2 mg de protéines totales pour 1 g de sol (Greenfield, 1972). Les particularités de la matrice sol rendent par ailleurs difficile l'extraction ainsi que la purification des extraits protéiques (Gobat et al., 2010). En effet, la composition géochimique (argile, matières organiques dont les acides humiques), l'hétérogénéité et les paramètres physicochimiques (granulométrie, porosité, température, pH, teneur en eau) sont des facteurs extrêmement variables d'un sol à autre. Dans ce cas, les protéines se trouvent généralement adsorbées aux particules colloïdales, argileuses ou aux substances humiques (Kiss et al., 1975; Schmidt and Martínez, 2016), avec lesquelles elles mettent en jeu de nombreuses interactions ioniques, hydrogènes et hydrophobes (Tomaszewski et al., 2011; Yu et al., 2013a). Une fois encapsulées ou piégées dans ces systèmes, les protéines résistent aux variations physicochimiques de

Introduction bibliographique

l'environnement qui ont tendance à les dénaturer, mais n'en deviennent que plus difficiles à analyser (Lähdesmäki et Pnspanen, 1992) (figure 32).

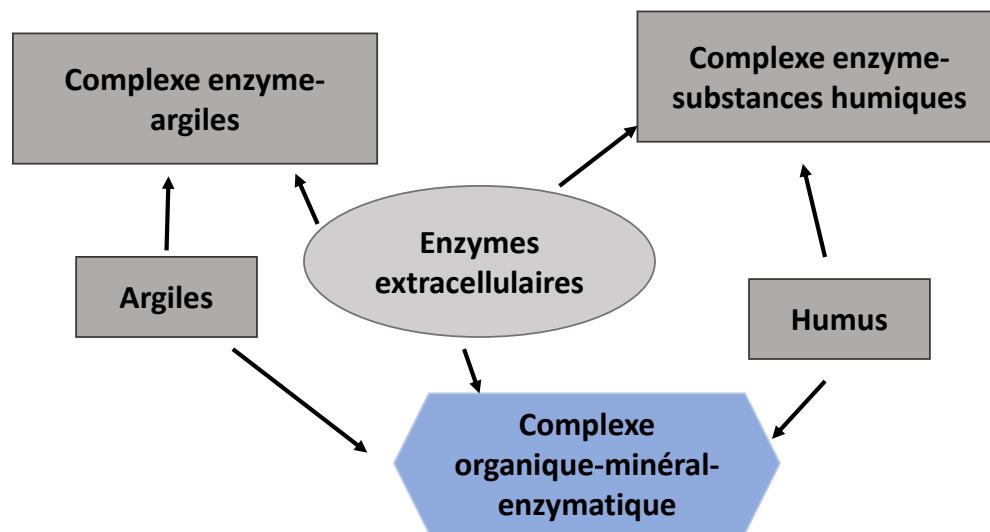


Figure 32. Complexation des enzymes extracellulaires aux composés de sols (Nannipieri et al., 1996)

L'extraction par solvant présente l'avantage de récupérer l'ensemble des protéines présentes dans un milieu donné. Cette technique se heurte cependant à sa non sélectivité lorsque la concentration des marqueurs protéiques d'intérêt est extrêmement faible au regard des protéines les plus abondantes dans le milieu. Face à ces verrous, des techniques d'extraction plus sensibles/ciblées telles que l'immunoextraction est préférentiellement employée.

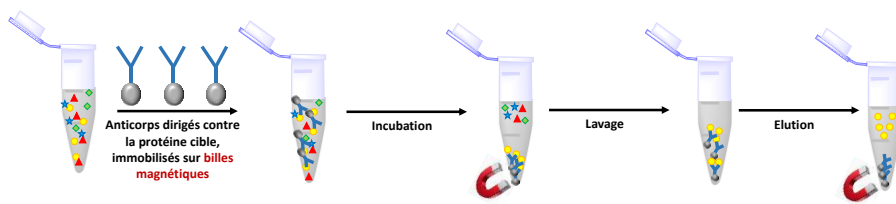
1.1.2. Extraction par immunoextraction

1.1.2.1. Principe

Les techniques d'immunoextraction reposent sur l'utilisation de supports d'extraction solides sélectifs, sur lesquels sont fixés des anticorps spécifiques dirigés vers les protéines d'intérêt (figure 33).

Introduction bibliographique

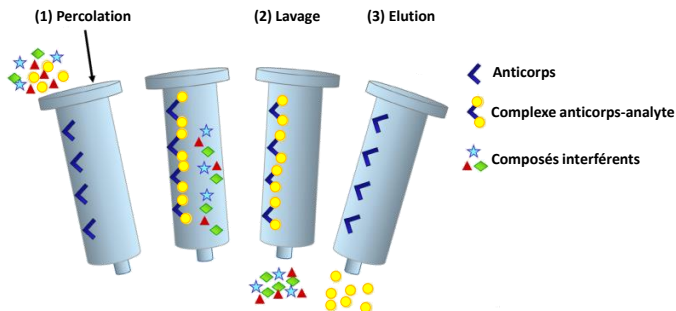
Immunoextraction sur billes magnétiques



Exemples de biomarqueurs

- Tyroglobuline (LOD pmol.mL^{-1}) (Hoofnagle et al., 2008)
- β -lactoglobuline et α -lactalbumine du lait (LOD < nmol.mL^{-1})

Immunoextraction sur support



Exemples de biomarqueurs

- Endomorphine (LOD ng.mL^{-1}); support : silice (Medina-Casanellas et al., 2012)
- Hormones œstrogènes (LOD pg.mL^{-1}); support : agarose (Hosogi et al., 2010)

Figure 33. Représentation schématique du principe de l'immunoextraction sur billes magnétiques et sur support (Płotka-Wasyłka et al., 2015)

Cette technique met en jeu des mécanismes de reconnaissance moléculaire entre un anticorps et antigène (représenté dans notre cas pas la protéine). La grande sélectivité du complexe anticorps-protéine permet d'obtenir de manière sélective un extrait pur sur le support et d'éliminer de façon efficace tous les autres composés interférents responsables des effets matrices qui interfèrent avec l'analyse par spectrométrie de masse (Hennion and Pichon, 2003).

L'étape d'éluion qui consiste à désorber les molécules cibles est cependant délicate. Elle doit être minutieusement contrôlée afin de ne pas affecter l'intégrité des anticorps immobilisés sur le support. Elle consiste dans la plupart du temps en une variation du pH du milieu afin d'induire une modification de l'état d'ionisation des composés et ainsi perturber les interactions du complexe anticorps-protéine. D'autres méthodes d'éluion peuvent être aussi employées. Celles-ci consistent à utiliser des mélanges de solvants hydro-organiques ou encore des agents chaotropes (Godfrey, 1998; Hennion and Pichon, 2003).

1.1.2.2. Applications

Le développement de cette méthode a signé une avancée remarquable dans le domaine médical. À titre d'exemple, Hoofnagle et al. (2008), ont extrait et détecté à des concentrations

Introduction bibliographique

de l'ordre du picomolaire la thyroglobuline (Tg), marqueur tumoral sérique du cancer de la thyroïde (Hoofnagle et al., 2008). De la même façon, l'hormone gonadotrophine humaine (*human chorionic gonadotropin*, β -hCG), marqueur de grossesse et de dopage, a été détectée dans deux fluides biologiques, urine et plasma humain, par immunocapture sur des billes magnétiques.

Parmi les marqueurs protéiques aujourd'hui détectés par immunoextraction, nous pouvons citer la ricine extraite à partir de la graine de ricin, protéine très toxique d'origine végétale, qui lorsqu'elle est inhalée entraîne le décès d'un individu après quelques jours (agent d'attaque bioterroriste préoccupant particulièrement les autorités politiques aujourd'hui (Becher et al., 2007)). Cette méthode repose sur la capture de la protéine cible, qui est ici la ricine, au moyen d'anticorps fixés sur des billes magnétiques activées par la protéine G. La ricine peut être aussi concentrée, après immunoextraction, grâce à un aimant puis hydrolysée par la trypsine libérant ainsi des marqueurs peptidiques, qui sont analysés par spectrométrie de masse.

Cette technique est peu rencontrée dans le domaine de l'environnement. Une liste de marqueurs spécifiques n'est en effet pas disponible à l'heure actuelle. Cependant, elle pourra dans les années à venir parfaitement s'intégrer en aval d'une phase d'extraction par solvant.

1.1.2.3. Verrous

Les techniques d'immunoextraction présentent une très forte spécificité vis-à-vis des molécules d'intérêt. Néanmoins, cette spécificité constitue également leur talon d'Achille, car même si ces techniques sont dotées d'une grande facilité de mise en œuvre, elles ne permettent pas d'accéder à l'ensemble du protéome présent dans une matrice donnée, ce qui limite l'extraction, à un nombre restreint de protéines cibles. De même les coûts liés à la production d'anticorps spécifiques sont relativement élevés.

1.2. Purification des protéines

1.2.1. Précipitation, lavage et solubilisation

L'extraction des protéines d'une matrice donnée conduit le plus souvent à une co-extraction de composés interférents, non protéiques tels que les sels, les acides nucléiques, les polysaccharides, les lipides, ou encore les détergents initialement présents dans les solutions d'extraction (Cañas et al., 2007). En effet, ces derniers sont généralement incompatibles avec les méthodes séparatives qui suivent ainsi qu'avec l'analyse par spectrométrie de masse. De ce

Introduction bibliographique

fait, la présence de détergents peut gêner la détection, voire conduire à une suppression du signal des composés recherchés (Zhou et al., 2012).

De nombreuses techniques peuvent être utilisées pour purifier l'extrait protéique. L'ultrafiltration (Jiang et al., 2004), la chromatographie ou encore la dialyse sont couramment utilisées (Voet and Voet, 2016). Toutefois, la précipitation reste incontestablement la méthode de purification la plus communément employée en protéomique (Feist and Hummon, 2015).

Généralement, la précipitation des protéines a lieu dans des solutions à force ionique élevée, en présence de forte concentration en solvants organiques (Beretov et al., 2015). Les agents de précipitation les plus utilisés dans les matrices biologiques et environnementales sont : l'acétone, additionnée ou non de sels (NaCl) (Crowell et al., 2013; Jiang et al., 2004; Sun et al., 2005), l'acétate d'éthyle (Yeung and Stanley, 2010), l'acide trichloroacétique (TCA/eau), (Manadas et al., 2006; Nandakumar et al., 2003, 2006; Pottiez et al., 2017; Rajalingam et al., 2009; Zhang et al., 2011), les mélanges TCA/acétone, en particulier sur les matrices végétales (R. Hao et al., 2015; Isaacson et al., 2006; Song et al., 2016, 2012a; Wu et al., 2014), le méthanol/chloroforme (Fic et al., 2010; Vertommen et al., 2010; Wessel and Flüggé, 1984) ou encore l'éthanol (Thongboonkerd et al., 2009).

Depuis les années 50 (Henry et al., 1956), le TCA (entre 10 et 25 % (w/v)), reste l'agent de précipitation le plus utilisé. Il a encore récemment démontré son efficacité pour la détection des marqueurs de la maladie d'Alzheimer dans le liquide céphalorachidien (protéine Tau et le peptide β -amyloïde (1-42)) à des concentrations de l'ordre du ng.mL^{-1} (Pottiez et al., 2017).

Une fois les protéines précipitées, une étape de lavage est réalisée afin d'éliminer l'excédent de sels ou d'agents de précipitation et ainsi garantir une bonne solubilisation des protéines dans un tampon d'analyse. En effet, certaines protéines restent insolubles formant ainsi un culot protéique compact (Feist and Hummon, 2015). Ceci est en particulier vrai lorsque la précipitation est réalisée avec du TCA (Nandakumar et al., 2003; Rajalingam et al., 2009; Zhang et al., 2011). Dans ce cas précis, des lavages successifs sont employés avec des solvants organiques (l'acétone ou l'éthanol) (Beretov et al., 2015; Murase et al., 2003; Pottiez et al., 2017; Song et al., 2016; Benndorf et al., 2007). L'acétone est le solvant le plus utilisé pour laver un culot protéique obtenu après précipitation au TCA. Néanmoins, de nombreux auteurs ont mis en évidence les limites liées à son utilisation (perte des protéines et mauvaise solubilisation

Introduction bibliographique

avant analyse) (R. Hao et al., 2015; Jiang et al., 2004; Nandakumar et al., 2003; Zhang et al., 2015).

Pour cette raison, des protocoles de solubilisation plus complexes ont été développés. Ils combinent à la fois des solutions de détergents et d'agents chaotropes connus pour améliorer, de manière considérable, la solubilité des protéines en milieu aqueux (Gong et al., 2008; Herbert, 1999).

Des traitements complémentaires aux ultrasons (Araújo et al., 2014; Manadas et al., 2006), aux microondes, par agitation vortex ou par chauffage (Ye and Li, 2012) sont aussi utilisés afin d'assurer une bonne solubilisation des protéines. La figure 34 récapitule les principaux protocoles utilisés pour la purification des protéines.

Introduction bibliographique

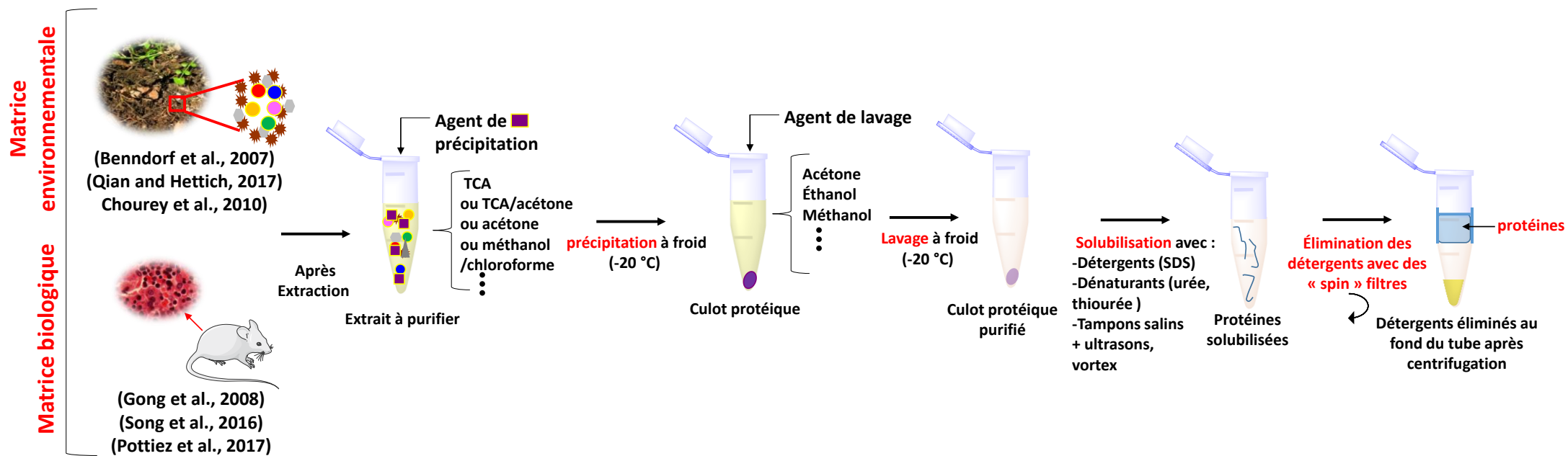


Figure 34. Purification des protéines à partir de matrices environnementales ou biologiques

Introduction bibliographique

Compte tenu, des limites liées à l'extraction des protéines, l'étape de purification se doit être la plus simple et la plus efficace possible afin de minimiser les pertes en protéines sans interférer avec la digestion enzymatique et enfin accroître la sensibilité de l'analyse.

1.2.2. Extraction sur phase solide (SPE)

L'extraction des composés sur phase solide ou SPE est une technique basée sur le partage des composés entre une phase liquide (l'échantillon) et une phase solide (adsorbant ou résine). Cette technique facile à mettre en œuvre repose sur un principe de séparation chromatographique qui utilise l'adsorbant comme phase stationnaire et l'éluant de l'échantillon comme phase mobile (Callesen et al., 2009).

Il existe une très grande variété de supports pour la purification des protéines, classés le plus souvent en fonction de la nature des interactions mises en jeu avec l'analyte. On distingue ainsi les supports à polarité inversée générant des interactions hydrophobes avec la molécule cible. Les plus utilisés sont généralement constitués de particules de silice greffées avec des chaînes alkyles allant de 4 à 18 atomes de carbone (C₄) et (C₁₈). La SPE en phase inverse est souvent réalisée en amont d'une analyse LC-MS. Cette technique, simple et reproductible, a notamment permis d'aller plus loin dans la lecture du protéome du plasma humain chez des individus sains ou malades (exemple cancer du sein)(Aresta et al., 2008; Callesen et al., 2009).

Des supports SPE polaires, mettant en jeux des interactions hydrophiles, sont également utilisés pour l'extraction des petits composés polaires tels que les sucres, les oligonucléotides ou encore les acides aminés. Cette technique est couramment employée lors d'approches de glycoprotéomique sur des matrices de type plasma, notamment pour renseigner sur certaines maladies comme le cancer du pancréas (Ahn et al., 2015; Callesen et al., 2009; Zhao et al., 2007).

En fonction du pH du milieu, les protéines ionisées, peuvent être séparées sur des phases stationnaires échangeuses d'ions. Différents supports d'extraction peuvent être employés (*silica strong anion exchange*, SAX ; *strong cation exchange*, SCX ; *weak cation exchange*, WCX ; *mixed cation exchange*, MCX) tels que des supports greffés avec des groupements de types ammonium quaternaire (SAX), ou sulfonates (SCX) ou acides carboxyliques (WCX), jouant respectivement le rôle d'échangeurs d'anions et de cations. En fonction des supports, les phases éluantes auront un pH croissant ou décroissant afin de décrocher les composés fixés. Ces

Introduction bibliographique

supports d'échangeurs d'ions peuvent être également combinés à une phase hydrophobe, on parle alors d'un mode mixte (Fortin et al., 2009).

Des supports sur lesquels sont immobilisées des molécules biologiques ayant une affinité pour une molécule d'intérêt permettent une purification spécifique des ligands synthétiques (dérivés de la purine et pyrimidine), biologiques (fragments d'ADN et d'ARN), des coenzymes, des anticorps, etc. (Callesen et al., 2009), peuvent être greffés sur les supports SPE. Parmi ces supports nous pouvons citer les supports d'affinité sur métal chélaté (*Immobilized metal ion affinity chromatography, IMAC*), mettant en jeu des mécanismes d'adsorption entre d'une part les groupements phosphates chargés négativement des protéines et des peptides phosphorylés, et d'autre part l'ion chélaté immobilisé sur le support (exemple l'oxyde de titanium (Yao et al., 2017)) (figure 35).

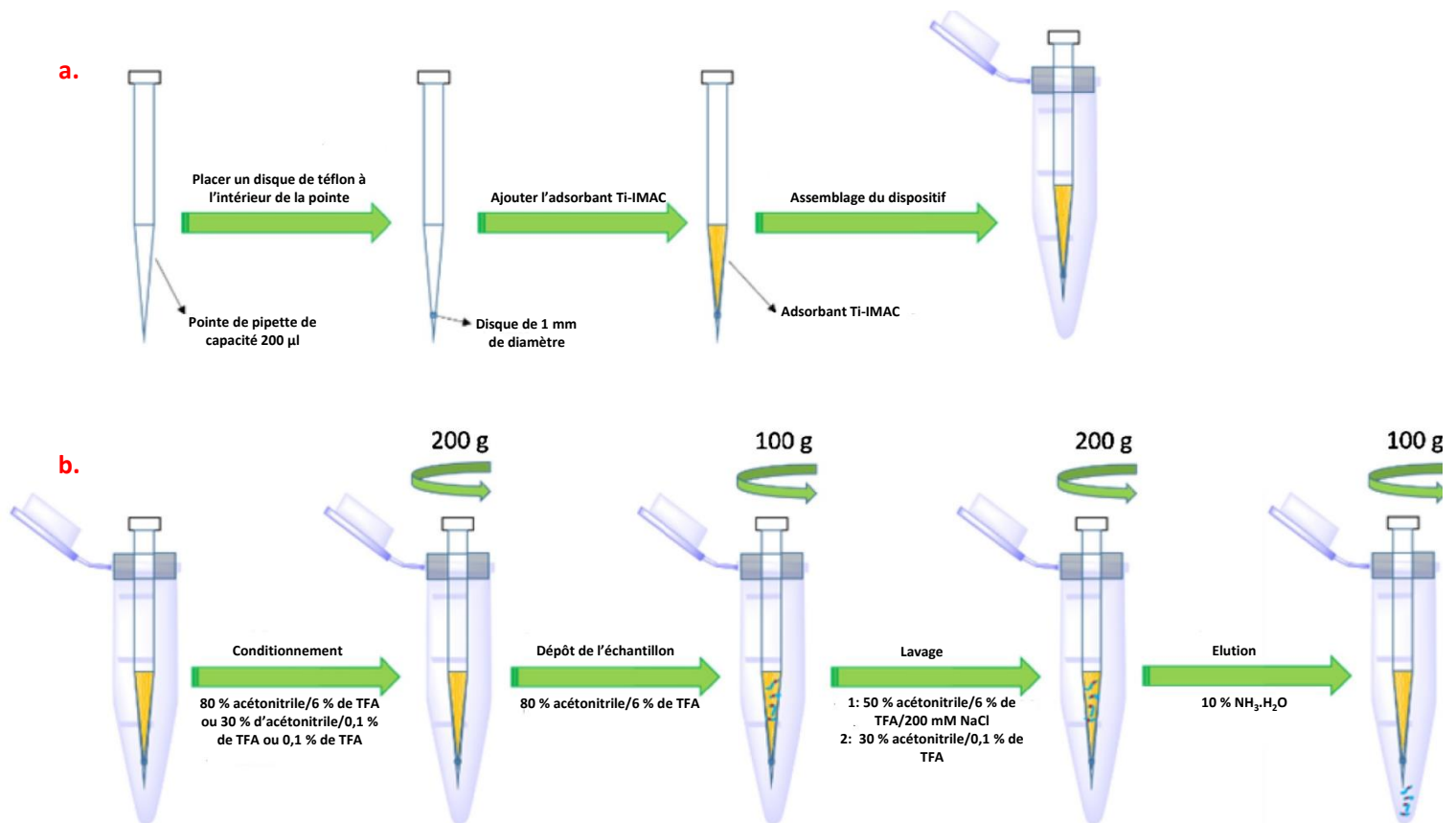


Figure 35. Préparation et application du support Ti^{4+} -IMAC ; a. Représentation schématique des étapes d'assemblage du dispositif Ti^{4+} -IMAC ; b. Centrifugation assistée par une extraction sur phase solide au moyen d'un Ti^{4+} -IMAC (Yao et al., 2017)

Grâce au développement des différentes phases stationnaires ainsi que leurs conditionnements, la SPE a pris une place importante dans la préparation des échantillons

Introduction bibliographique

protéiques. Cette méthode présente l'avantage d'être automatisable permettant ainsi d'extraire, de purifier et de concentrer plus facilement un grand nombre de molécules de natures chimiques différentes. Toutefois, les principaux inconvénients de la SPE restent indéniablement le prix des colonnes qui sont dans la plupart du temps à usage unique, également la perte en composés d'intérêts lors des différentes étapes de traitement ou encore le temps de réalisation relativement long lorsqu'elle n'est pas automatisée.

1.3. Hydrolyse enzymatique des protéines

L'approche protéomique « *bottom-up* » passe par une étape d'hydrolyse enzymatique. Cette hydrolyse des protéines, aussi appelée digestion, consiste à rompre certaines de leurs liaisons peptidiques de manière à obtenir des fragments plus petits, des peptides de 4 à plus de 50 acides aminés en fonction de la taille de la protéine. Cette opération est réalisée au moyen d'une protéase (généralement la trypsine, la papaïne, la chymotrypsine ou la pepsine). Avant toute digestion enzymatique et afin de faciliter l'accès de la protéase aux sites spécifiques de clivage, une étape de réduction des ponts disulfures, en présence de dithiothréitol (DTT), est réalisée. La deuxième étape consiste ensuite à bloquer les fonctions thiols et éviter la reformation des ponts disulfures. Classiquement, l'iodoacétamide (IAA) est utilisé comme agent alkylant (Hustoft et al., 2012).

Parmi les enzymes utilisées, la trypsine (endopeptidase) est la plus fréquemment employée. Elle coupe spécifiquement les liaisons peptidiques après les acides aminés chargés positivement, lysine (Lys) et arginine (Arg), sauf s'ils sont suivis d'une proline (Pro) (figure 36). L'activité optimale de la trypsine se situe à pH légèrement basique entre 7,5 et 8,5 et à une température de 37 °C (Hustoft et al., 2012). En règle générale, le rapport massique (masse de l'enzyme E (μg) par rapport à la masse de la protéine P (μg), (E/P)) utilisé est compris entre 1/10 et 1/100 (Lebert et al., 2015).

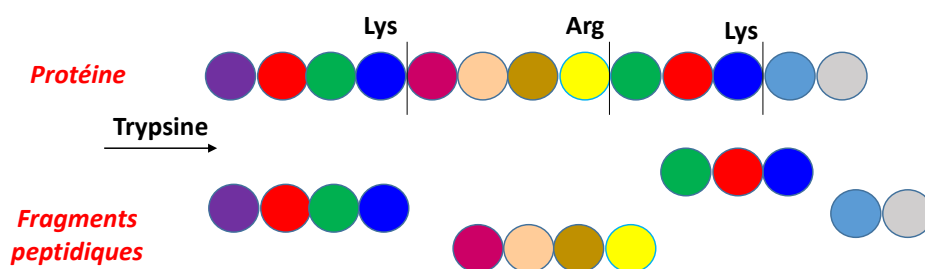


Figure 36. Digestion typique d'une protéine

Introduction bibliographique

La digestion trypsique peut être réalisée en solution ce qui implique une hydrolyse en milieu liquide, après réduction et alkylation, suivie généralement d'une analyse des peptides obtenus par LC-MS (Medzihradzky, 2005). Une seconde approche consiste en une digestion des protéines sur gel d'électrophorèse (Shevchenko et al., 2007). Après séparation électrophorétique, les spots d'intérêt sont découpés en petits cubes, puis lavés afin d'éliminer l'excédent de sels. Les protéines présentes dans chaque spot sont soumises aux étapes de réduction et d'alkylation citées ci-dessus. Les peptides obtenus sont également analysés par LC-MS.

Un des inconvénients majeurs de la digestion trypsique sur gel correspond à la perte des protéines lors des différentes opérations. Néanmoins, le fractionnement des protéines sur gel réduit la complexité du mélange protéique à analyser et simplifie ainsi l'analyse LC-MS des fragments peptidiques obtenus. À l'inverse, la digestion en solution est réalisée dans un tube et permet une récupération plus simple de la fraction liquide qui est directement analysable par LC-MS (Feist and Hummon, 2015).

1.4. Séparation des protéines

La digestion trypsique des protéines présentes, augmente de manière considérable la complexité du mélange protéique, puisqu'elle génère un grand nombre de peptides différents. Ces peptides peuvent avoir des tailles très proches, voire similaires, et être caractérisés par des spectres de masse « quasi-superposables » (Karpievitch et al., 2010). Par ailleurs, l'infusion directe d'un mélange complexe de peptides dépasse les capacités techniques des spectromètres de masse en termes de résolution et de gamme dynamique de protéines. Pour toutes raisons, différentes techniques séparatives sont employées, telles que l'électrophorèse sur gel ou encore la chromatographie en phase liquide, afin de faciliter et de simplifier l'analyse MS d'un échantillon peptidique complexe.

1.4.1. Electrophorèse bidimensionnelle (2D)

L'électrophorèse bidimensionnelle sur gel de polyacrylamide (*Two-dimensional polyacrylamide gel electrophoresis*, 2D-PAGE), représente la technique électrophorétique la plus communément utilisée pour l'analyse des protéines (Monteoliva and Albar, 2004). Une telle approche permet de séparer plusieurs milliers de protéines. Développée en 1975 (O'Farrell, 1975), cette méthode combine deux principes de séparation. Sur la première dimension, les protéines sont soumises à un champ électrique dans un gel présentant un gradient de pH continu.

Introduction bibliographique

Au cours de cette étape, appelée focalisation isoélectrique (IEF), les protéines migrent sur le gel jusqu'à une zone de pH égale à leur pH isoélectrique, où leur charge globale est nulle. La deuxième dimension permet de séparer les protéines selon leur poids moléculaire. Elle est réalisée en conditions dénaturantes en présence généralement de SDS (SDS-PAGE). Le SDS confère à la protéine une charge globale négative afin qu'elles puissent migrer dans le tamisage du gel sans contrainte spatiale, seulement en fonction de leur taille. Ainsi les protéines ayant été focalisées dans la même zone de pH isoélectrique, seront ensuite séparées suivant leur dimension, dans la seconde dimension en fonction de leur poids moléculaire (figure 37).

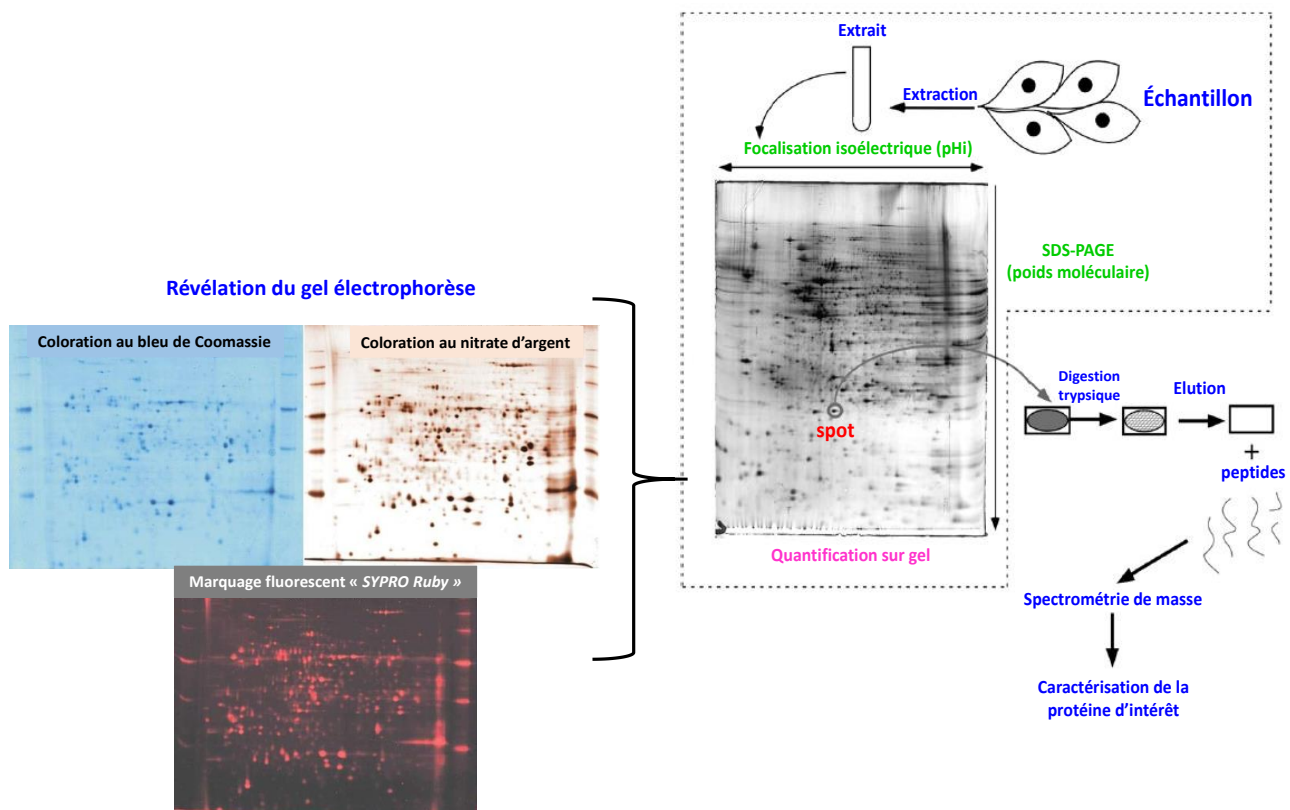


Figure 37. Schéma d'une analyse protéomique basée sur l'électrophorèse bidimensionnelle (Chevalier, 2010; Rabilloud and Lelong, 2011)

La révélation des protéines après migration est effectuée *via* des méthodes de coloration utilisant le bleu de Coomassie (Neuhoff et al., 1988), les colorations à l'argent (Chevallet et al., 2006) ou encore des sondes fluorescentes (Chevalier et al., 2006) (figure 37). Toutefois certains colorants notamment la coloration à l'argent sont incompatibles avec la spectrométrie de masse (Chevalier et al., 2006). Il existe d'autres techniques de détection plus spécifique, basées sur l'immunodétection. Les protéines révélées sur le gel d'acrylamide sont d'abord transférées sur une membrane par application d'un champ électrique. Des anticorps dits « primaires » dirigés

Introduction bibliographique

vers l'antigène (la protéine d'intérêt) sont dans un premier temps, incubés. Une fois le complexe protéine-anticorps formé, les anticorps dits « secondaires » sont ensuite incubés. Ces derniers sont le plus souvent couplés à une enzyme ou un fluorophore, qui sera détecté en UV ou par fluorescence.

Les premières applications de l'électrophorèse bidimensionnelle furent réalisées dans les années 70, où Patrick O'Farrell a mis en avant le haut pouvoir résolutif de cette technique qui avait alors permis de caractériser 1100 protéines majeures d'*Escherichia coli* (figure 38).

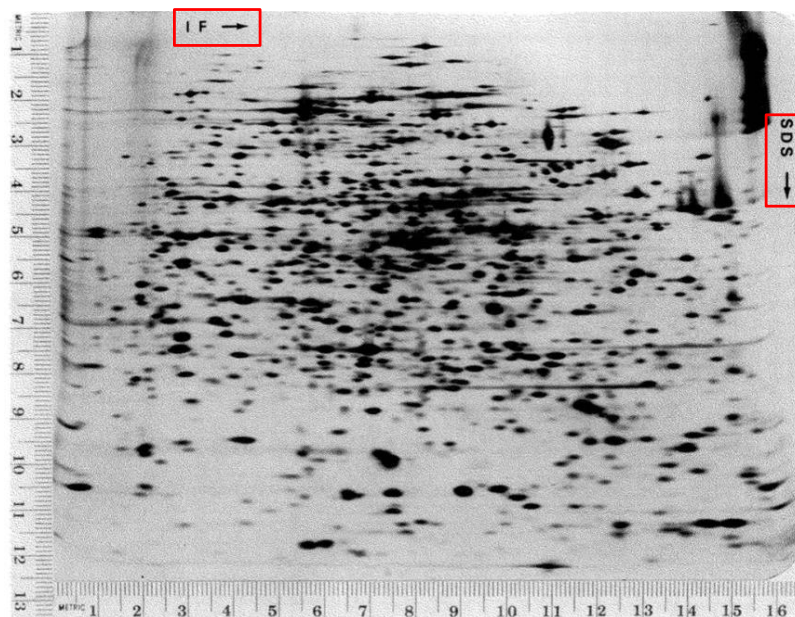


Figure 38. Séparation des protéines d'*E.coli* par électrophorèse bidimensionnelle. 1100 protéines ont été séparées. Les cellules bactériennes ont été lysées par sonication puis traitées avec de la Dnase et de la Rnase avant d'être solubilisées dans un tampon de lyse. 25 μ l d'échantillons ont été déposés sur le gel d'électrophorèse (O'Farrell, 1975)

Bien que l'électrophorèse 2D-PAGE soit une méthode reproductible et robuste (Rabilloud and Lelong, 2011), elle est néanmoins limitée par l'étape de préparation de l'échantillon, qui doit garantir une solubilisation totale des protéines et empêcher leur dégradation lors de leur migration dans le gel. De même, la présence des protéines majoritaires (dans le plasma) peut gêner la détection des protéines cibles présentes à de très faibles concentrations (Magdeldin et al., 2014a; Pragma et al., 2017).

1.4.2. Techniques séparatives « hors gel » : la chromatographie liquide

Tout comme l'électrophorèse 2D, la chromatographie liquide, est une technique d'analyse qui permet de séparer des mélanges peptidiques extrêmement complexes (des capacités de pic entre 1500 jusqu'à environ 1800, (Horie et al., 2012; Shen et al., 2005)). Couplée à la spectrométrie de masse, cette méthode est devenue un outil incontournable pour l'analyse des protéines dans des champs d'activités allant du diagnostic médical au diagnostic environnemental. Une telle méthode s'applique le plus souvent à la séparation d'un grand nombre de peptides issus de la digestion trypsique des protéines d'intérêt avant analyse par spectrométrie de masse.

Quatre types de mécanismes d'interaction sont souvent utilisés en protéomique : l'échange d'ions, l'interaction hydrophile (*Hydrophilic Interaction Liquid Chromatography*, HILIC), l'exclusion stérique ou encore l'interaction hydrophobe appelée (*Reverse Phase Liquid Chromatography*, RPLC). La figure 39, représente la sélectivité de six phases stationnaires en fonction de la polarité, de l'hydrophobicité, de la charge positive ou négative des peptides, et permet ainsi de sélectionner la phase stationnaire la plus adaptée à l'analyse du protéome dans sa globalité.

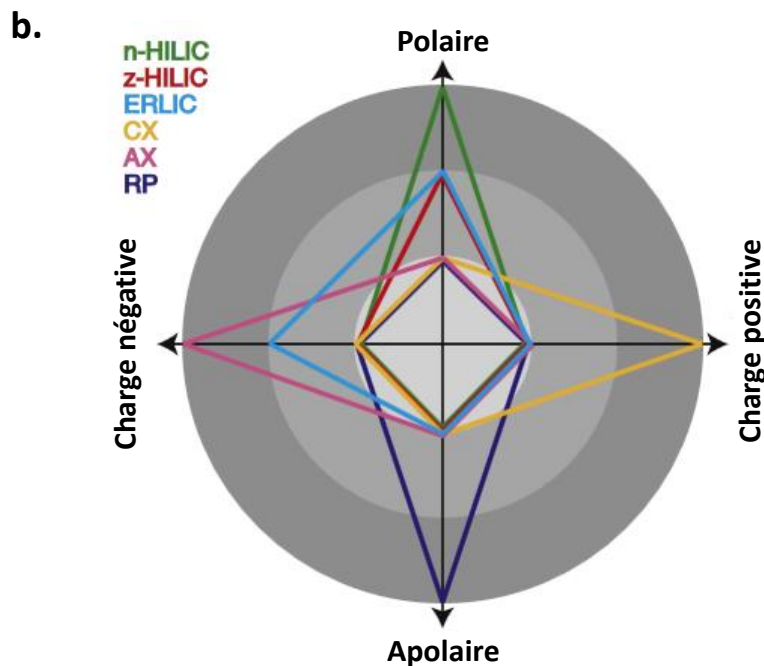
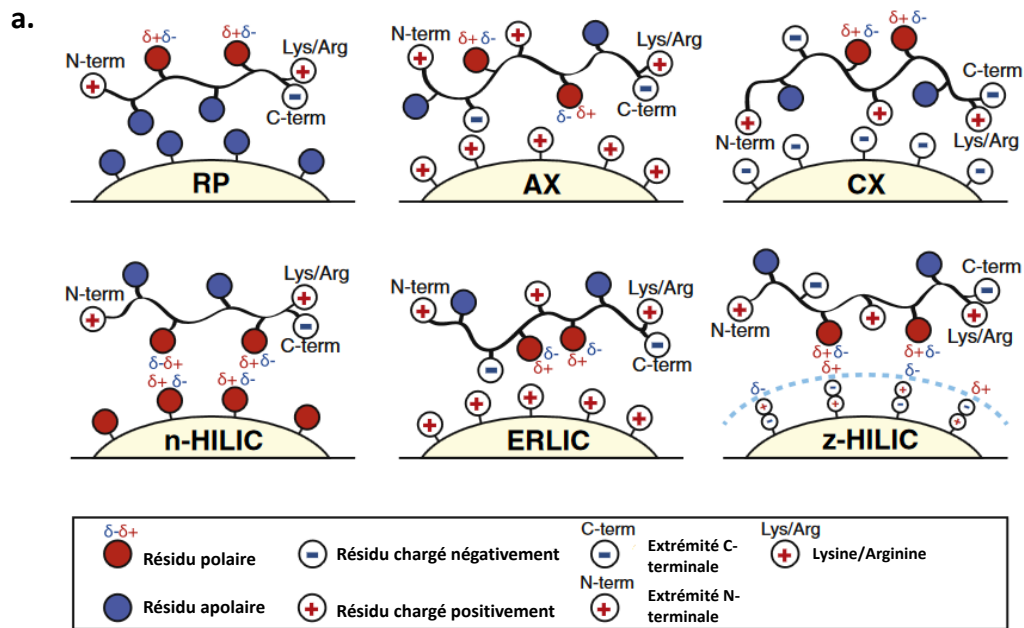


Figure 39. a. Représentation schématique des interactions établies entre les différents sites des peptides tryptiques et les phases stationnaires utilisées en chromatographie liquide : RP : phase inverse ; AX : échange d'anions ; CX : échange de cations ; n-HILIC : interaction hydrophile-neutre ; ERLIC : répulsion électrostatique-interaction hydrophile ; z-HILIC : interaction hydrophile-zwitterionique. b. Représentation schématique de la sélectivité des phases stationnaires vis-à-vis des charges positives, négatives, groupements apolaires et polaires. Le cercle interne/gris clair : faible sélectivité ; les cercles externes/gris : forte sélectivité (Di Palma et al., 2012)

Introduction bibliographique

En règle générale, la chromatographie en phase inverse (RPLC) est la méthode la plus utilisée en analyse protéomique. Introduite en 1976, pour la séparation des peptides (Gruber et al., 1976), cette technique repose sur une rétention des peptides en fonction des interactions hydrophobes qu'ils établissent avec la phase stationnaire. Les phases mobiles sont le plus souvent constituées d'un solvant aqueux et d'un solvant organique acidifié (acétonitrile ou méthanol acidifié). De nouvelles méthodes chromatographiques reposant sur des phases stationnaires en carbone graphite (colonne *Hypercarb* commercialisée par thermo Fisher Scientific), avec une structure en feuillets plans agencés de façon hexagonale sont depuis peu utilisées. Selon Jansen et al. (2009), les colonnes *Hypercarb* permettent une rétention des analytes *via* : (a) des interactions hydrophobes avec la surface apolaire du carbone graphité poreux ; (b) des interactions électrostatiques (type dipôle-charge) grâce à la présence d'électrons libres à la surface du graphite la rendant ainsi polarisable ; (c) des interactions stériques engendrées par la surface plane du carbone graphite permettant la séparation des stéréoisomères. Les colonnes *Hypercarb* permettent également de dessaler et retenir préférentiellement les peptides hydrophiles *via* des interactions polaires (Chin and Papac, 1999; Larsen et al., 2004; Minshull et al., 2016).

Des colonnes dites monolithiques sont également utilisées. Elles sont caractérisées par un double réseau poreux formé de macropores et de mésopores (figure 40). Ceux-ci assurent respectivement une réduction de la pression en tête de colonne et une augmentation de la surface spécifique de la phase stationnaire permettant ainsi une amélioration de l'efficacité de séparation (Jiang et al., 2015). Les colonnes monolithiques peuvent être utilisées comme réacteurs enzymatiques pour la digestion en ligne des protéines (immobilisation de la trypsine sur support monolithique) (Jiang et al., 2015; Šlechtová et al., 2017).

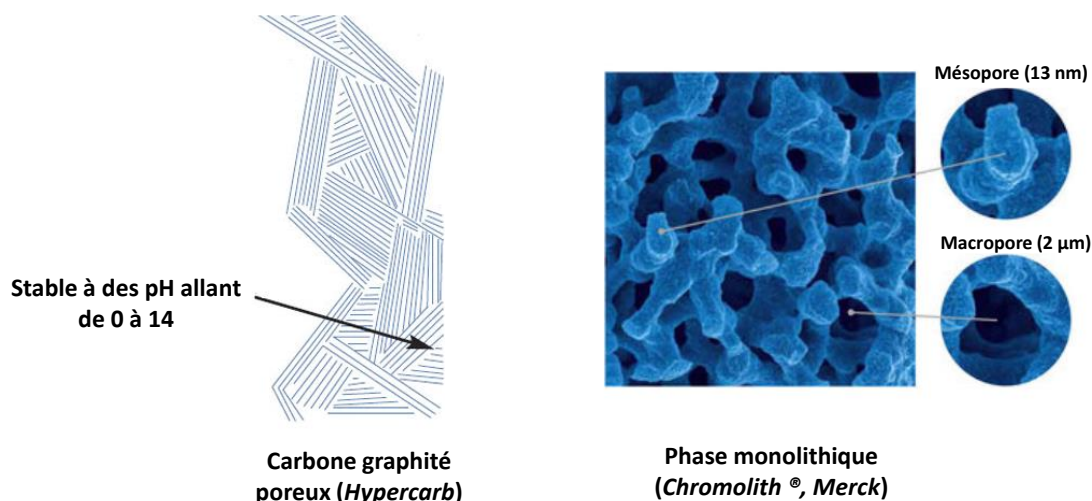


Figure 40. Représentation d'une phase en carbone graphité poreux (*Hypercarb*) poreux et d'une phase monolithique « *Chromolith* ® » commercialisée par Merck

1.4.2.1. La chromatographie Liquide à Très Haute Pression (UHPLC)

Plusieurs éléments sont à prendre en compte dans le choix de la colonne analytique : la longueur, le diamètre interne de la colonne (d.i), la viscosité du solvant, le débit de la phase mobile, la taille des particules, la pression ou encore la température de la colonne. En effet, une diminution de la taille des particules des colonnes de 3,5 à 1,7 μm , entraîne une augmentation de l'efficacité de séparation du même facteur. Ceci conduit par conséquent à une augmentation de la résolution d'environ 1,4 (Wilson et al., 2005). Une élévation de la pression d'environ 8 fois est observée. Grâce à l'avènement des techniques de séparation, les colonnes chromatographiques type UHPLC (Chromatographie Liquide à Très Haute Pression, *Ultra High Pressure Liquid Chromatography*), furent développées pour des applications en protéomique. Ces colonnes analytiques permettent d'augmenter de manière considérable le pouvoir résolutif, d'améliorer la sensibilité et de diminuer la durée de l'analyse à quelques minutes par rapport à une analyse par HPLC, qui utilise des colonnes ayant des tailles de particules plus importantes (Zhao and Lin, 2014). Les caractéristiques des colonnes HPLC et UPLC sont résumées dans le tableau 3.

1.4.2.2. La nanochromatographie

Les limites posées par l'analyse des protéines dans les milieux complexes nécessitent à l'heure actuelle toujours plus de sensibilité qui a depuis peu été considérablement améliorée avec le développement de la nanochromatographie liquide (Nano-LC) (Pottiez et al., 2017;

Introduction bibliographique

Benndorf et al., 2007). Les colonnes utilisées en nano-LC sont caractérisées par des faibles diamètres internes inférieurs ou égaux à 0,1 mm (Wilson et al., 2015). La réduction du diamètre interne des colonnes permet d'obtenir un gain important en sensibilité et en intensité de signal, à de très faibles débits d'élution ($\text{nL}\cdot\text{min}^{-1}$). Néanmoins, la nano-LC nécessite généralement des étapes de purification préalables afin d'obtenir un mélange peptidique extrêmement pur (dépourvu de sels, interférents, etc.), et éviter ainsi les phénomènes de colmatage de la colonne (Uchida et al., 2013). En outre, l'injection de faibles volumes d'échantillon (nL) limite la détection des cibles. Par ailleurs, les systèmes de nano-LC ne bénéficient pas toujours d'une bonne robustesse en raison de la fragilité des nano-composants (Gama et al., 2013; Sinnaeve and Van Bocxlaer, 2004). Une comparaison entre les différents systèmes chromatographiques est présentée sur le tableau 3.

Tableau 3. Comparaison des caractéristiques techniques de l'UPLC, de la NanoLC par rapport à l'HPLC (Gumustas et al., 2013; Wilson et al., 2015)

	HPLC	UPLC	NanoLC
Pression maximale (Bar)	600	2000	1200 ou plus
Débit ($\mu\text{L}\cdot\text{min}^{-1}$)	1000-2000	10-500	0,1-0,3
Volume injecté (μL)	5-100	2-20	0,01
Diamètre de la colonne (d.i) (mm)	2-4,6	1-2,2	$\leq 0,1$
Taille des particules	3-5	1,4-2	< 2
Longueur de la colonne (nm)	150-300	50-200	50-150

1.4.2.3. La Chromatographie multidimensionnelle (MDLC : Multidimensional Liquid Chromatography)

La très grande complexité des échantillons biologiques et environnementaux nécessite à l'heure actuelle, la mise en œuvre d'approches chromatographiques plus performantes afin de réduire de manière significative la complexité des mélanges peptidiques et ainsi améliorer la séparation de ces composés (Valeja et al., 2015). La chromatographie multidimensionnelle est actuellement en plein essor et largement appliquée pour l'analyse de divers composés

Introduction bibliographique

(métabolites polaires/apolaires, peptides, etc.) (Wang et al., 2015; Zhou et al., 2013; Stephan et al., 2016; Wang et al., 2017). Cette technique consiste à coupler des systèmes chromatographiques impliquant des mécanismes de rétention différents. L'objectif est d'augmenter la capacité de pics (nombre de pics pouvant être séparés dans une fenêtre de temps avec une résolution donnée) et d'améliorer la résolution des peptides avant une analyse par spectrométrie de masse (Law et al., 2015; Vonk et al., 2015; Zhang et al., 2010). Lors d'une séparation bidimensionnelle, chaque fraction de l'échantillon subit deux modes de séparation différents (critère d'orthogonalité) (Wu et al., 2012). Plusieurs modes de couplage sont utilisés pour la séparation des peptides. Une configuration bidimensionnelle (2D) comprenant une colonne d'échange cationique et une colonne à polarité inverse est le plus souvent employée pour la séparation des peptides (figure 41) (Gong et al., 2008; Mitulović, 2015; Qian and Hettich, 2017). Une séparation bidimensionnelle (SCX + RPLC) d'un mélange de peptides complexes a permis d'augmenter la capacité de pics, d'environ 10 fois (Frahm et al., 2006), comparée à un système classique (RPLC).

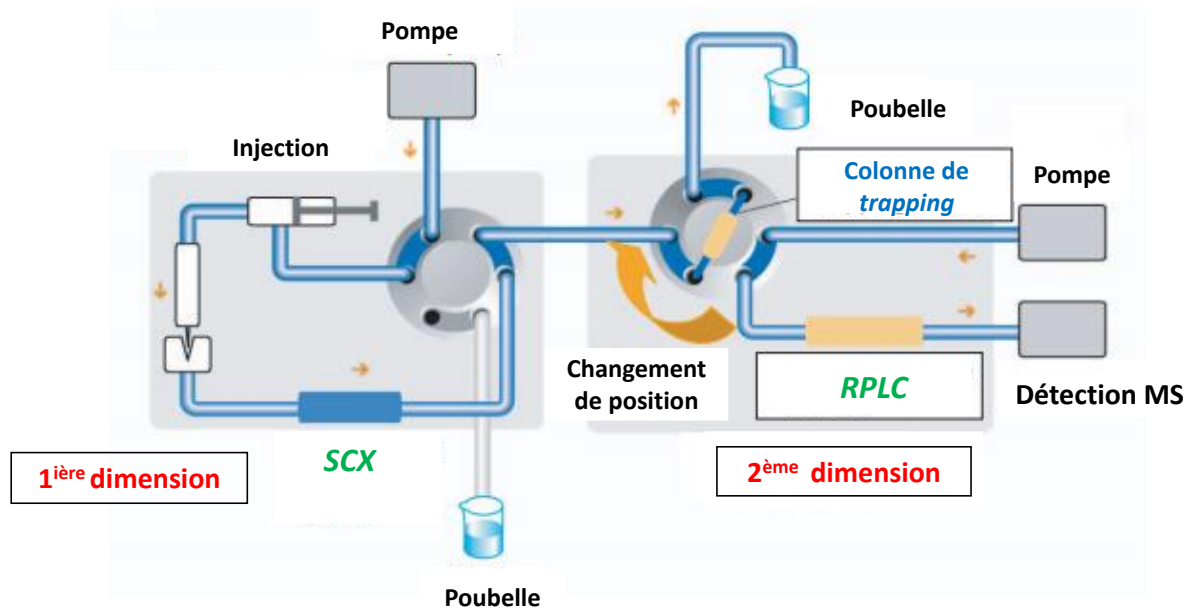


Figure 41. Principe d'une chromatographie liquide 2D (Nägele et al., 2004)

Dans les études protéomiques non ciblées, la chromatographie 2D comprehensive est utilisée. Dans ce cas, l'ensemble de l'échantillon est séparé suivant la première dimension puis transféré dans la seconde dimension. À l'inverse, dans les études de protéomiques ciblées, l'approche dite de « *heart-cutting* » est employée. Une ou plusieurs fractions de l'échantillon

Introduction bibliographique

sont éluées par la première dimension puis transférées dans la deuxième colonne (Sandra et al., 2017; Sarrut et al., 2014).

Il existe d'autres systèmes multidimensionnels permettant la préconcentration puis la séparation de mélanges peptidiques. À titre d'exemple, Wang et al. (2010) ont mis en place un couplage en ligne d'une colonne de préconcentration biphasique (RP C₁₈-SCX monolithique) et d'une séparation analytique sur une phase inverse C₁₈. Environ 1000 protéines (étude du carcinome hépatocellulaire) ont été quantifiées grâce à ce système multidimensionnel (Wang et al., 2010) (figure 42).

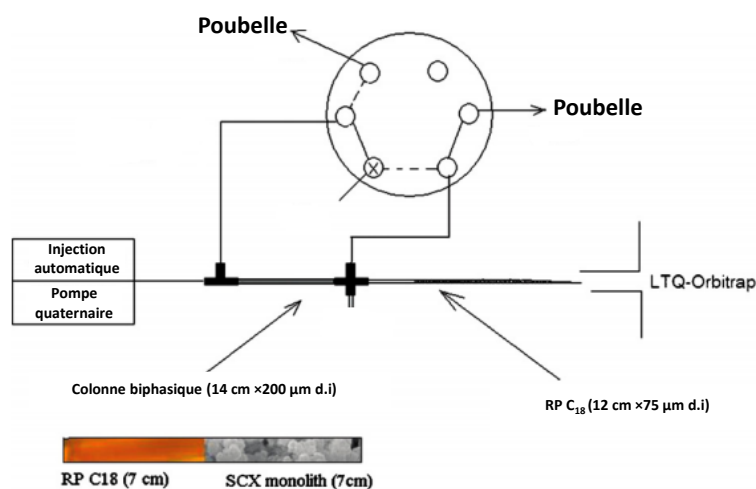


Figure 42. Représentation schématique du système multidimensionnel développé par Wang et al. (2010)

Néanmoins, la mise en place d'un système de chromatographie multidimensionnel fait face à de nombreuses difficultés liées notamment à des problèmes de miscibilité entre les phases mobiles lors d'un couplage entre une phase échangeuse d'ions et une phase inverse. En outre, l'interface entre les deux systèmes (vannes, joints de raccordements entre les colonnes, etc.) est susceptible d'augmenter le volume mort, qui lorsqu'il est associé à des injections de faibles volumes (de l'ordre du nL) peut conduire à une diminution de la sensibilité de l'analyse (Koppen et al., 2014). Le tableau 4 ci-dessous présente quelques exemples de séparation des peptides par chromatographie liquide dans le domaine de l'environnement et de la santé.

Introduction bibliographique

Protéome	Phase stationnaire	Phase mobile	Références
Protéome du milieu marin	Phase inverse silice greffée C ₁₈ (12 cm×75 mm d.i, 5 µm, 100 Å)	Eau (0,1 % acide formique) ; acétonitrile (0,1 % acide formique)	(Morris et al., 2010)
Protéome du sol	Phase inverse C ₁₈ (25 cm×75 mm d.i, 5 µm)	Eau (0,1 % acide formique) ; acétonitrile (0,1 % acide formique)	(Zampieri et al., 2016)
Protéome des eaux douces	-1 ^{ère} dimension : <i>Hypercarb</i> (colonne en carbone graphite poreux) (50 ×2.1 mm, 3 mm, 5 mm). -2 ^{ème} dimension : Phase inverse C ₁₈ (75 µm×105 mm, 2 µm, 100 Å).	<i>Hypercarb</i> : 3 % acétonitrile (0,1 % acide trifluoroacétique) ; 97 % acétonitrile (0,1 % acide trifluoroacétique). C ₁₈ : 3 % acétonitrile (0,1 % acide formique) ; 97 % acétonitrile (0,1 % acide formique).	(Russo et al., 2016)
Peptides amyloïdes (Alzheimer)	Phase inverse C ₁₈ (50 cm×200 µm d.i, 3 µm, 300 Å).	Formiate d'ammonium dans l'eau (10 mM). Formiate d'ammonium dans l'acétonitrile (90 %).	(Pottiez et al., 2017)
Protéome du liquide céphalorachidien	-1 ^{ère} dimension : Echange cationique (SCX) (Polysulfo-éthyle, 200 mm×2.1 mm). -2 ^{ème} dimension : Phase inverse C ₁₈ (15 cm × 75 µm).	SCX : 10 mM KH ₂ PO ₄ et 25% acetonitrile (pH < 3). 10 mM KH ₂ PO ₄ , 1 M KCl, et 25% acétonitrile (pH < 3). C ₁₈ : 95 % acétonitrile (0,1 % acide formique).	(Waller et al., 2008)
Protéome du foie de rat	-1 ^{ère} dimension : Echange cationique (SCX) (150 mm × 2.1 mm, 5 µm, 300 Å). -2 ^{ème} dimension : Phase inverse C ₁₈ (100 mm × 2.1 mm, 1,7 µm, 100 Å).	SCX : 10 mM dihydrogénophosphate de potassium dans 25 % acétonitrile (pH 2,75); 1 M chlorure de potassium (pH 2,75). C ₁₈ : Eau (0,1 % acide formique) ; acétonitrile (0,1 % acide formique)	(Golizeh et al., 2015)

Tableau 4. Exemples de séparation des peptides par chromatographie liquide (1D et 2D)

Introduction bibliographique

- Les caractéristiques des colonnes sont données dans l'ordre suivant (longueur de la colonne × diamètre interne, granulométrie, porosité)

2. Analyse des protéines par spectrométrie de masse

2.1. Principe et description

La spectrométrie de masse permet de détecter, d'identifier et de quantifier des molécules ionisées et/ou leurs produits de fragmentation en fonction de leur rapport masse sur charge (m/z) (Gillet et al., 2016). Couplée le plus souvent à une technique séparative comme la chromatographie liquide (LC), l'électrophorèse bidimensionnelle ou encore la chromatographie en phase gazeuse (GC), la spectrométrie de masse représente le dernier maillon de la chaîne analytique. Une telle technique permet une analyse et une quantification sensible de l'ensemble des protéines exprimées à l'état de traces dans des systèmes complexes. La spectrométrie de masse permet aussi l'identification des modifications post-traductionnelles d'acides aminés et la localisation des sites modifiés au sein d'une protéine donnée (Aebersold and Mann, 2016). Un spectromètre de masse est constitué d'une source d'ionisation et d'un ou plusieurs analyseurs qui vont permettre de séparer, sélectionner et fragmenter les ions en fonction de leur rapport m/z , et d'un détecteur qui collecte les ions séparés. (Di Girolamo et al., 2013) (figure 43).

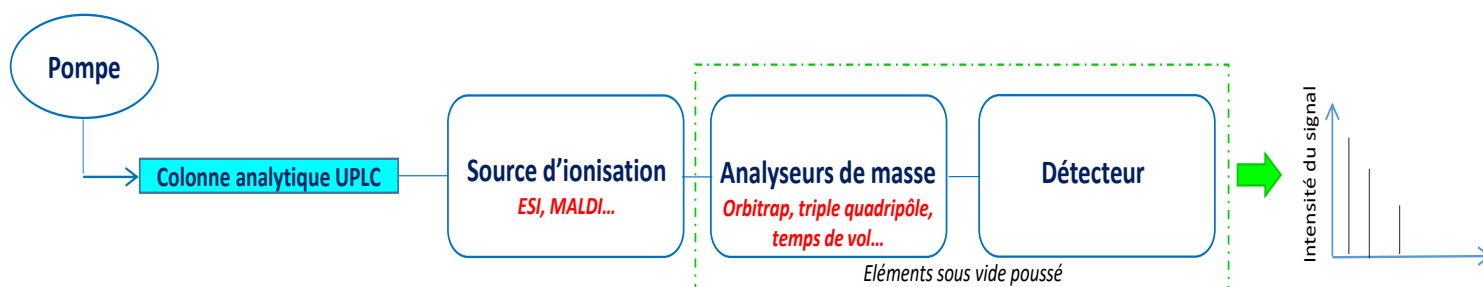


Figure 43. Représentation générale d'un spectromètre de masse

2.2. Les sources d'ionisation

La source d'ionisation consiste à ioniser des analytes contenus dans un échantillon. En protéomique des techniques d'ionisation douces « *soft ionization* » sont utilisées pour l'analyse des macromolécules biologiques, contrairement aux techniques d'ionisation par impact électronique (EI) qui conduisent à une fragmentation de la molécule au cours de l'ionisation. Il existe de nombreuses sources d'ionisation dont le choix va dépendre principalement de la nature physicochimique des molécules d'intérêts. Les mécanismes d'ionisation les plus utilisées en protéomique sont : l'ionisation par électrospray ou électronebulisation (ESI), l'ionisation chimique à pression atmosphérique (APCI), la désorption ionisation laser assistée par matrice

Introduction bibliographique

(MALDI). De nouvelles technologies d'ionisation sont également employées pour l'analyse des protéines, telles que la désorption ionisation par électronébulisation (DESI) (Douglass and Venter, 2013) ou la technique LESA (*Liquid extraction surface analysis*) (Mikhailov et al., 2016; Sarsby et al., 2015).

2.2.1. Ionisation électrospray (ESI)

L'ionisation par électrospray ou électronébulisation a été développée en 1984 par le groupe Fenn (Yamashita and Fenn, 1984), qui l'a ensuite appliquée à l'identification et l'analyse des macromolécules biologiques. En 2002, ce travail de recherche a été récompensé par le prix Nobel de chimie. L'un des principaux avantages de ce mode ionisation réside dans la formation d'ions multichargés, permettant ainsi l'analyse de molécules de hauts poids moléculaires. Par ailleurs, le processus d'électronébulisation permet de mettre en œuvre une méthode d'ionisation douce afin d'éviter une éventuelle fragmentation moléculaire (Wilm, 2011). En ESI, le transfert des ions en phase gazeuse, implique trois principales étapes (Bruins, 1998; Wilm, 2011) ; (1) le processus nébulisation ou création de microgouttelettes chargées. Les molécules en solution pénètrent dans la source *via* un capillaire métallique. Sous l'action d'un gaz inerte appelé aussi gaz de nébulisation (N₂ en général), un champ électrique (potentiel entre 2,5 et 6 kV) est établi entre le capillaire et une électrode. Ce champ électrique intense charge les gouttelettes de la solution qui s'accumulent au niveau de l'extrémité du capillaire. Il s'en suit, la formation d'un cône nommé cône de Taylor résultant de la déformation du solvant. Ce cône s'étire entraînant une émission d'un jet de gouttelettes chargées. Les gouttelettes formées se déplacent sous l'action d'un champ électrique, le solvant contenu dans celles-ci est évaporé par collision avec un gaz inerte (N₂). Au fur et à mesure de cette désolvatation, la taille des gouttelettes diminue, augmentant ainsi la répulsion électrostatique entre charges de même signe. Lorsque cette répulsion électrostatique devient supérieure aux forces de tension superficielle (limite de Rayleigh), la gouttelette se scinde en gouttelettes de plus petite taille. Ce processus se produit jusqu'à conduire à des ions totalement désolvatés (figure 44).

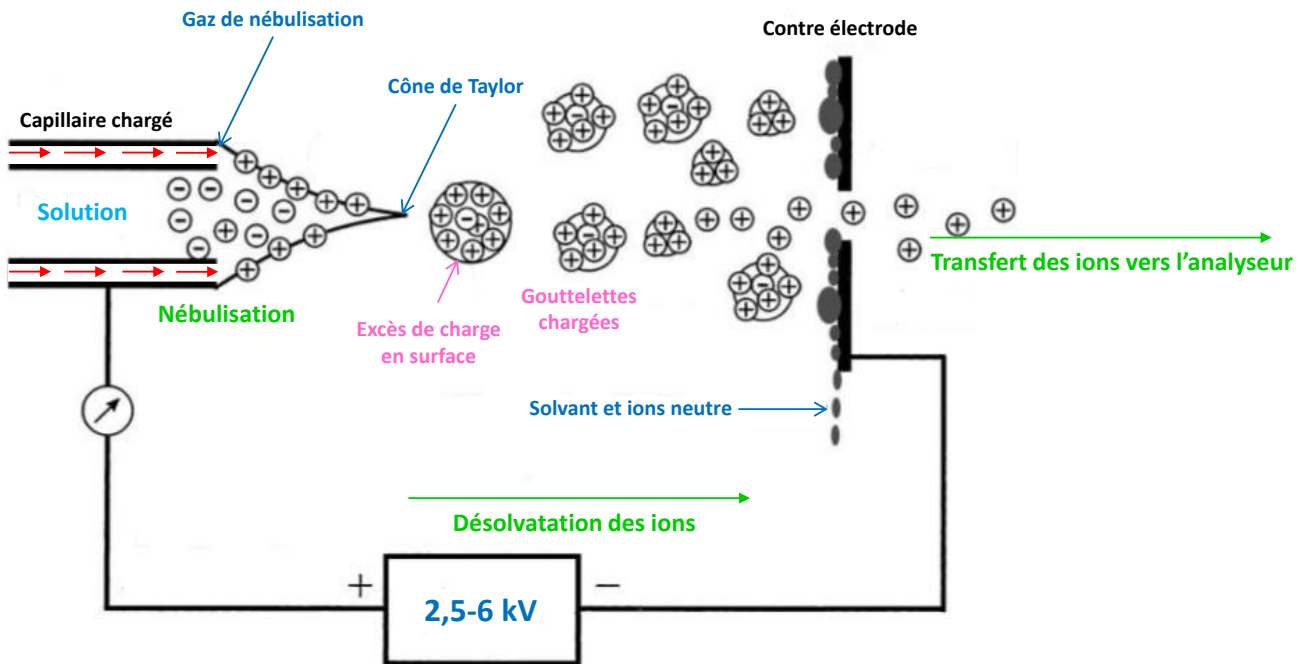


Figure 44. Représentation schématique de la source à l'électrospray en mode positif (Cech and Enke, 2001)

Le mécanisme d'ionisation-désolvatation de l'ESI peut être expliqué par deux théories. La première propose un modèle de l'évaporation ionique appelé aussi « *Ion evaporation model* ». Tandis que la deuxième théorie implique un modèle de la charge résiduelle ou « *Charged residue model* ».

Le premier modèle représenté dans la figure 45.a, prend en compte les propriétés géométriques des gouttelettes. De ce fait, lorsque les gouttelettes chargées atteignent un rayon suffisamment petit de l'ordre de 20 nm après évaporation du solvant, la densité de charge de surface est suffisamment élevée pour que l'ion soit expulsé en phase gazeuse. Le modèle de la charge résiduelle (figure 45.b), implique quant à lui une succession de fissions des gouttelettes suite à l'évaporation du solvant et une série d'explosions coulombiennes, jusqu'à l'obtention d'une dernière gouttelette ne contenant qu'un seul ion (Wilm, 2011).

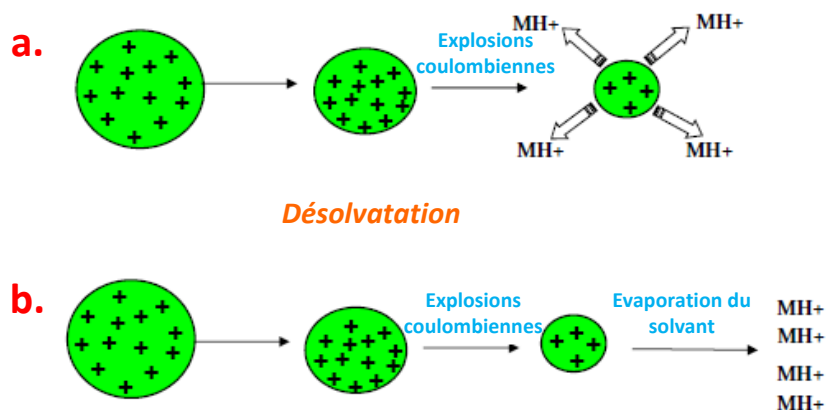


Figure 45. Représentation schématique du mécanisme d'ionisation-désolvatation ; a. modèle de l'évaporation ionique, b. modèle de la charge résiduelle (Wilm, 2011).

L'ionisation par électrospray est devenue aujourd'hui, la méthode de référence pour l'analyse des molécules de hauts poids moléculaires, peptides et protéines. Par ailleurs, la génération d'une interface entre la phase liquide et la phase gazeuse rend la méthode d'ionisation-désolvatation parfaitement adaptée pour son couplage avec la chromatographie liquide.

L'ionisation par électrospray des peptides en mode positif ou négatif permet de générer, respectivement, des ions multichargés, $[M+nH]^{n+}$ ou $[M-nH]^{n-}$. En outre, les peptides générés après digestion trypsique contiennent des sites basiques (résidus aminoacides basiques (lysine, arginine et histidine) ou fonction N-terminale), ce qui facilite leur ionisation en mode positif. Ainsi, les phases mobiles utilisées en électrospray sont acidifiées par des acides organiques volatils tels que l'acide acétique ou l'acide formique afin de favoriser la protonation (Johnson et al., 2013b) (Exemple de spectre de masse d'un peptide obtenu après digestion trypsique, figure 46). Les liquides injectés dans la source d'électrospray doivent être soigneusement contrôlés. L'injection de détergents, habituellement utilisés pendant l'étape de préparation de l'échantillon en protéomique, comme le SDS ou le CHAPS, aboutit le plus souvent à une suppression du signal des ions caractéristiques des peptides à analyser (Scheerlinck et al., 2015). Pour cette raison, il est le plus souvent conseillé d'inclure une étape de lavage (élimination de l'excès de détergents), utilisant par exemple, une cartouche utilisant un filtre à membrane en polytétrafluoroéthylène (Crowell et al., 2015).

Introduction bibliographique

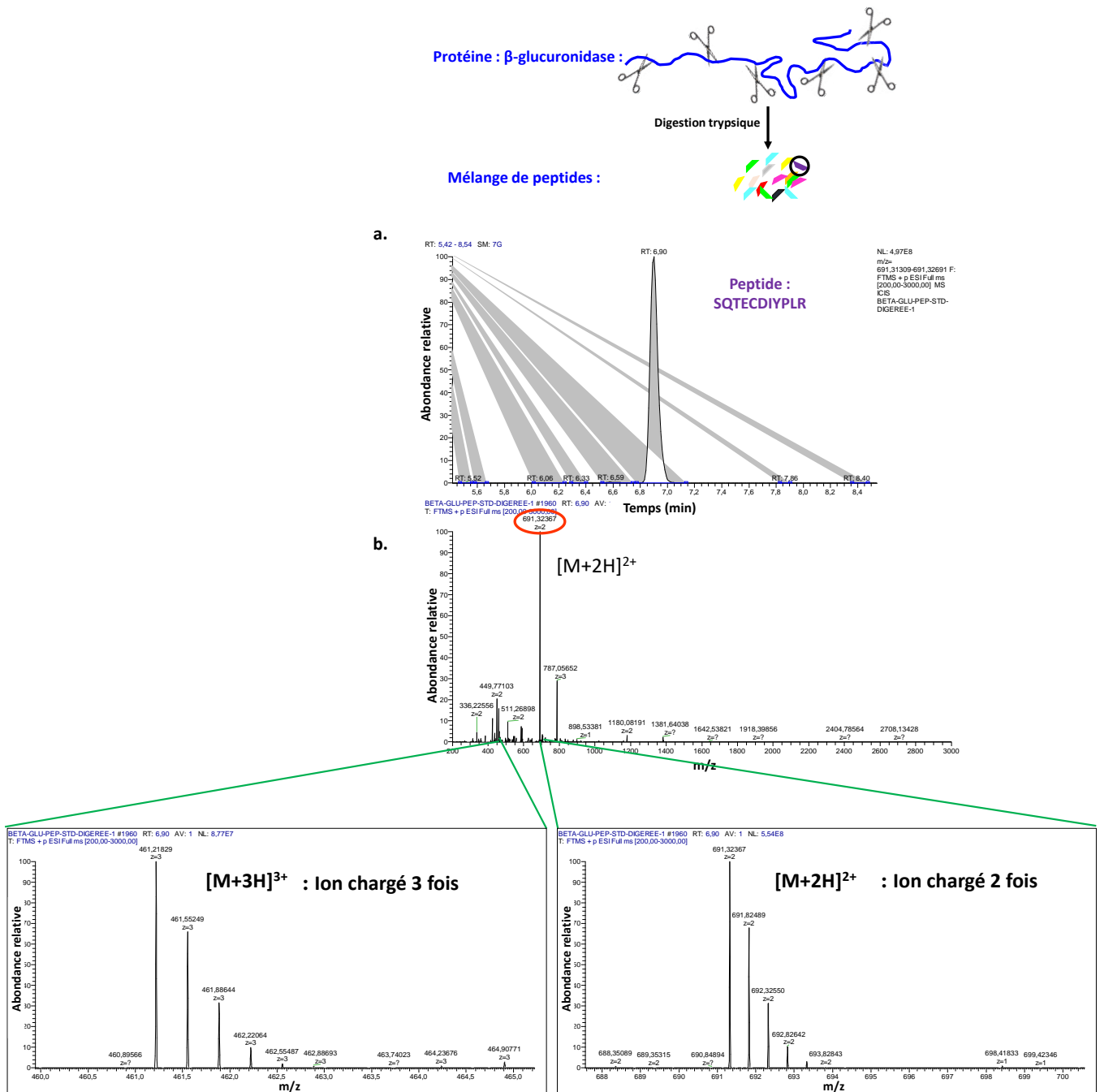


Figure 46. a. Chromatogramme du peptide ‘SQTECDIYPLR’ issu de la digestion trypsique de la β -glucuronidase; b. spectre de masse (MS) du peptide $[M+2H]^{2+}$, $m/z=691,32367$ et du $[M+3H]^{3+}$, $m/z=461,21829$

Il existe d’autres méthodes d’ionisation plus récentes telles que la désorption-ionisation par électronébulisation (DESI). Tout comme l’ESI, la DESI est une technique d’ionisation douce permettant de détecter des molécules intactes. Ainsi, un spray de très fines gouttelettes, produits par électronébulisation, est envoyé à pression atmosphérique directement sur la surface de l’échantillon à analyser (exemple tissus biologiques). L’impact du spray sur la surface permet

Introduction bibliographique

d'extraire les composés qui seront ensuite ionisés et acheminés jusqu'au spectromètre de masse. Un des avantages de cette technique d'ionisation est qu'elle ne nécessite aucune étape de préparation d'échantillon. Cette méthode est depuis peu utilisée pour faire de l'imagerie sur tissus dans des conditions ambiantes (Douglass and Venter, 2013) (figure 47).

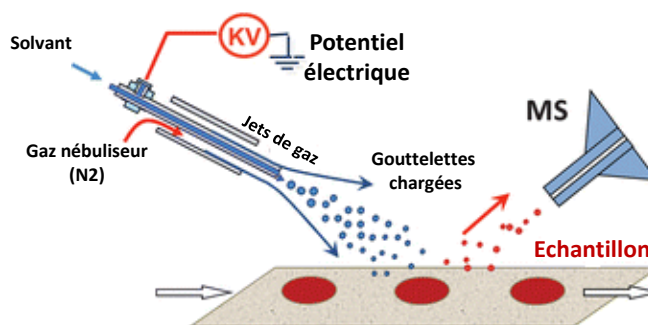


Figure 47. Représentation schématique du principe de la DESI (Ifa et al., 2010)

Dans le cadre de notre étude, nous avons choisi d'utiliser l'ionisation par électrospray (ESI) pour l'ionisation des peptides d'intérêt.

2.2.2. Ionisation désorption laser assistée par matrice (Matrix Assisted Laser Desorption Ionisation, MALDI)

La source MALDI est généralement utilisée pour l'analyse des protéines intactes (approche « *top-down* »). Le développement de la technique d'ionisation MALDI remonte aux années 80, où un groupe de chercheurs a pu démontrer son intérêt pour l'analyse de molécules de hauts poids moléculaires jusqu'à plusieurs milliers de daltons (Karas and Hillenkamp, 1988). La source MALDI consiste à évaporer et à ioniser les protéines natives (non hydrolysées), déposées sur une matrice, à l'aide d'un faisceau laser. La matrice (Acide α -cyano-4-hydroxycinnamique (CHCA), Acide 2,5-dihydroxybenzoïque (DHB)) permet de protéger les analytes d'une possible dégradation au contact du laser et de favoriser l'ionisation. L'ionisation se déroule en plusieurs étapes : (1) la première étape consiste à mélanger la matrice en excès avec la solution protéique. Ce mélange est par la suite déposé sur une plaque métallique. (2) Dans une deuxième étape, la plaque est placée sous vide et un faisceau laser pulsé, travaillant le plus souvent en UV (environ 337 nm), est envoyé sur le mélange matrice/protéines co-cristallisé. (3) L'énergie transmise par le faisceau laser est absorbée provoquant l'ionisation et la sublimation du mélange

Introduction bibliographique

de la matrice, entraînant avec elle les analytes d'intérêt (Di Girolamo et al., 2013). La figure 48 représente le principe de l'ionisation MALDI.

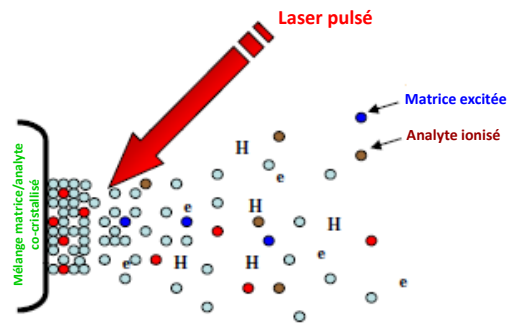


Figure 48. Principe de l'ionisation MALDI (Di Girolamo et al., 2013)

En ionisation MALDI, les ions générés sont généralement monochargés (des ions multichargés peuvent être également observés). Couplée à un analyseur de masse temps de vol (TOF), cette technique est principalement utilisée pour l'ionisation des protéines intactes présentes dans un échantillon donné.

Parmi les principales découvertes liées à l'utilisation du MALDI-TOF, nous pourrions citer la détection de l'insuline, facteur de croissance (Nelson et al., 2004), ou encore l'entérotoxine B staphylococcique (toxine responsable des intoxications alimentaires) dans les matrices biologiques, à des quantités de l'ordre du nanogramme (Schlosser et al., 2007) (figure 49).

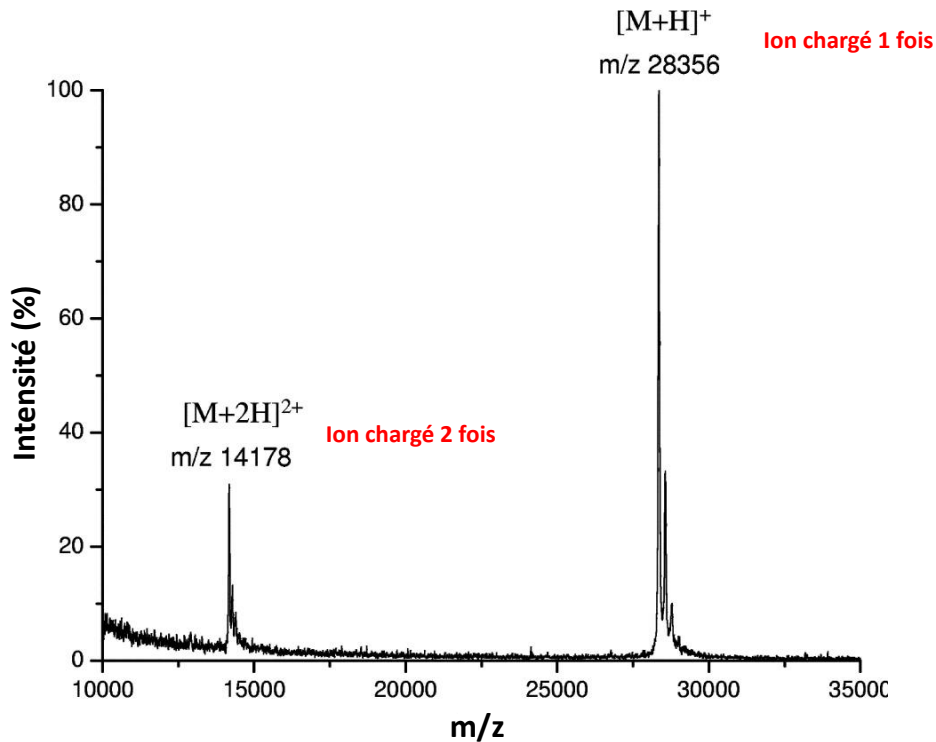


Figure 49. Spectre de masse de la protéine (entérotoxine B staphylococcique) obtenu par MALDI-TOF/MS (85 fmol ou 2,5 ng). Les pics correspondent à l'ion moléculaire de la protéine mono et di-chargée (Schlosser et al., 2007)

Une variante proche de la technologie MALDI, est la méthode SELDI (*Surface Enhanced Laser Desorption Ionization*) couplée à l'analyseur temps de vol (SELDI-TOF). Ce type d'approche, consiste à déposer l'échantillon sur une surface (puce à protéine ou *ProteinChip Array*), activée chimiquement et capable ainsi d'interagir de manière spécifique avec la molécule cible *via* des interactions, type antigène-anticorps, enzyme substrat, etc. (Di Girolamo et al., 2013) (figure 50). Une fois générés, les ions traversent l'interface entre la source et l'analyseur de masse. Cette interface assure la transmission des ions de la pression atmosphérique vers une région où règne un vide poussé. Elle permet aussi de focaliser le faisceau d'ions et de garantir une désolvatation complète des ions avant qu'ils n'atteignent l'analyseur de masse. L'intérêt majeur de cette méthodologie réside dans sa rapidité et sa spécificité pour l'établissement de profils d'expression protéique d'un échantillon à un autre et l'identification de nouveaux biomarqueurs potentiels (Muthu et al., 2016).

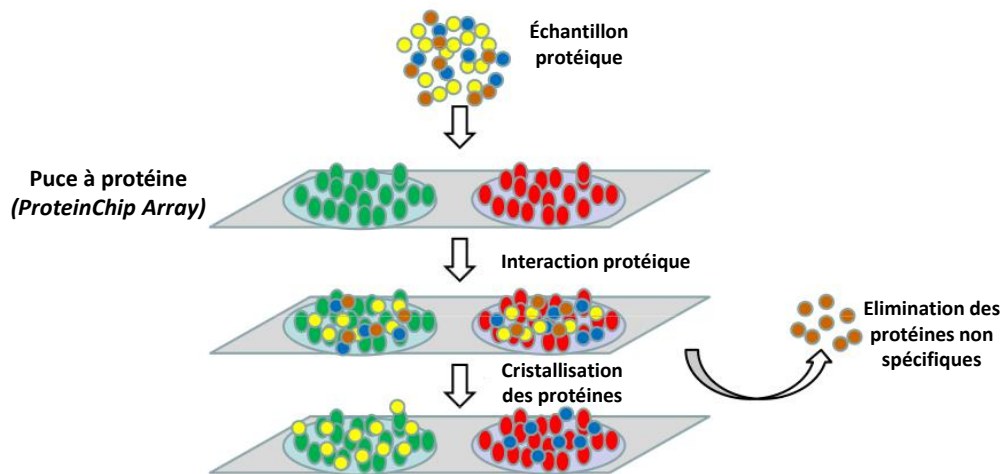


Figure 50. Représentation schématique du processus du SELDI (Di Girolamo et al., 2013)

2.3. Les analyseurs de masses

L'analyseur de masse a pour rôle de séparer les ions formés dans la source en fonction de leur rapport m/z . Plusieurs types d'analyseurs, couplés à une source ESI ou MALDI, sont utilisés en protéomique (Temps de vol (TOF), triple quadripôle, orbitrap, résonance cyclotronique ionique à transformée de Fourier (FT-ICR)). Le choix est réalisé en fonction de l'application qualitative ou quantitative. Plusieurs paramètres sont aujourd'hui utilisés pour juger de la performance des analyseurs, tels que le pouvoir de résolution, la sensibilité, la précision de mesure, la gamme de masse détectable.

Le pouvoir de résolution d'un analyseur R est défini comme la capacité à discriminer deux ions ayant un rapport masse sur charge voisin. Il correspond au rapport entre la masse m et la largeur du pic à mi-hauteur, Δm : $R = m/\Delta m$. Généralement la résolution est calculée en prenant la largeur du pic à mi-hauteur (*Full Width at Half Maximum*, FWHM) (figure 51).

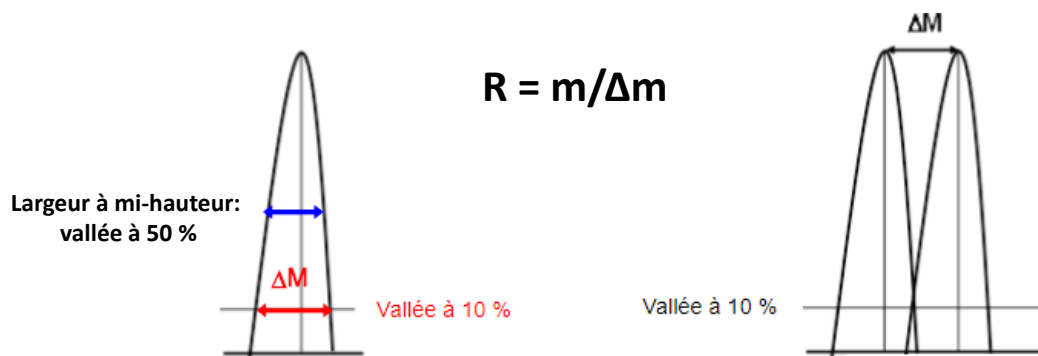


Figure 51. Représentation schématique de la résolution d'un analyseur de masse

Introduction bibliographique

La précision sur la mesure du m/z , permet d'évaluer l'erreur entre la masse théorique et la masse donnée par l'appareil. Elle le plus souvent exprimée en partie par million (ppm). La gamme de masse analysable doit être prise en compte en fonction des composés recherchés. Les analyseurs de masse sont aussi caractérisés par leur sensibilité, c'est-à-dire par la quantité minimale d'analyte détectable par l'appareil. D'autres critères peuvent être aussi pris en compte, notamment la vitesse de balayage, ou encore la possibilité de réaliser des analyses en tandem.

Les analyseurs en tandem, les plus utilisés en protéomique, seront principalement présentés dans la partie suivante.

2.3.1. Les analyseurs en tandem

La spectrométrie en tandem (MS/MS) consiste à coupler plusieurs analyseurs entre eux afin d'obtenir des appareils dits hybrides. L'utilisation de la (MS/MS) permet de mesurer la masse de l'ion d'intérêt appelé aussi ion précurseur ou encore ion parent, mais également la masse des ions fils, obtenus après fragmentation de cet ion parent. Une telle analyse se déroule en trois étapes : (1) un premier analyseur sélectionne un ou plusieurs ions précurseurs (parents) avec un certain rapport m/z , (2) l'ion précurseur est ensuite fragmenté dans une cellule de collision, (3) enfin, un deuxième analyseur consiste à mesurer les rapports m/z des ions fragments (fils) (figure 52). L'intérêt de la spectrométrie de masse en tandem réside dans la possibilité de confirmer l'identité de l'ion parent par l'intermédiaire des ions fils, de le quantifier de façon précise et d'augmenter ainsi la sensibilité de détection.

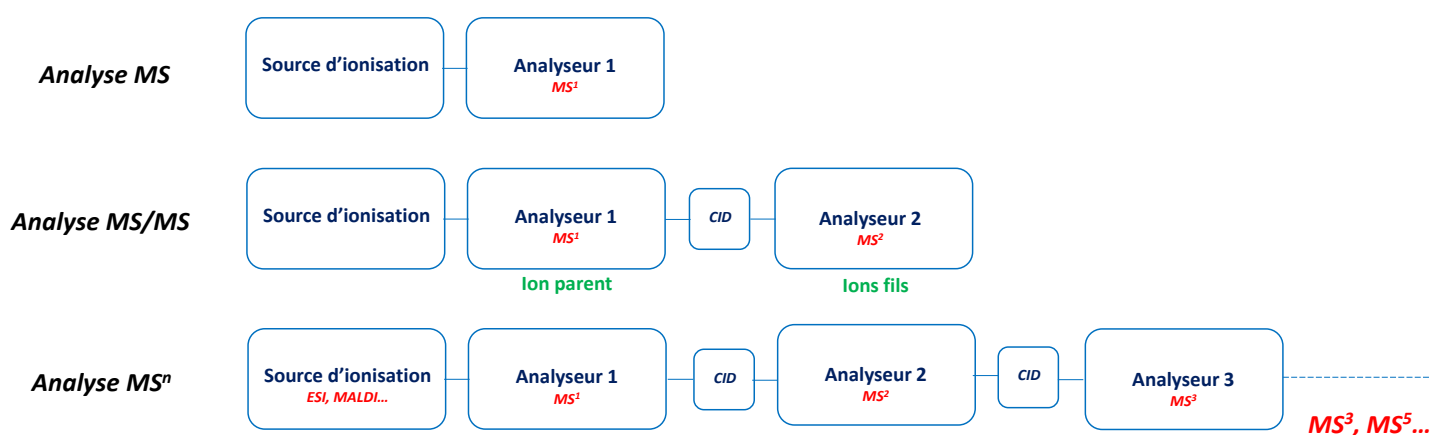


Figure 52. Analyse en tandem MS/MS

Introduction bibliographique

Nous détaillerons dans cette partie l'analyse en tandem par triple quadripôle, le Q-TOF et par l'association quadripôle orbitrap (Q-exactive). Le tableau 5 résume les principales caractéristiques techniques des analyseurs de masse utilisés actuellement en protéomique.

Tableau 5. Performances techniques des principaux analyseurs de masse (Holčapek et al., 2012)

Analyseur	Résolution	Précision (ppm)	Gamme de masse
Triple quadripôle	7500 (m/z 508)	5	10-3000
Q-TOF	22,500 (m/z 956)	<1	20-16,000
Orbitrap	100,000 (m/z 200)	<2	50-4000
Q-exactive	140,000 (m/z 200)	<1	50-4000
Q-ICR (Résonance Cyclotronique Ionique)	2500,000 (m/z 400)	<0,25	100-10,000

2.3.1.1. Le triple quadripôle

Un triple quadripôle, comme son l'indique, est constitué de trois quadripôles placés en série. Le premier quadripôle Q1 permet la sélection de l'ion parent. Le second quadripôle Q2 joue le rôle de cellule de collision. Ainsi, au moment où l'ion précurseur traverse le Q1, il subit une série de collisions au moyen d'un gaz inerte (généralement hélium ou argon). Ces collisions ont pour effet d'augmenter l'énergie interne de l'ion parent. De ce fait, lorsque l'énergie est suffisamment élevée, celui-ci subira une série de fragmentations (Paizs and Suhai, 2005). Les fragments produits sont ensuite analysés par le troisième quadripôle Q3 en fonction de leur rapport m/z. Ce mode d'acquisition sélective, caractéristique du triple quadripôle est appelé « *Single Reaction Monitoring* » (Aebersold and Mann, 2016; Lange et al., 2008; Ronsein et al., 2015; Gillet et al., 2016). Ainsi, dans le mode de balayage SRM, le premier et le troisième quadripôle agissent comme des filtres de manière à sélectionner spécifiquement les valeurs m/z prédéfinies et qui correspondent à l'ion précurseur et à son fragment spécifique (figure 53) (Lange et al., 2008). Une transition SRM correspond à un couple ion parent de rapport (m/z)₁ et d'un ion fils de rapport (m/z)₂.

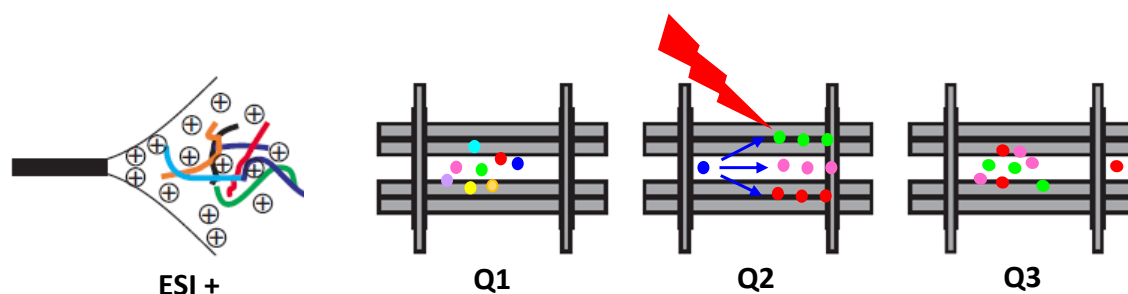


Figure 53. Représentation schématique du triple quadripôle (Mode *SRM*) (d'après Aebersold and Mann, 2016)

Un gros avantage de la SRM est une diminution des limites de détection du nanomole à l'attomole (Shi et al., 2016). Trois peptides dérivés du β -amyloïde ont été détectés dans le plasma humain à des concentrations infimes de l'ordre du $100 \text{ attomol/mL}^{-1}$. L'utilisation de la SRM pour la détection de ces biomarqueurs a permis d'augmenter de manière remarquable la sensibilité d'analyse sans passer par une étape d'immuno-enrichissement (Sano et al., 2014).

Néanmoins, lors de l'analyse des milieux biologiques, des ions interférents sont susceptibles d'être détectés par les quadripôles Q1 et Q3, en raison notamment du faible pouvoir résolutif de l'analyseur quadripolaire (Ronsein et al., 2015). C'est pourquoi des analyseurs à haute résolution doivent être considérés pour l'analyse des protéines dans les matrices complexes.

2.3.1.2. Le Q-TOF

Le Q-TOF est un spectromètre de masse hybride constitué de deux analyseurs, un quadripôle (Q) et un analyseur à temps de vol (TOF), séparés par une cellule de collision. Les ions formés dans la source ESI sont d'abord sélectionnés dans le quadripôle avant d'être fragmentés dans la cellule de collision. Les fragments ainsi obtenus sont focalisés au moyen d'un hexapôle et d'une série de lentilles. Les ions sont ensuite impulsés perpendiculairement à leur direction de départ et sont transmis par paquet vers le TOF (Lacorte and Fernandez-Alba, 2006) (figure 54). Le couplage quadripôle et TOF permet de gagner en termes de résolution (22,500 pour un m/z égale à 956) et de précision sur la mesure de la masse ($< 1 \text{ ppm}$) (Gillet et al., 2016) (tableau 5).

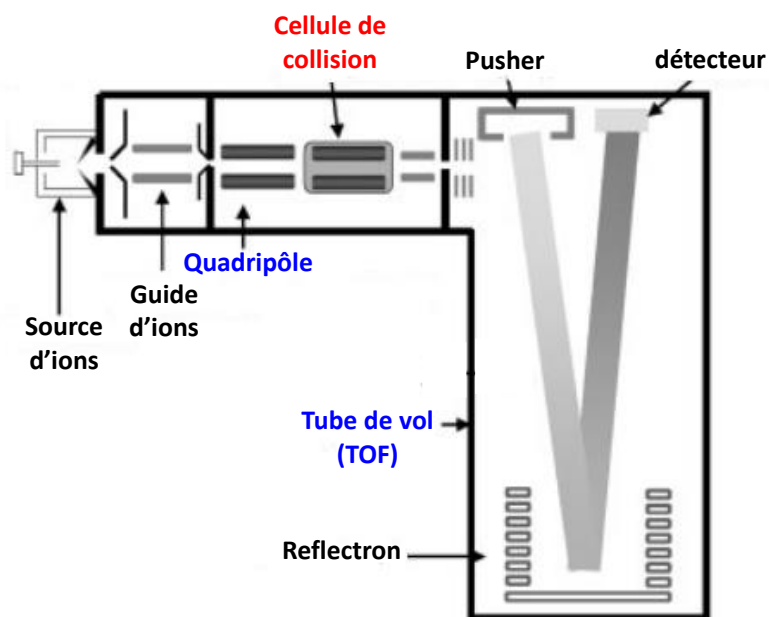


Figure 54. Représentation schématique du spectromètre hybride Q-TOF (Lacorte and Fernandez-Alba, 2006)

Le principe du TOF repose sur la mesure du temps de vol d'ions, préalablement accélérés par un champ électrique, dans une zone libre de champ (tube de vol) avant d'atteindre le détecteur. Le TOF est composé de deux régions dans lesquelles règne un vide poussé : une région d'accélération des ions sous l'action d'un champ électrique et une région libre de champ. En mode linéaire, les ions sont transmis dans le tube de vol avec une énergie cinétique qui dépend de la charge (z) de l'ion et de sa masse (m). Cela a pour conséquence de limiter les capacités du TOF à séparer des ions avec des rapports m/z proches. Néanmoins, le développement de l'extraction retardée et du mode réflectron (figure 54) ont permis d'homogénéiser les vitesses des ions lors de leur trajet ce qui se traduit par un gain de la résolution du TOF.

2.3.1.3. Le Q-exactive

Ce type de spectromètre combine les performances d'un analyseur quadripolaire utilisé pour sélectionner les ions parents, et d'un second analyseur orbitrap qui assure une détection à très haute résolution (figure 55).

Introduction bibliographique

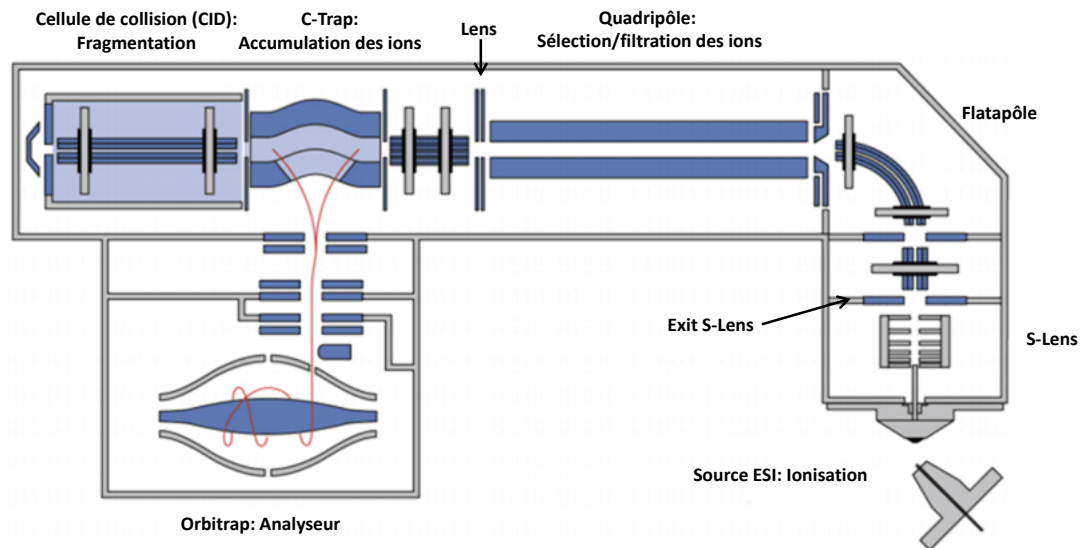


Figure 55. Représentation du Q-exactive (d'après *Thermo Fisher*)

Les ions multi-protonés générés à pression atmosphérique par une source ESI, sont dirigés vers une série de lentilles (*S-lens*) connectées au capillaire de transfert chauffé (275 °C, dans le cas de notre étude). Ces lentilles constituent le point d'entrée des ions vers le système et vont permettre ainsi d'accélérer, de focaliser et de transmettre les ions jusqu'au flatapôle *via* une lentille de sortie « *Exit S-Lens* ». Le processus de désolvatation se poursuit tout au long de ce déplacement par collision avec le gaz résiduel. Le flatapôle est constitué de quatre barreaux plats avec un espacement de 2 mm entre les barreaux. La courbure de ses barreaux (90 °C) favorise l'élimination des composés neutres et des gouttelettes de solvant. Les ions sont ensuite focalisés et transmis *via* une lentille « *Lens* » vers le premier analyseur, quadripôle. Celui-ci va permettre de sélectionner un ou plusieurs ions afin d'obtenir la gamme de masse désirée, généralement en protéomique entre 200 à 2000 m/z pour un « *Full-scan* » (Détection de tous les ions dans la gamme m/z de l'analyseur). Les ions précurseurs sont ensuite accumulés dans le piège à ions « *C-Trap* » ou fragmentés dans la cellule de collision (*High Energy Collision*, HCD), qui fonctionnent de façon synchronisée, avant d'être analysés par l'orbitrap. Les ions parents entrent en collision avec les molécules de gaz (azote), entraînant une rupture des liaisons chimiques résultant de la formation d'ions fils. L'énergie de collision appliquée est généralement comprise entre 20 et 50 eV pour la fragmentation des peptides. Avant leur détection dans l'orbitrap, les ions fragments ainsi formés sont retransmis et accélérés vers la « *C-Trap* », en appliquant une différence de potentiel (Michalski et al., 2011). Ainsi et contrairement au triple quadripôle qui fonctionne de manière continue, les ions sont accumulés dans la *C-Trap* avant analyse par l'orbitrap qui donne accès à la masse exacte des composés

Introduction bibliographique

d'intérêt. Le Q-exactive est caractérisé par une précision sur la masse mesurée inférieure à 1 ppm (tableau 5), ce qui lui permet de différencier deux ions avec un m/z très proches (exemple : 400,04965 et 400,04930 (0,00035 Da)).

L'orbitrap a été développé dans les années 2000 par Alexander Makarov (Makarov, 2000). Il est constitué d'une électrode externe creuse à l'intérieur de laquelle est placée coaxialement une électrode interne centrale en forme de fuseau. Les ions sont injectés en paquet tangentiellement à l'électrode interne au moyen d'interstices situés sur l'électrode externe. Un champ électrostatique quadripolaire est maintenu entre les deux électrodes permettant de piéger les ions autour de l'électrode centrale. Chaque ion entrant dans ce piège aura à la fois un mouvement circulaire autour de l'électrode, ainsi qu'un mouvement axial le long de cette électrode (figure 56).

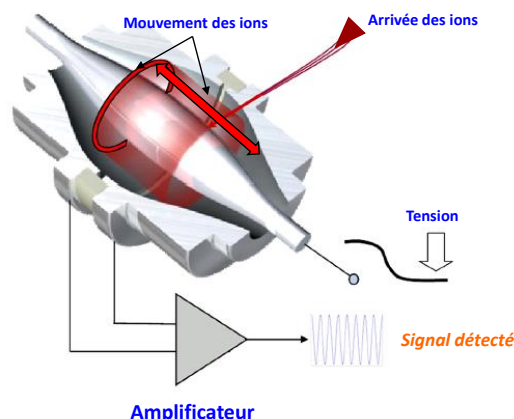


Figure 56. Représentation schématique de l'analyseur orbitrap (Zubarev and Makarov, 2013)

Les oscillations axiales des ions avec une fréquence f , dépendent particulièrement du rapport m/z (Hu et al., 2005; Makarov et al., 2009). Les courants induits par les oscillations des ions entre les deux électrodes, sont ensuite amplifiés et convertis par transformée de Fourier. Il en résulte alors la masse et l'intensité des ions, formant ainsi un spectre de masse (Hu et al., 2005; Makarov et al., 2009; Scigelova and Makarov, 2006).

L'orbitrap connaît un réel engouement dans le domaine de la protéomique. En effet, cette technologie a révolutionné l'analyse par spectrométrie de masse et s'est imposée très rapidement pour l'analyse des macromolécules biologiques. En grande partie grâce à ses performances techniques, un tel analyseur permet d'atteindre une résolution d'environ 100 000 ($m/z=200$) et une précision sur la mesure de masse inférieure à 2 ppm (Scigelova and Makarov,

2006). Le tableau 5 résume les principales caractéristiques techniques des analyseurs de masse utilisés actuellement en protéomique.

Les différents modes d'acquisition du Q-exactive

Le Q-exactive est adapté aussi bien aux analyses non ciblées (« *full Scan* ») pour l'identification de l'ensemble des protéines présentes dans un échantillon donné, qu'aux analyses semi-quantitatives et ciblées de type PRM « *parallel reaction monitoring* » (Ronsein et al., 2015; Schiffmann et al., 2014). Le mode balayage « *Full Scan* », permet la détection de tous les ions dans la gamme m/z sélectionnée par le quadripôle. Ce mode fournit une grande quantité d'information sur l'échantillon analysé. Le mode PRM (figure 57) très sélectif, repose sur la sélection de l'ion précurseur cible dans le quadripôle et de sa fragmentation dans la cellule de collision. Ainsi, les ions fragments générés sont ensuite transférés vers l'Orbitrap *via* la C-Trap. Il est alors possible d'identifier le rapport m/z de l'ion parent $(m/z)_1$ et celui de l'ion fils $(m/z)_2$. Par ailleurs, il est aussi possible de suivre de façon simultanée plusieurs transitions PRM et de réaliser ainsi un multiplexage dans le but d'analyser différents peptides d'intérêts (Shi et al., 2016). À ce jour le mode PRM présente de nombreux avantages par rapport au mode SRM (Shi et al., 2016). En effet, il a été démontré récemment que les performances analytiques pour la quantification de 35 peptides introduits dans un échantillon d'urine sont meilleures en mode PRM. Ainsi, ce mode a permis d'atteindre des limites de quantification plus faible pour 61 % des transitions par rapport au mode SRM du triple quadripôle (Gallien et al., 2013).

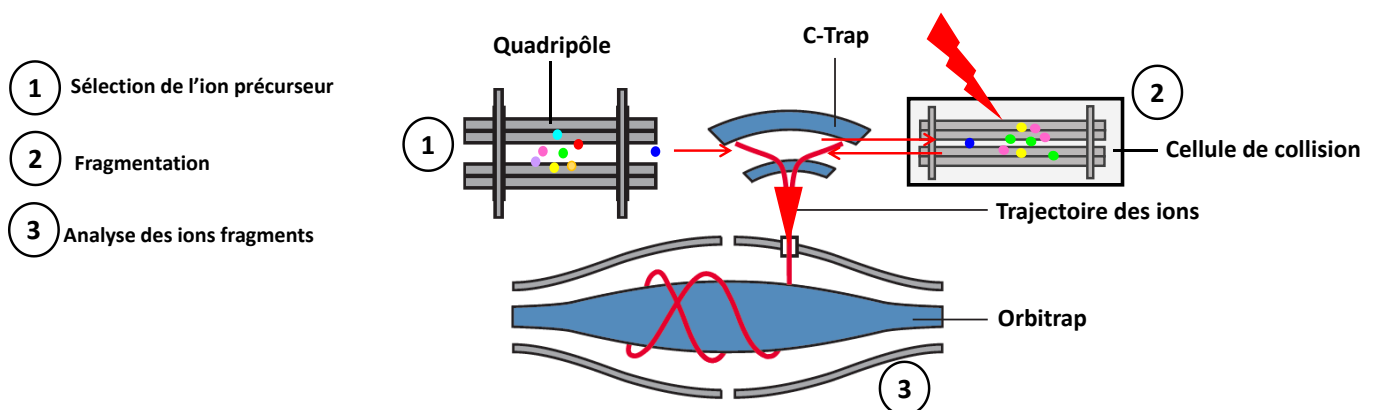


Figure 57. Principe du mode PRM (Aebersold and Mann, 2016)

Grâce à ce mode PRM, Kim et al. (2016), ont développé une approche multiplexe sur un Q-exactive, pour la quantification des tyrosines kinases (PTKs) (Kim et al., 2016).

Introduction bibliographique

De nouveaux modes d'acquisition des données en protéomique sont en plein développement avec l'essor de la technologie Q-orbitrap. Le mode *data dependent acquisition* (DDA) et le mode *data independent acquisition* (DIA). Ces modes d'acquisition concernent soit la sélection des ions précurseurs en fonction de leur intensité pour produire leurs spectres de fragmentation (fragmentation successive des ions) (DDA), soit la fragmentation de l'ensemble des ions précurseurs au cours de l'analyse (DIA) (Aebersold and Mann, 2016; Gillet et al., 2016).

Pour le mode DDA, les spectres MS et MS/MS sont générés de façon cyclique. Lors de l'éluion des ions peptides, une succession de scan MS et MS/MS est alors établie. Un scan MS est généralement suivi de plusieurs scans MS/MS sous la forme d'un TopN (N signifie le nombre d'ions précurseurs les plus intenses pouvant être fragmentés). Concernant le mode DIA, tous les peptides élués sont fragmentés en même temps sans aucune sélection préalable de l'ion précurseur. Ainsi, la fragmentation des ions précurseurs transmis à l'analyseur peut être réalisée de façon simultanée ou par fenêtre réduite d'environ 25 m/z (Aebersold and Mann, 2016; Shi et al., 2016). La figure 58 récapitule les différents modes d'acquisition des données MS/MS.

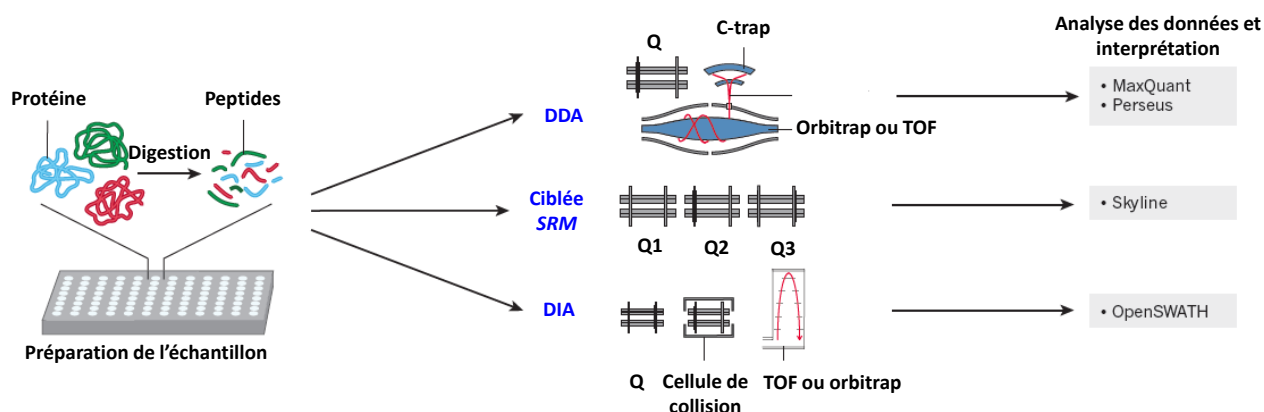


Figure 58. Les différents modes d'acquisition des données protéomiques. L'analyse des données est réalisée au moyen de logiciels dédiés à chaque mode d'acquisition (d'après Aebersold and Mann, 2016)

Bruderer et al. (2015), ont comparé les données de quantification générées par un Q-exactive entre deux analyses, réalisées en mode DDA et en mode DIA. Cette étude a montré que le mode DIA permet de gagner en sensibilité et en reproductibilité par rapport au DDA. Cette étude a permis d'établir un profil protéique de tissus de foie humain traités au paracétamol (Bruderer et al., 2015).

2.4. Interprétation des analyses de protéines par spectrométrie en tandem

La spectrométrie de masse tandem (MS/MS) appliquée à l'analyse des peptides issus de la digestion enzymatique des protéines, permet d'avoir des informations sur la séquence des acides aminés. L'ion peptidique est d'abord sélectionné par l'analyseur avant d'être fragmenté dans la cellule de collision. Les fragments ou les ions fils générés sont ensuite analysés pour donner le spectre MS². La fragmentation des peptides a lieu principalement au niveau de la chaîne peptidique et génère des fragments N-terminaux et C-terminaux. Lorsque la charge est portée par l'extrémité N-terminale, les ions formés sont appelés a,b et c. Tandis que, quand la charge est portée par l'extrémité C-terminal, les ions sont nommés x,y et z. Par ailleurs, lors de la fragmentation à basse énergie, la majorité des ions obtenus sont de type b et y (fragmentation au niveau de la liaison peptidique) (Paizs and Suhai, 2005) (figure 59).

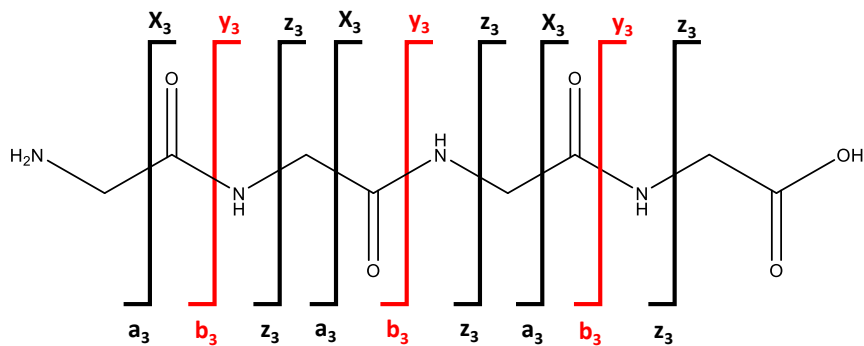


Figure 59. Nomenclature des ions peptidiques générés par fragmentation MS/MS

Ainsi, en analysant les différences de masse entre les différents ions détectés sur le spectre MS/MS, il est alors possible de retrouver, en se référant à une banque de données, la séquence en acides aminés et donc de déduire la séquence du peptide analysé. Un exemple d'identification est représenté dans le spectre MS/MS (figure 60).

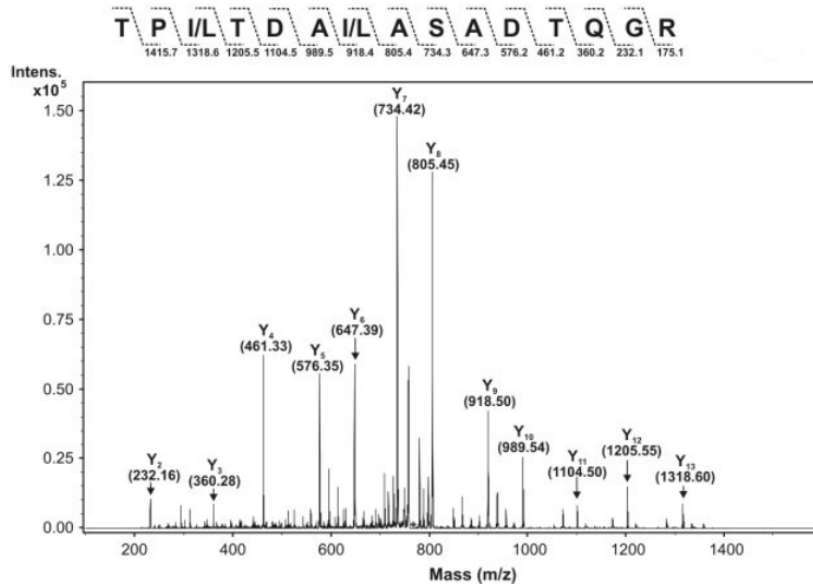


Figure 60. Exemple de spectre de fragmentation (Rinalducci et al., 2009)

Pour conclure, la caractérisation des protéines dans des mélanges hétérogènes et hyperdilués est à ce jour un réel défi scientifique. En effet, dans l'approche « *bottom-up* » l'étape de digestion trypsique augmente de manière considérable la complexité intrinsèque du mélange puisque le nombre de peptides générés est très conséquent. De ce fait, l'analyse s'avère complexe et requiert par conséquent l'utilisation de spectromètres de masse ultra-sensibles, capable de fournir des masses exactes comme la technologie hybride dotée d'un système orbitrap pour une analyse à haute résolution.

2.5. Stratégies de quantification

Il existe différentes stratégies de quantification des protéines par spectrométrie de masse. Elles peuvent être relatives ou absolues, avec ou sans marquage isotopique (Kito and Ito, 2008). La quantification relative consiste à comparer l'expression des protéines entre deux conditions différentes, généralement par le biais d'un marquage isotopique. Cette stratégie est la plus utilisée en protéomique différentielle (analyse des protéines dont les taux d'expression sont significativement différents entre deux échantillons). Différentes techniques de marquage peuvent être utilisées. (1) le marquage métabolique impliquant l'incorporation d'acides aminés marqués dans le milieu qui seront donc inclus dans la séquence de la protéine lors de la traduction de son ARN messager correspondant, (2) le marquage chimique des fonctions amines au moyen de réactifs commerciaux tel que le iTRAQ (*Isobaric Tags for relative and Absolute Quantitation*), ou encore le marquage chimique des résidus cystéines par le réactif

Introduction bibliographique

ICAT (*Isotope-Coded Affinity Tags*) et (3) le marquage protéolytique au cours de la digestion trypsine (en remplaçant l'eau désionisée de la digestion trypsique par de l'eau lourde ($H_2^{18}O$) (Castillo et al., 2014).

Les stratégies de quantification des protéines sans marquage, appelées aussi « *Label-free* » reposent sur la mesure de l'aire du pic chromatographique ou spectral (ions parents ou ions fils), ou encore sur le calcul de l'abondance relative des protéines (emPAI : *exponentially modified protein abundance index*). En ce qui concerne la quantification absolue des protéines, celle-ci est généralement réalisée par étalonnage interne avec ou sans marquage, ou étalonnage externe. Néanmoins, la mise au point d'une telle méthode se heurte généralement à la complexité des milieux étudiés. En effet, ils génèrent un effet matrice empêchant d'atteindre les niveaux de sensibilité recherchés (González-Antuña et al., 2013). La co-élution de composés interférents conduit à une diminution de l'efficacité d'ionisation au niveau de la source ESI ainsi qu'à des problèmes de séparation chromatographique des peptides ou protéines cibles (González-Antuña et al., 2013; Panuwet et al., 2016; Rogers et al., 2016) (figure 61). L'exemple le plus frappant reste celui de l'analyse des protéines minoritaires du plasma, qui en raison de la présence de protéines en très fortes concentrations (exemple albumine) et de nombreux métabolites (lipides, sucres, etc.) génère un effet significatif qui risque par conséquent de supprimer ou de masquer le signal de peptides d'intérêt (Trufelli et al., 2011). Ainsi, l'effet matrice constitue un point très important à prendre en compte dans l'analyse des protéines, puisqu'il peut impacter de manière significative les performances de quantification en termes de sensibilité et de répétabilité.

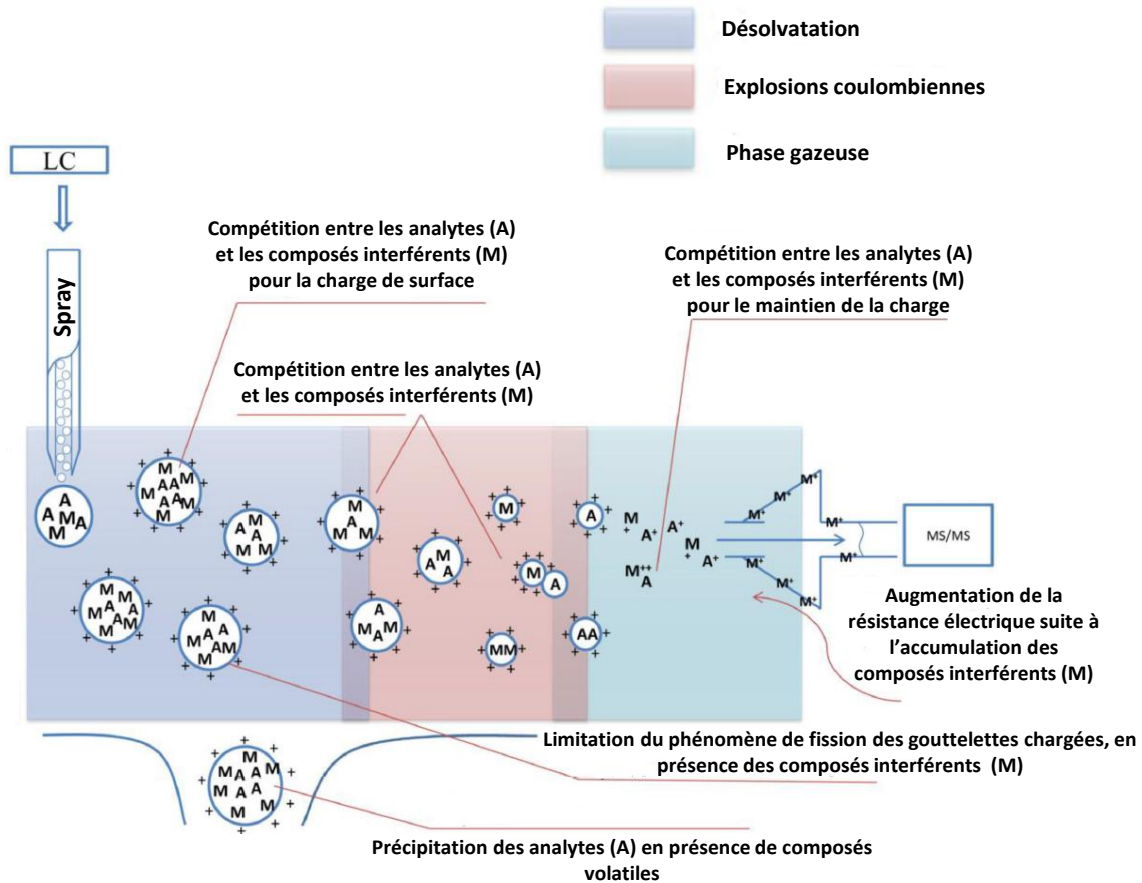


Figure 61. Représentation schématique de l'effet matrice sur l'ionisation des analytes par électrospray (Panuwet et al., 2016)

3. Autres méthodes d'analyse des protéines

3.1. Méthodes immunologiques

Les techniques immunologiques consistent à utiliser des anticorps en tant que réactifs analytiques pour la détection de protéines cibles. Très spécifiques, les anticorps permettent de détecter une protéine donnée présente dans un échantillon complexe. La technique immunoenzymatique ELISA consiste à visualiser le complexe anticorps-antigène (ici protéine cible) *via* une réaction colorée. Lorsque l'anticorps se fixe sur la cible protéique, une enzyme qui lui est liée est catalysée en présence de substrat aboutissant à la formation d'un produit coloré.

Les avantages de la technique de dosage par ELISA sont multiples : spécifique, simple et rapide, elle permet d'avoir des résultats semi-quantitatives et d'atteindre des limites de détection de l'ordre du nanogramme (i.e. détection de l'entérotoxine A staphylococcique avec une LOD

Introduction bibliographique

de $0,02 \text{ ng.mL}^{-1}$) (Kuang et al., 2013). Toutefois, cette technique est limitée à la détection d'une cible unique, elle nécessite également un équipement spécialisé et est fortement dépendante de la méthode de révélation (réaction enzyme-substrat).

Il existe d'autres techniques telles que le *western blotting* pour le dosage immunologique de protéines séparées sur gel SDS-PAGE. Les bandes de protéines cibles sont transférées vers une membrane de nitrocellulose, on parle de *blotting*. Un anticorps spécifique dirigé contre la protéine d'intérêt est alors additionné. Le complexe anticorps-antigène ainsi formé peut être détecté par un second anticorps marqué par une molécule fluorescente, permettant de ce fait la détection de la protéine d'intérêt. Cette technique longtemps utilisée dans les laboratoires de biochimie a notamment permis le dépistage des infections provoquées par le virus de l'hépatite C (VHC) qui détecte la protéine de la capsid (Yeh et al., 1994). Elle permet d'atteindre des limites de détection de l'ordre du picogrammes (Ye et al., 2013).

3.2. Détection des protéines en microscopie à fluorescence

Contrairement aux techniques de détection biochimiques réalisées le plus souvent en milieu liquide ou sur gel d'électrophorèse, les marqueurs fluorescents permettent d'étudier les protéines dans leur milieu biologique. Des cellules, contenant les protéines d'intérêt peuvent être colorées, analysées par microscopie à fluorescence et ainsi renseigner sur la localisation de la protéine (figure 62).

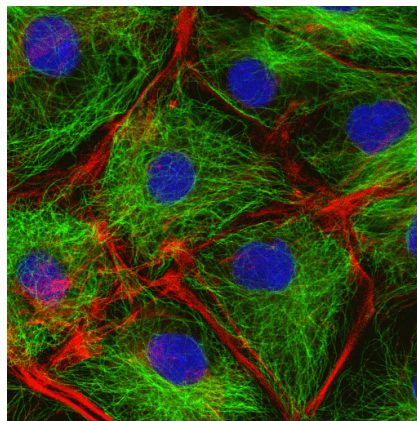


Figure 62. Micrographie d'une cellule animale après marquage par des fluorescents spécifiques. Les filaments d'actine et les microtubules sont colorés respectivement en rouge et en vert (Risler, 2009). Source: Mark Shipman, James Blyth and Louise Cramer, MRC-Laboratory Molecular Cell Biology and Cell Biology Unit, UCL, London UK

3.3. Etude de la structure des protéines

Le repliement des protéines est un processus très complexe mettant en jeu différents états de repliement, de la formation des structures secondaires jusqu'à l'obtention de structures tertiaires ou quaternaires de plus en plus complexes. Ainsi et afin de mieux appréhender la dynamique de repliement ou dénaturation des protéines, il paraît important d'étudier les structures intermédiaires d'une protéine soumise à différentes conditions expérimentales. Pour une protéine donnée, le choix de la technique spectrale dépendra de ce fait, des objectifs de l'expérience que l'on souhaite atteindre.

3.3.1. La cristallographie aux rayons X

De nombreuses techniques permettent d'évaluer la conformation des protéines et de mieux appréhender leur fonction biologique. Au niveau atomique, la cristallographie aux rayons X fut la première méthode développée pour la détermination des protéines. Cette technique permet de visualiser la disposition tridimensionnelle de la plupart des atomes d'une protéine. En effet, parmi tous les types de radiations, les rayons X peuvent atteindre les meilleures résolutions atomiques pour l'interprétation des structures moléculaires. L'analyse cristallographique aux rayons X de la structure tridimensionnelle d'une protéine requiert plusieurs éléments essentiels : un cristal de protéine, une source de rayon X ainsi qu'un détecteur (figure 63).

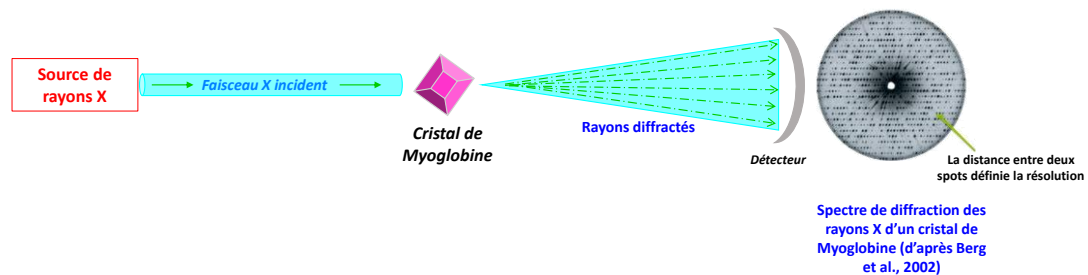


Figure 63. Principe de la cristallographie aux rayons X, la source de rayons X produit un faisceau incident qui sera diffracté par le cristal. Le spectre de diffraction est ensuite recueilli par un détecteur

L'obtention d'un cristal de protéine est une étape limitante qui va dépendre de la pureté de la protéine. En effet, lors du processus de cristallisation toutes les molécules de protéines doivent s'orienter suivant une organisation répétée et constante.

3.3.2. Spectroscopie RMN

Dans les années 80, une autre méthode permettant d'élucider la structure tridimensionnelle des protéines fut introduite. Il s'agit de la spectroscopie par résonance magnétique (RMN). Son principe repose sur la propriété que possèdent certains noyaux atomiques, appelée *spin* nucléaire (^1H , ^2H , ^{13}C , ^{14}N etc.). À l'état naturel, ces spins (petits aimants) sont caractérisés par une très faible intensité. Dès lors qu'ils sont soumis à l'action d'un champ magnétique, les spins nucléaires s'alignent soit dans la même direction de ce champ soit dans une direction opposée. L'énergie apportée au système sous forme d'ondes électromagnétiques correspondantes à des fréquences permet de provoquer une variation du comportement magnétique de l'échantillon (aimantation nucléaire) qui est à la base de la résonance magnétique. L'interaction entre les spins des noyaux fournit des informations structurales décisives pour la détermination des structures tridimensionnelles des protéines. Cette technique unique révèle la structure atomique et la dynamique des protéines en solution. Récemment, la RMN bidimensionnelle (2D) réalisée au sein des cellules vivantes (RMN en cellule) a permis d'explorer les structures protéiques, l'échelle moléculaire (Müntener et al., 2016).

3.3.3. Spectrofluorométrie

3.3.3.1. Fluorescence intrinsèque

Les protéines présentent des fluorophores dans leur structure primaire qu'il est possible de détecter par spectrofluorométrie. Les fluorophores sont des entités chimiques, qui passent d'un état fondamental à un état excité, lors de l'absorption d'un photon d'énergie $h\nu$ (lumière d'excitation, dans le domaine de l'UV-visible). Cet état étant instable, la molécule retourne à son état fondamental en restituant une partie de l'énergie absorbée sous forme de lumière fluorescente (lumière d'émission).

La spectrofluorométrie 3D est plus particulièrement utilisée dans le domaine de l'environnement. Outre sa rapidité d'analyse, elle est employée pour caractériser la matière organique dissoute (MOD) dans les milieux sédimentaires ou aquatiques (Fu et al., 2005). Les fluorophores représentatifs de structures types humiques ou protéiniques sont ciblés. La figure 64, adaptée de Chen et al. (2003), représente les différentes zones de fluorophores qui correspondent à des composés organiques tels que les protéines et les substances humiques.

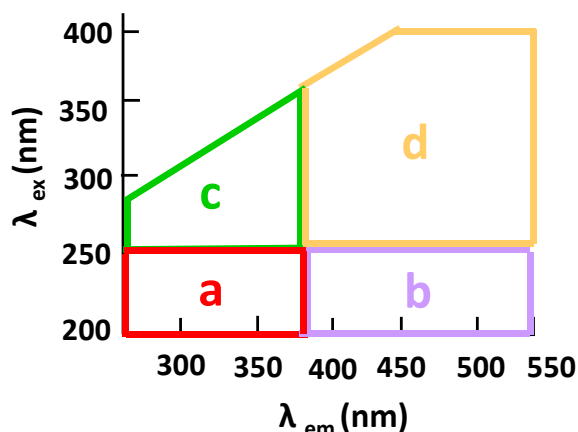


Figure 64. Représentation schématique des différentes zones de fluorophores; a : zone caractéristique des protéines (acides aminés aromatiques) ; b : zone fluorophore caractéristique des acides humiques ; c : zone fluorophore caractéristique des sous-produits protéiques microbiens ; d : zone fluorophore caractéristique des acides humiques (Chen et al., 2003)

Il existe diverses méthodes utilisées en spectrofluorométrie pour le suivi du repliement des protéines (Ghisaidoobe and Chung, 2014; Uversky and Dunker, 2010) :

La fluorescence intrinsèque des protéines est due à la présence des résidus d'acides aminés aromatiques tryptophane, tyrosine et phénylalanine qui ont la propriété d'absorber dans le domaine de l'UV et d'émettre des photons de fluorescence. Parmi les trois acides aminés aromatiques, le tryptophane, grâce à son noyau indolique, participe généralement à 90 % de la fluorescence totale des protéines (tableau 6 ; figure 65) (Ghisaidoobe and Chung, 2014).

Tableau 6. Données spectrales du tryptophane, tyrosine et phénylalanine dans une solution d'eau à pH neutre (Ghisaidoobe and Chung, 2014)

Acide aminé	Absorption		Fluorescence	
	λ_{\max} (nm)	ϵ ($M^{-1}cm^{-1}$)	λ_{\max} (nm)	(Φ_F)
Tryptophane (W)	280	5600	348	0,2
Tyrosine (Y)	274	1400	303	0,14
Phénylalanine (F)	257	200	282	0,04

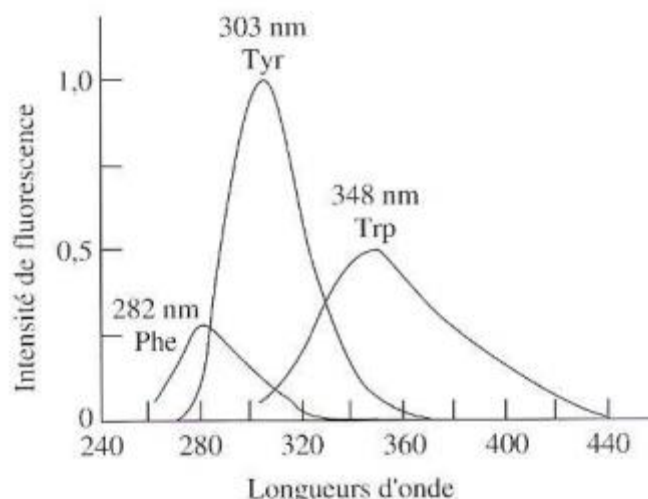


Figure 65. Spectre de fluorescence des acides aminés aromatiques dilués dans un tampon phosphate à pH 7, λ_{ex} (nm) = 270 nm (Albani JR, 2001)

Le résidu tryptophane peut être utilisé pour suivre des changements de conformation d'une protéine donnée. À titre d'exemple, Qadeer et collaborateurs (Qadeer et al., 2012) ont étudié la dynamique de dénaturation de la chymopapaïne à différents pH. Les résultats ont mis en évidence l'effet de l'environnement local (polaire ou apolaire) sur l'intensité de la fluorescence intrinsèque due à ses quatre résidus tryptophanes. Toutefois, la sensibilité des fluorophores naturels reste relativement faible. En effet, l'intensité de la fluorescence dépend fortement de leur nombre, de leur localisation/exposition par rapport au solvant. Des fluorophores extrinsèques ont ainsi été développés dans le but d'améliorer la sensibilité de détection de ces techniques et de fournir des informations pertinentes sur la structure protéique.

3.3.3.2. Fluorescence extrinsèque

Généralement, la fluorescence extrinsèque repose sur l'utilisation de fluorophores susceptibles de se combiner à certains groupements des protéines pour les rendre fluorescentes. Ainsi le spectre d'émission va dépendre non seulement de la nature de la molécule fluorescente, mais également des interactions mises en jeu entre le fluorophore et les constituants du milieu (protéique et non protéique).

De nombreux fluorophores peuvent être utilisés, comme les dérivés réactifs de la fluorescéine ou de la rhodamine, l'acide cis-parinarique, ou encore le 1-anilino-8-naphthalène sulfonate (ANS) (Lakowicz, 2006). Le 1-anilino-8-naphthalène sulfonate (ANS) est le plus utilisé pour l'étude des processus de dénaturation des protéines. Lorsqu'il est en solution

Introduction bibliographique

aqueuse, il émet un signal fluorescent faible. L'intensité de fluorescence de l'ANS augmente avec l'hydrophobicité du milieu, comme le montre la figure 66.

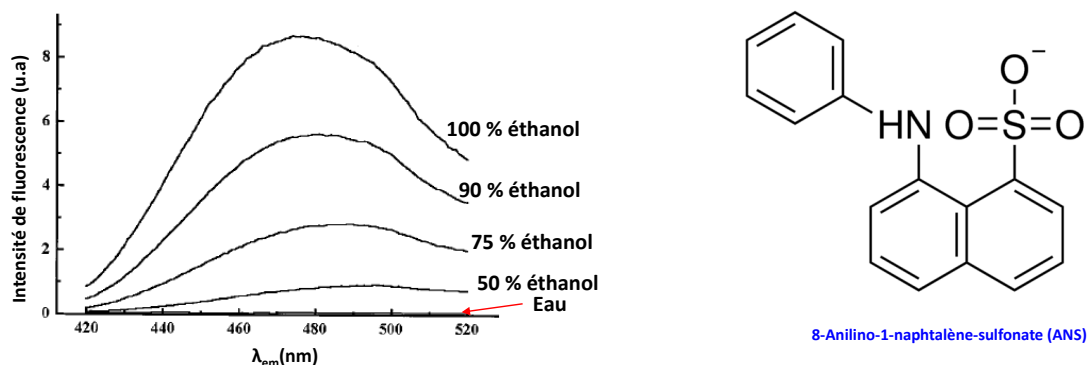


Figure 66. Spectre de fluorescence de l'ANS dans différents environnements : polaires et apolaires. L'intensité du signal de l'ANS augmente en fonction de l'hydrophobicité du milieu

Lorsqu'il est ajouté dans un échantillon protéique, il va se fixer aux régions hydrophobes et former ainsi un complexe fluorescent après excitation entre 370 et 400 nm. L'ANS, a permis pour la première fois de suivre l'enfouissement des résidus hydrophobes au cours du processus de repliement des protéines. Il a été également utilisé par Semisotnov et al. (Semisotnov et al., 1991) pour mettre en évidence l'état *molten globule* des protéines (défini dans le chapitre I, 1.2.1) en fonction des conditions physicochimiques de leur milieu (Semisotnov et al., 1991).

Nous utiliserons l'ANS lors de notre étude pour identifier les changements de structure des protéines modèles lors des différents traitements de purification.

3.3.4. Spectroscopie Raman

La spectroscopie Raman consiste à mesurer la réponse vibrationnelle de structures moléculaires ou cristallines présentes dans un échantillon. Pour cela l'intensité de la lumière diffusée après excitation par un rayonnement électromagnétique.

Le principe physique de la spectroscopie Raman repose sur la diffusion inélastique, par la matière, d'une radiation monochromatique. Découvert en 1928 par le physicien Raman (Raman and Krishnan, 1928) lors de l'étude de la diffusion de la lumière par des liquides, ce travail a été ensuite récompensé en 1930 par un prix Nobel en physique. Lorsqu'un échantillon est éclairé par un faisceau de lumière monochromatique (source laser) de fréquence connue, une

Introduction bibliographique

grande partie du rayonnement émis est généralement transmise sans modification de la fréquence. Ainsi, la plupart des photons conservent la même énergie et sont donc diffusés, avec la même fréquence que les photons incidents. On parle alors de phénomène de diffusion élastique sans changement de fréquence (diffusion Rayleigh). Toutefois, pour environ 1 photon sur 10^8 , un changement de fréquence est observé. Dans ce cas, les photons sont émis avec une longueur d'onde différente de la longueur d'onde incidente. Ce phénomène inélastique correspond à l'effet Raman (Butler et al., 2016) (figure 67).

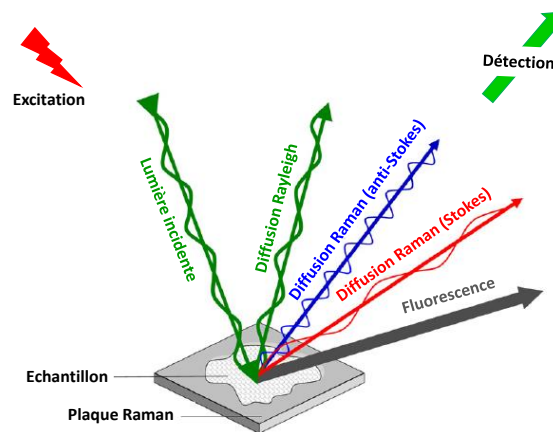


Figure 67. Représentation schématique de la diffusion de la lumière d'un échantillon éclairé par une source laser. Lorsqu'une partie des photons est diffusée avec une fréquence inférieure à la fréquence incidente, on parle de diffusion Stokes. A l'inverse, si les photons sont diffusés avec une fréquence supérieure à la fréquence incidente, la diffusion Raman est de type anti-Stokes

Un spectre Raman comporte un certain nombre de raies (ou bandes) dont les fréquences sont reliées aux vibrations des différents groupes moléculaires constituant le matériau. Les différentes structures protéiques donnent lieu à des bandes de vibrations, qui sont attribuées d'une part à la chaîne polypeptidique et d'autre part aux chaînes latérales de la protéine. En Raman trois types de bandes dites amides sont majoritaires : la bande amide I correspond principalement à la vibration d'élongation du groupement carbonyle C=O du squelette peptidique, elle est observée entre $1600-1690\text{ cm}^{-1}$. La bande amide II, entre $1480\text{ et }1580\text{ cm}^{-1}$, ainsi que la bande III, entre $1230-1300\text{ cm}^{-1}$, correspondent à la vibration de déformation de la liaison N-H couplée à la vibration d'élongation de la liaison C-N (Rygula et al., 2013). La

Introduction bibliographique

figure 68 illustre deux spectres Raman de deux protéines, que sont le lysozyme (L) et la chymotrypsine (C) (Hedoux et al., 2013).

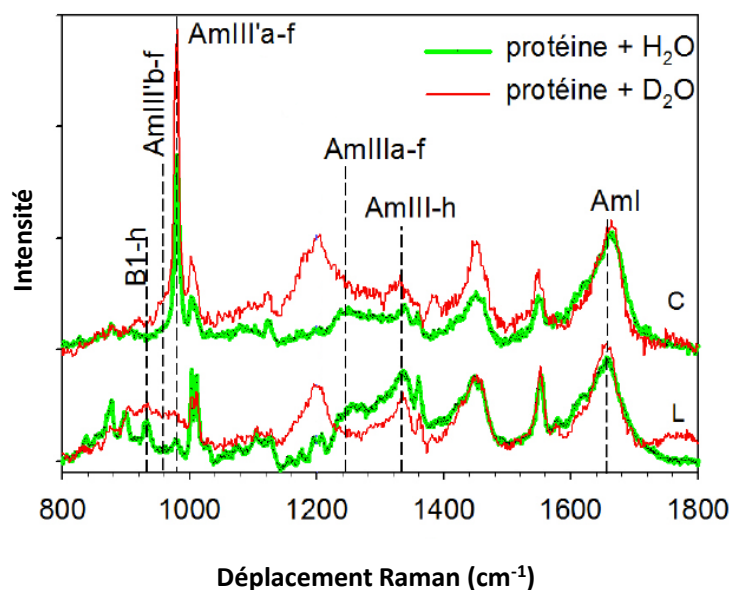


Figure 68. Spectre Raman des bandes amides de deux protéines solubilisées dans le H₂O et le D₂O ; L : Lysozyme ; C : Chymotrypsine. L'attribution des différentes bandes de vibration est comme suit : B1h (940 cm⁻¹) : squelette N-C α -C (Mode d'étirement relatif à l'hélice α), AmIIIb'-f (960 cm⁻¹) : Amide III' feuillet- β de l'échange isotopique, AmIIIa'-f (980 cm⁻¹) : Amide III' feuillet- β de l'échange isotopique, AmIIIa-f (1240 cm⁻¹) : Amide III feuillet- β , AmIII-h (1340 cm⁻¹) : Amide III hélice- α , AmI (1650-1660 cm⁻¹) : Amide I (Hedoux et al., 2013)

Les bandes amides I et III contribuent considérablement à la caractérisation de la structure secondaire des protéines (Rygula et al., 2013). De manière générale, la bande amide I est la plus intense. Cette bande caractéristique est relativement complexe et peut être décomposée en différentes sous-bandes ou composantes : les hélices α , les feuillets β , les structures désordonnées ou encore les coudes β (David et al., 2008). Outre les bandes amides, des vibrations attribuées aux chaînes latérales peuvent être aussi observées. Ces vibrations dépendent particulièrement de l'environnement hydrophobe ou polaire de la protéine. Ainsi on distingue, les vibrations des résidus tryptophane, tyrosine ou encore des vibrations associées à l'élongation des liaisons C-S et S-S (Rygula et al., 2013).

La spectroscopie Raman est une technique parfaitement adaptée pour l'étude de la conformation des protéines. Elle offre l'avantage d'être une technique non destructive et non

Introduction bibliographique

invasive (Tuma, 2005). L'étude des protéines par spectrométrie Raman suscite un grand intérêt parmi la communauté scientifique. En effet, les informations fournies par cette technique ont permis de modéliser les différents états conformationnels de certaines protéines dans le but de mieux comprendre leur activité et leur la fonction biologique au sein des cellules (Oshokoya and JiJi, 2015). À titre d'exemple, dans le domaine pharmaceutique la spectroscopie Raman a été utilisée pour étudier le rôle des polysaccharides dans la stabilisation de protéines à visée thérapeutique (Hédoux et al., 2009). D'autres études se sont intéressées aux processus de dénaturation et de stabilisation des protéines soumises à différentes conditions physicochimiques (Ashton and Blanch, 2010; Hédoux et al., 2010; Jordan et al., 1995). Dans le domaine de la santé, la spectroscopie Raman a été aussi utilisée pour étudier les maladies ayant pour origine un mauvais repliement des protéines (exemple, peptide β -amyloïde dans la maladie d'Alzheimer ou encore la maladie de Parkinson) (Kurouski et al., 2015).

Objectifs et stratégies expérimentales

Objectifs et stratégies expérimentales

L'étude bibliographique a mis l'accent sur les challenges analytiques liés à l'extraction et la purification des protéines qui impliquent généralement des protocoles longs et multi-étapes (pour certains, jusqu'à 10 étapes sont proposées). Par ailleurs, ces approches sont basées le plus souvent sur des études « essai-erreur » au cours desquelles les interactions établies entre les protéines et leurs milieux ne sont pas considérées. L'effet de ces différents agents sur la structure même des protéines n'est pas plus examiné alors que cela permettrait probablement de comprendre l'efficacité ou l'inefficacité de tel ou tel agent et donc d'orienter la stratégie expérimentale à suivre. En outre, les protocoles développés en protéomique sont pour la plupart spécifiques à un nombre restreint de protéines ou à une seule matrice. Ils utilisent de manière générale de nombreux agents chimiques agressifs (détergents, agents chaotropes, etc.), voire même toxiques (le phénol par exemple est très utilisé pour purifier les protéines des sols), incompatibles avec une analyse par spectrométrie de masse.

Pour contourner ces verrous analytiques, une voie consiste à travailler sur les protocoles amont. Dans ce contexte, mon travail de thèse a pour objectif d'étudier la possibilité d'une approche simplifiée et non ciblée, permettant l'extraction, la purification et l'analyse des protéomes dans leur ensemble. En outre, elle devra être facilement reproductible et transposable à différentes matrices (biologiques ou environnementales), sans pour autant diminuer la sensibilité et la robustesse de l'analyse par UPLC-HRMS/MS. Mon travail de thèse s'articule autour de trois objectifs principaux :

- Le premier correspond à l'optimisation des paramètres d'UPLC-HRMS/MS en vue d'une identification et quantification optimales des protéines, et compatibles avec une stratégie protéomique « *bottom-up* ». Au cours de cette étude, un système chromatographique « *dual trap* » permettant la concentration en ligne des protéines faiblement abondantes et l'amplification de manière considérable de leurs signaux en spectrométrie de masse a été mis au point. Dans une deuxième partie, les paramètres relatifs à l'analyse des protéines par spectrométrie de masse haute résolution ont été optimisés. Les peptides prototypiques issus de la digestion des protéines standards, utilisées pour développer les protocoles de préparation d'échantillon amont, ont été déterminés. Une méthode de quantification relative des peptides basée sur les spectres des ions fragments (MS/MS) a ensuite été optimisée. Enfin, les différents paramètres influençant la digestion trypsique des protéines ont été étudiés.

- Le deuxième volet de ce travail est consacré au développement d'une approche de purification simplifiée mais efficace quelles que soient les matrices étudiées

Objectifs et stratégies expérimentales

(environnementales ou biologique). Dans le cadre de cette étude, un nouvel agent de lavage et de solubilisation, l'éthanol/HCl a été proposé. L'effet attendu de cet agent était de déstabiliser de la structure compacte dite de *molten globule* caractéristique des protéines lorsque celles-ci sont précipitées avec du TCA. Après avoir démontré son efficacité sur solutions standards, nous avons étudié à l'échelle moléculaire le mécanisme d'action de l'éthanol/HCl. L'objectif ici était de comprendre comment il modifiait la structure secondaire des protéines, et en quoi ces modifications amélioreraient leur solubilité dans les tampons aqueux utilisés pour la digestion trypsique. En outre, cette étude nous a permis de généraliser le mode d'action de l'éthanol/HCl sur des protéines ayant pourtant des structures primaires et secondaires différentes. Pour finir, l'efficacité de ce nouvel agent de purification a été vérifiée sur des extraits protéiques obtenus à partir de matrices complexes.

- Une fois les conditions d'analyse UPLC-HRMS/MS et l'étape de purification optimisées, nous nous sommes intéressées à la première étape de prétraitement d'échantillon, à savoir, l'extraction. Dans le cadre de cette dernière étude, les différents paramètres physiques et chimiques influençant l'extraction des protéines cibles (dopées) ont été étudiés. L'idée était ici de faire un premier screening des paramètres les plus influents. Bien que des pistes d'amélioration des protocoles existants aient été étudiées, un protocole d'extraction optimal ne peut encore être proposé.

Cette séquence d'étude (1. analyse, 2. purification, 3. extraction) a été choisie suivant l'objectif ultime d'avoir des signaux en masse amplifiés. L'étape extraction étant la principale étape limitante, celle-ci a été étudiée une fois l'analyse, la digestion et la purification optimisée.

Ces trois parties ont été construites sur la base d'une même démarche expérimentale. Tout d'abord, les protocoles développés devant répondre à une forme de « polyvalence » en termes de protéines ciblées et de matrices étudiées, ils ont tous été optimisés et validés sur des protéines et matrices complexes de nature et composition très diverses.

S'agissant des protéines, quatre protéines standards ont été utilisées. Ces macromolécules diffèrent par leurs propriétés physicochimiques (structures primaires, secondaires et tertiaires, poids moléculaire, pH isoélectrique, etc), ainsi que par leur application. L'exoglucanase-1 est une enzyme extracellulaire d'origine bactérienne ou fongique fréquemment retrouvée dans le sol et responsable de la décomposition de la cellulose. Le lysozyme est une enzyme qui catalyse l'hydrolyse des peptidoglycanes chez les bactéries Gram (+), et est utilisé comme modèle dans

Objectifs et stratégies expérimentales

de nombreuses études protéomiques en raison de sa forte stabilité (Aslebagh et al., 2016; Tomaszewski et al., 2011). La ribonucléase est une nucléase qui catalyse la dégradation de l'ARN. Elle est présente chez tous les organismes vivants. Enfin, l'albumine sérique humaine (HSA) est une protéine soluble très abondante dans le plasma humain. Elle est produite par le foie et présente des propriétés de transport d'hormones, de médicaments (diazépam, ibuprofène), d'acides gras etc...(Ghuman et al., 2005). Ces quatre marqueurs protéiques sont hydrophiles, caractérisés par un nombre élevé en acides aminés hydrophiles (indices d'hydropathie, -0,35 (HSA) ; -0,15 (lysozyme) ; -0,213 (Rnase) ; -0,43 (exogulucanase-1), respectivement) (Magdeldin et al., 2012). Les différentes caractéristiques physicochimiques des protéines cibles sont illustrées sur la figure 69.

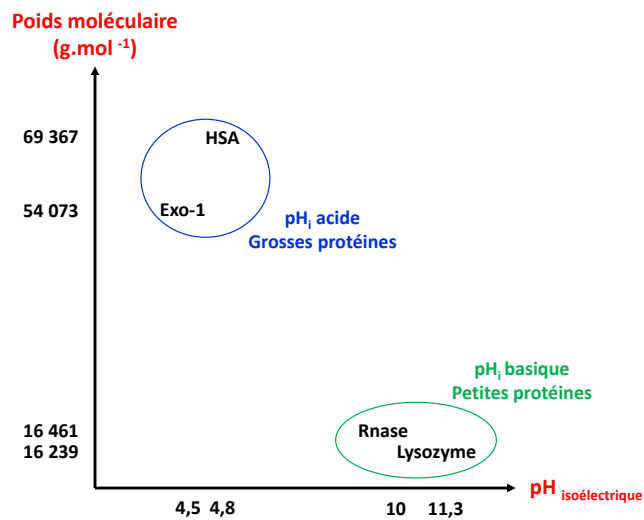


Figure 69. Caractéristiques physicochimiques des cibles protéiques

S'agissant des milieux, les stratégies développées pour purifier puis extraire les protéines ont toujours été testées et optimisées sur solutions standards et/ou matrices modèles, puis validées sur matrices complexes. Là encore, dans un souci de versatilité, des échantillons complexes biologiques et environnementaux ont été sélectionnés.

Enfin, au cours de mon projet de thèse, j'ai utilisé les compétences, outils et stratégies des chimistes pour mieux comprendre ces marqueurs biologiques. L'objectif était plus particulièrement ici d'apporter de nouvelles connaissances fondamentales pour comprendre la dynamique de ces molécules au cours de leur extraction, purification et analyse. Dans ce contexte, nous avons effectué une étude mécanistique afin de valider l'étape de purification, *via* l'arsenal spectroscopique (spectrofluorométrie en présence d'ANS, spectroscopie Raman, la

Objectifs et stratégies expérimentales

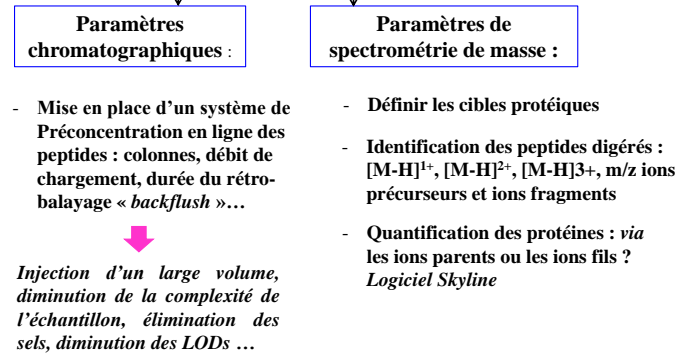
microscopie de force atomique) dont nous disposons au laboratoire. A l'issue de cette étude fondamentale, nous avons démontré au niveau moléculaire le « fonctionnement » du nouvel agent chimique que nous avons proposé.

Le diagramme représenté ci-dessous récapitule l'organisation des trois chapitres de résultats et discussions.

Chapitre I

Développement analytique

1 Mise en place d'une méthode analytique pour des applications en protéomique « hors gel »



2 Optimisation des paramètres de digestion trypsique

- Effet du tampon, ratio enzyme/substrat ...

Chapitre II

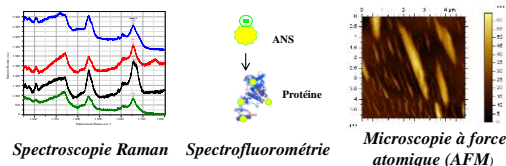
Développement d'une approche simple pour la purification des protéines

1 Mise en place d'un protocole de purification des protéines

Choix du solvant de purification

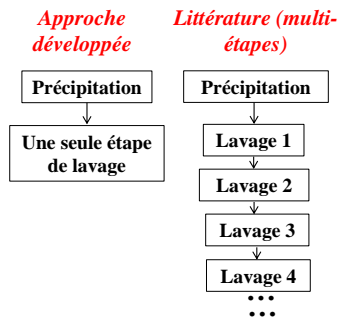
2 Etude de la dynamique conformationnelle des protéines au cours de la purification

Justification du choix du solvant de purification



3 Application de l'approche développée sur matrices réelles

Comparaison avec des approches publiées dans la littérature



Chapitre III

Influence des différents paramètres expérimentaux lors de la mise en place d'un protocole d'extraction des protéines

1 Etude de l'extraction des protéines sur matrice simple

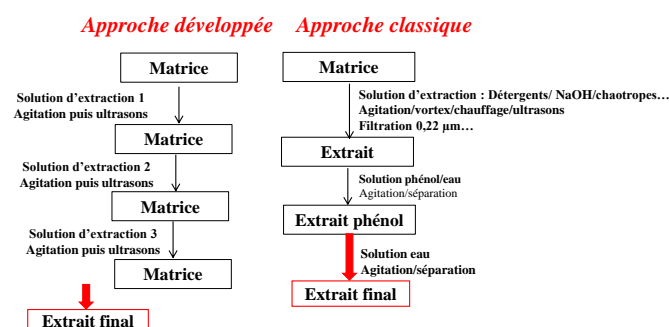
Effet de la filtration; choix du solvant d'extraction...

2 Etude de l'extraction des protéines sur matrice complexe : environnementale et biologique

Paramètres physiques et mécaniques Paramètres chimiques

Filtration, température, temps d'extraction, agitation vs ultrasons

Solvants...



Partie expérimentale

Partie expérimentale

1. Matériels

1.1. Protéines, peptides et acides aminés standards

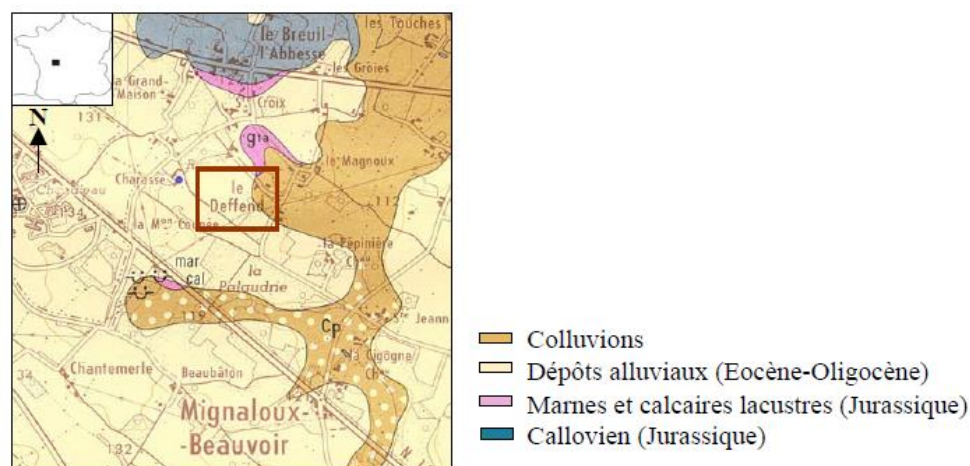
Quatre protéines ont servi de modèles standards au cours de ce travail de thèse. La ribonucléase pancréatique bovine Rnase (EC 3.1.27.5), la cellulase de *Trichoderma reesei* ATCC 26921 (qui est en fait un mélange de plusieurs exo- et endoglucanases, 1,4- β -D-glucane-4-glucanohydrolase (EC 3.2.1.4), le lysozyme (N-acétylmuramoylhydrolases du peptidoglycane) (EC 3.2.1.17) et enfin l'albumine humaine (HSA). Elles ont toutes été fournies par Sigma Aldrich (Saint Louis, USA).

Des acides aminés basiques ont également été utilisés au cours des étapes d'optimisation de l'extraction protéique. Il s'agit de l'histidine (99 %), de l'arginine (99 %) et de la lysine (99 %). Une solution de cinq peptides a été également utilisée comme standard analytique (Glycine-tyrosine, valine-tyrosine-valine, leucine-enképhaline, méthionine-enképhaline et standard de digestion, angiotensine-II). Ces produits sont fournis par Sigma Aldrich (Saint Louis, USA).

1.2. Matrices environnementales et biologiques étudiées

Plusieurs matrices ont été étudiées au cours de ce travail de thèse :

- Matrices environnementales : Sol du Deffend : le site du Deffend est localisé à Mignaloux Beauvoir à 2 km à l'Est de l'Université de Poitiers (figure 70). Il s'agit d'un sol limono-argileux caractérisé par une faible stabilité structurale. La composition lipidique et en acides humiques a fait l'objet d'une thèse (Célérier, 2008).



Partie expérimentale

Le site du Deffend est situé à la charnière du Massif Armoricaïn et du Massif Central et des Bassins Aquitain et Parisien, au niveau de la zone surélevée du seuil du Poitou. La couverture sédimentaire est composée de terrains calcaires et marno-calcaires jurassiques. Au Tertiaire, l'exondation (retrait de l'eau d'une prairie inondée) de ce domaine a mené à la karstification des calcaires. À l'époque de l'Eocène-Oligocène, la région de Poitiers est dominée par un paysage fluvio-lacustre sur lequel viennent se répandre les produits de l'altération issus du démantèlement du Massif Central. De ce fait, les formations superficielles possèdent une distribution spatiale hétérogène alternant des limons argileux riches en argiles gonflantes et des matériaux détritiques de nature argilo-sableuse. Les sols ont été échantillonnés à la tarière (0-15 cm) puis congelés à -20 °C, avant d'être lyophilisés à une pression 1,03 mbar pendant 12h00 (Christ, modèle Alpha 1-4). Ils ont été ensuite tamisés à 500 µm afin d'éliminer tous les résidus végétaux.

La deuxième matrice environnementale utilisée correspond à du sable (matrice minérale pure) fourni par Fisher Scientific (Illkirch, France).

- Matrices biologiques : les échantillons de foie de souris SWISS et des xénogreffes de tumeurs du carcinome pulmonaire de souris C57BL/6, proviennent du Pôle Biologie-Santé de Poitiers. Ils ont préalablement été congelés à -20 °C avant l'expérience. Ces échantillons nous ont été délivrés par Jonathan Clarhault, chercheur au CHU de Poitiers.

- Biofilm : cette matrice a été échantillonnée dans la rivière de la Vienne localisée du côté de Clupeau (D109E) dans la commune de Tarnac. Il s'agit de biofilms épilithiques qui colonisent donc la surface des roches immergées. Les roches situées dans le lit de la rivière ont d'abord été prélevées. La surface de chaque roche a ensuite été grattée dans le but de décrocher le maximum de biofilm puis rincée avec de l'eau ultrapure. Les solutions de biofilm ont été congelées à -20 °C avant d'être lyophilisées (lyophilisateur Cosmos 20 K Cryotec, France). L'échantillon de biofilm nous a été délivré sous forme lyophilisée par Jérôme Labanowski, et Leslie Mondamert, chercheurs à l'IC2MP.

1.3. Tampons et solvants utilisés pour la préparation d'échantillon

Différentes solutions ont été utilisées tout au long de l'étude. Le diméthylsulfoxyde (DMSO), C₂H₆OS (>99 %); l'hydroxyde de sodium (NaOH); le trishydroxyméthyl aminométhane-acide chlorhydrique (Tris-HCl) (99 %); le chlorure de magnésium, MgCl₂ (99 %); Hydroxide de TétraMéthylAmmonium (TMAH), C₄H₁₃NO (99 %), l'isopropanol, C₃H₈O

Partie expérimentale

(99 %), l'acide trichloroacétique $C_2HCl_3O_2$, (99,9 %), l'acétone C_3H_6O (99 %), le méthanol CH_3OH (> 99,9 %), le chloroforme $CHCl_3$ (> 99,9 %), l'éthanol C_2H_6O (99 %), 1,25 M d'acide chlorhydrique dans l'éthanol pur.

Tout le matériel utilisé dans notre étude a été fourni par Sigma Aldrich (Saint Louis, USA) et de Fisher Scientific (Illkirch, France).

1.4. Réactifs et tampons utilisés pour la digestion trypsique et l'analyse UPLC-MS/MS

Pour l'hydrolyse trypsique des protéines, nous avons utilisé des solutions de DL-dithiothréitol (DTT) (> 99 %), de l'iodoacétamide (IAA) (> 99 %) qui permettent la réduction des ponts disulfures puis l'alkylation des groupements thiols libres. Le bicarbonate d'ammonium NH_4HCO_3 (99 %) ainsi que le trishydroxyméthyl aminométhane (Tris) (> 99,8 %) ont été utilisés comme tampon. Nous avons utilisé la trypsine porcine (20 μ g) de Promega (Madison, WI, USA).

Les phases mobiles utilisées pour l'analyse UPLC-MS étaient composées d'acétonitrile et d'acide formique de qualité HPLC. De l'eau ultrapure (résistivité 18,2 M.Ohm.cm) a été également employée lors des expérimentations afin de préparer certaines solutions et rincer la verrerie.

Tout le matériel utilisé dans notre étude a été fourni par Sigma Aldrich (Saint Louis, USA) et Fisher Scientific (Illkirch, France).

1.5. Réactifs et tampons utilisés pour la spectrofluorométrie

Afin de réaliser les expériences de spectrofluormétrie, l'acide 1-anilino-8-naphthalène sulfonate (ANS) $C_{16}H_{13}NO_3S$ (≥ 97 %), l'acétate de sodium (>99 %) et l'acide acétique (≥ 99.7 %), ont été employés. Tout le matériel utilisé dans notre étude a été fourni par Sigma Aldrich (Saint Louis, USA), Fisher Scientific (Illkirch, France) et Acros-Organics (Geel, Belgium).

2. Méthodes

2.1. Analyse UPLC-MS/MS

2.1.1. Paramètres UPLC-MS/MS

2.1.1.1. Mode d'analyse direct

Partie expérimentale

L'analyse des peptides après digestion a été réalisée avec un système de chromatographie liquide type UPLC Acela 1250 pump, couplé à un spectromètre de masse hybride Quadripôle/Orbitrap type Q-Exactive, équipé d'une source à électrospray (Thermo Fisher Scientific, Waltham, MA, USA). La séparation des peptides a été effectuée sur une colonne analytique *Hypersil Gold Q* (50×2,1 mm i.d, de granulométrie de 1,9 µm, et porosité de 175 Å) fournie par Thermo Scientific. L'élution des peptides a été réalisée à 40 °C pendant 17 min, à un débit de 500 µL.min⁻¹, selon un gradient linéaire allant de 5 à 90 % de solvant A en 15 min, suivie d'un reconditionnement avec 5 % de A et 95 % de B et un retour aux conditions initiales. Le solvant A et B sont composés respectivement d'acétonitrile et d'eau ultrapure additionnés de 0,1 % d'acide formique. Le volume d'injection était de 5 µL. Les peptides élués ont été analysés en ligne en mode positif. Les paramètres d'analyse MS sont détaillés ci-après.

2.1.1.2. Système « *dual trap* »

Ce système impliquait le fonctionnement en parallèle de deux pompes quaternaires (pompes Acela 600 et 1250). La pompe 600 (Thermo Fisher Scientific) permet de charger l'échantillon dans le système de *trapping*, tandis que la pompe 1250 dirige l'élution en mode *backflush* et la séparation analytique des composés. Dans cette étude, deux colonnes de préconcentration ont été utilisées : une colonne semi-polaire *Hypersil Gold aQ* (20 x 2,1 mm, 12 µm, 175 Å; Thermo Fisher Scientific) et une colonne *Hypercarb* 100% carbone graphité poreux (20 x 2,1 mm, 7 µm, 250 Å; Thermo Fisher Scientific) permettant la rétention des composés polaires grâce à la surface polarisable du graphite. Ces colonnes ont été testées seules ou en série. Elles étaient connectées à la colonne analytique au moyen d'une vanne de commutation à 6 voies type Cheminert® (VICI AG International, Schenkton, Switzerland).

Le fonctionnement du système « *dual trap* » sera détaillé dans le chapitre I de résultats, sachant que quatre paramètres différents, ayant chacun une influence sur la rétention des peptides, ont été plus particulièrement étudiés :

- Nature des phases stationnaires des pré-colonnes (colonne *Hypersil Gold aQ* ou *Hypercarb*).
- Configuration des pré-colonnes : seules ou en séries.
- Débit de chargement entre 125 et 1000 µL.min⁻¹, et temps de chargement entre 60 et 480 sec.

Partie expérimentale

- Le temps du *backflush* : 60, 240 et 900 sec.

Au moyen d'un plan expérimental, nous avons étudié 21 configurations analytiques différentes. Elles ont été comparées suivant leur capacité à retenir les peptides par rapport à au mode d'analyse direct.

2.1.1.2.1. Analyses statistiques

Afin de comparer les 21 configurations analytiques, des analyses de la variance multivariée (MANOVAs) ont été réalisées avec un risque α de 5 % avec le logiciel XL-Stat (Addinsoft, Paris, France) :

- MANOVA (1) : Analyse MANOVA sur la reproductibilité des rendements analytiques pour les 21 configurations testées.
- MANOVA (2) : Effet de l'arrangement des colonnes de préconcentration, du débit et du temps de chargement et du temps de *backflush*.

Pour la validation de la méthode analytique, une analyse en composante principale (ACP) a été réalisée, avec pour variables continues, les rendements analytiques (de rétention) de chaque peptide standard et pour individus, les configurations chromatographiques testées.

2.1.1.2.2. Analyse quantitative

Deux droites d'étalonnage ont été réalisées pour chaque peptide. La première allant de 0,1 à 6 $\mu\text{g}\cdot\text{mL}^{-1}$ a été utilisée pour évaluer les performances du mode d'analyse direct (sans « *dual trap* »). La seconde allant de 0,0025 à 0,05 $\mu\text{g}\cdot\text{mL}^{-1}$ nous a permis d'évaluer les performances du système « *dual trap* ». Ces droites de calibration ont été obtenues par régression linéaire *via* le logiciel d'acquisition des données *Xcalibur* et sont considérées comme valides lorsque le coefficient de détermination (R^2) est supérieur à 0,9. Les limites de détection (LODs : $3.3\sigma/S$, S : pente) et de quantification (LOQs : $10\sigma/S$) ont ensuite été déterminées.

2.1.2. Analyse par spectrométrie de masse haute résolution (UPLC-HRMS/MS)

Avant chaque séquence d'analyses, le spectromètre de masse était calibré avec des composés de référence (Thermo Scientific™ Pierce™ Calibration Solutions). Les peptides étaient analysés en mode positif. Les paramètres généraux de la source électrospray étaient les suivants : le potentiel d'ionisation était de 4 kv ; la température de la source et la température

Partie expérimentale

du capillaire étaient respectivement 300 et 275 °C. Le gaz de nébulisation, le *sweep* gaz ainsi que le gaz auxiliaire (N₂) ont été optimisés respectivement à 35, 10 et 20 (unité arbitraire). Le système fonctionne avec un passage automatique entre le mode MS et le mode MS/MS au moyen du logiciel de pilotage du spectromètre de masse. L'acquisition est pilotée par le logiciel *Xcalibur* (Thermo Fisher Scientific). Elle a ici été réalisée en mode « Data Dependant Acquisition » associée à un balayage complet « Full Scan » de l'échantillon injecté avec 1 microscan.s⁻¹ et une résolution de 70,000. La gamme de masse a été fixée entre 200 et 3000 m/z. Les 4 ions (TopN = 4) les plus intenses ont été sélectionnés pour une analyse en mode fragmentation MS/MS dans la CID avec 1 microscan.s⁻¹ et une résolution de 35,000. L'énergie de collision normalisée a été fixée à 35 % et un temps d'exclusion dynamique de 10 sec a été utilisé.

2.1.3. Analyses bio-informatiques

Des recherches à partir des spectres MS² avec la banque de données SwissProt ont été réalisées pour les échantillons réels (sol, biofilm, foie de souris et xénogreffes de tumeurs). La comparaison des protocoles de purification des protéines a été effectuée *via* l'algorithme *Mascot distiller* (v. 2.6.0), développé par Matrix science. Les paramètres nécessaires suivants ont été fixés : le spectromètre « ESI-Q-Haute résolution » a été choisi comme instrument, la carbamidométhylation comme modification fixe des cystéines, la trypsine comme enzyme de digestion et deux mauvais clivages sont permis. Les tolérances d'erreur de masse pour les ions précurseurs et les ions fragments étaient respectivement de 5 ppm et 0,3 Da.

2.2. Optimisation de la digestion trypsique

Différents paramètres influençant la digestion trypsique ont été étudiés et optimisés, y compris le rapport trypsine/protéine, la solution tampon de digestion, le temps de digestion ainsi que la concentration d'acide formique à ajouter à la fin de digestion. Cette optimisation a été réalisée sur une solution standard contenant cinq peptides à une concentration de 30 µg.mL⁻¹, dont l'angiotensine, seul peptide potentiel pouvant être digéré par la trypsine. À partir de 100 µL d'échantillon, 10 µL de DTT à 15,4 mg.mL⁻¹ ont été additionnés, afin de rompre les ponts disulfures. Les échantillons ont ensuite été mis dans un rack thermostaté à 60 °C pendant 1h00. Dix microlitres d'iodoacetamide à 44,4 mg.mL⁻¹ ont été ensuite ajoutés, puis les échantillons ont été placés dans le noir pendant 45 min à température ambiante. Une fois les 45 min écoulées, 100 µL de tampon bicarbonate d'ammonium à 50 mM ou de tampon 10 Mm Tris-bicarbonate

Partie expérimentale

d'ammonium à 50 mM ont été ajoutés. Les protéines ont été ensuite digérées de 2h00 à 24h00 à 37 °C en utilisant différents ratios de trypsine/protéine (w/w) : 1/6 ; 1/20 ; 1/30 ; 1/50. La digestion trypsine a été enfin stoppée par l'acide formique à 0,4 % (Lebert et al., 2015) ou à 2,5 %. La digestion trypsique a été évaluée par l'analyse de l'angiotensine II.

2.3. Mise au point d'un protocole de purification des protéines

La purification des protéines a été réalisée à partir d'une solution standard à 35 µg.mL⁻¹. Un volume de 375 µL de solution a été utilisé pour toutes les procédures étudiées. Cent vingt-cinq microlitres de TCA à 100 % a été alors ajouté pour précipiter les protéines (concentration finale de TCA dans le tube est égale à 25 %). Le mélange a été agité au vortex puis les protéines ont été précipitées pendant une nuit à -20 °C. Les tubes ont ensuite été centrifugés à 16600 g pendant 15 min à 4 °C. Le surnageant a été éliminé délicatement afin d'éviter toute perte de protéines durant l'étape de pipetage. Le culot protéique a été repris dans 500 µL de solvant de lavage : acétone, éthanol ou éthanol/HCl. Les échantillons ont été ensuite incubés pendant 1h00 à -20 °C avant d'être centrifugés à 16600 g pendant 15 min à 4 °C. Le surnageant a été à nouveau éliminé délicatement, transféré dans un autre tube puis évaporé sous vide à l'aide du concentrateur Speedvac.

Les protéines ont enfin été solubilisées dans 100 µL de tampon d'analyse, 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris après évaporation totale du solvant de lavage, puis quantifiées soit par dosage Bradford et électrophorèse monodimensionnelle (1D), soit par UPLC-HRMS/MS.

2.3.1. Quantification des protéines par dosage de Bradford et électrophorèse monodimensionnelle (1D)

La quantité de protéines totales (40 µL de l'échantillon purifié) a été déterminée selon la méthode Bradford (DC Protein Assay kit, Bio-Rad). Pour l'électrophorèse 1D, un volume de 25 µL d'échantillon a été déposé sur un gel Novex © Tris-glycine, 10 % de polyacrylamide (Invitrogen) puis soumis à une tension constante de 70 mV. Les protéines ont ensuite été visualisées après coloration au nitrate d'argent. Brièvement, les gels ont été trempés dans une solution de fixation contenant 50 % de méthanol, 10 % d'acide acétique et 100 mM d'acétate d'ammonium avant d'être rincés à l'eau (2 fois). Pour la révélation des protéines, les gels ont d'abord été incubés pendant 15 min dans de thiosulfate de sodium à 0,005 %, puis 15 min dans le nitrate d'argent à 0,1 %. Le développement de la coloration (2 min) a été réalisé en présence

Partie expérimentale

du formaldéhyde à 0,036 % et du carbonate de sodium à 2 %. La réaction a été stoppée avec une solution d'EDTA à 50 mM. Le traitement des images obtenues après électrophorèse a été réalisé à l'aide du logiciel ImageJ.

2.3.2. Quantification des protéines par UPLC-HRMS/MS

Les protéines ont été solubilisées dans 100 μL de tampon d'analyse, puis diluées au 1/3 en utilisant le même tampon et enfin digérées avec de la trypsine. Un volume de 1 μL d'une solution à 1 $\text{mg}\cdot\text{mL}^{-1}$ de peptides standards a été ajouté. À partir de 100 μL d'échantillon, 10 μL de DTT à 15,4 $\text{mg}\cdot\text{mL}^{-1}$ ont été additionnés. Les échantillons ont été ensuite incubés à 60 °C pendant 1h00. Dix microlitres d'iodoacetamide à 44,4 $\text{mg}\cdot\text{mL}^{-1}$ ont été ajoutés, puis les échantillons ont été placés dans le noir pendant 45 min à température ambiante. Une fois les 45 min écoulées, 100 μL de 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris ont été ajoutés. Enfin, les protéines ont été digérées par la trypsine avec un ratio trypsine/protéine (w/w) égal à 1/6, avant d'être analysées par UPLC-MS/MS. La digestion a été réalisée pendant 24h00 à 37°C. L'activité de la trypsine a été stoppée par de l'acide formique à 0,4 % (concentration finale).

2.3.3. Analyses statistiques

Afin de comparer les rendements de purification à partir des solutions standards de protéines, une analyse de la variance (ANOVAs) a été réalisée suivant un risque α de 5 % avec le logiciel XL-Stat (Addinsoft, Paris, France).

2.3.4. Spectrofluorométrie

Les mesures de fluorescence ont été réalisées avec un spectromètre de type Fluoromax-4 spectrophotometer (HORIBA Scientific, Kyoto, Japan) qui permet d'obtenir des spectres de fluorescence à trois dimensions (3D). L'échantillon a été placé dans une cuve en quartz, de trajet optique 1 cm, puis introduit dans un compartiment thermostaté à 20 °C. La source lumineuse utilisée est une lampe à arc Xénon 450 watts, émettant un rayonnement polychromatique continu dans l'UV-Visible, entre 250 et 800 nm. L'intensité de fluorescence a été mesurée en réalisant un balayage de longueur d'onde d'excitation et d'émission respectivement entre 240 à 450 nm et 260 à 600 nm par pas de 5 nm. Ce type de balayage permet ainsi d'obtenir une matrice d'émission-excitation.

Partie expérimentale

Dans notre étude, l'acide 1-anilino-8-naphthalène sulfonate (ANS) a été choisi comme fluorophore extrinsèque. Toutes les expériences ont été réalisées dans le tampon acétate à 0,1 M (pH 5). Dans un premier temps, les spectres de fluorescence du tampon acétate (blanc), de l'ANS et des protéines seules ont été tracés. Par la suite, un volume de 500 μM d'ANS a été mélangé avec un volume de 20 μM de la protéine d'intérêt (figure 71).

Les spectres des différents échantillons ont ensuite été obtenus après soustraction du spectre correspondant au blanc.

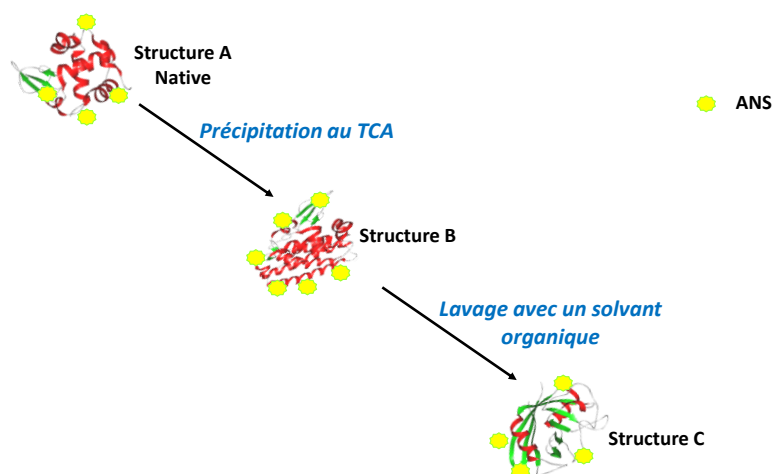


Figure 71. Traitements successifs appliqués à la protéine sous forme native

2.3.5. Micro-spectroscopie Raman

Les analyses en micro-spectroscopie Raman ont été réalisées à l'aide d'un microscope confocal HORIBA JOBIN YVON Labram HR800UV équipé d'un détecteur CCD refroidi par effet Peltier. La longueur d'onde excitatrice est de 632,8 nm est produite par un laser He-Ne (Melles Griot). La puissance laser délivrée à l'échantillon varie de 6 à 1,5 mW (utilisation de filtres de densité optique). L'appareil est équipé d'un microscope confocal Olympus qui permet de travailler en rétrodiffusion. Un réseau de diffraction à 1800 traits. mm^{-1} est utilisé et l'ouverture du trou confocal est de 200 μm . Le spectromètre est calibré avec un échantillon de silicium. Le logiciel LabSpec 5 permet l'acquisition et le traitement des résultats.

Une faible quantité d'échantillon (quelques mg) a été déposée sur une lame de verre. Les protéines natives ont été directement analysées sous forme solide. Les protéines précipitées au TCA et lavées au solvant organique (éthanol-HCl ou acétone) ont également été analysées sous

Partie expérimentale

forme solide après évaporation du solvant. Les conditions opératoires relatives à l'appareillage sont résumées dans le tableau 7.

Tableau 7. Conditions opératoires

Laser (nm)	Filtre	Trou (µm)	Réseau (g/mm)	Objectif	Gamme balayée (cm ⁻¹)	Temps d'exposition (s)	Nombre d'acquisitions	Mode d'acquisition	Filtre spike
632,8	DO ou D0,6	200	1800	50x LMP	1150-1800	180-400	4	Multiple windows	Multi (auto add)

Trois protéines à une concentration de 220 mg.mL⁻¹, la HSA, le lysozyme ainsi que la Rnase, ont été étudiées sous forme native et après traitement avec les différents réactifs selon le schéma illustré sur la figure 72.

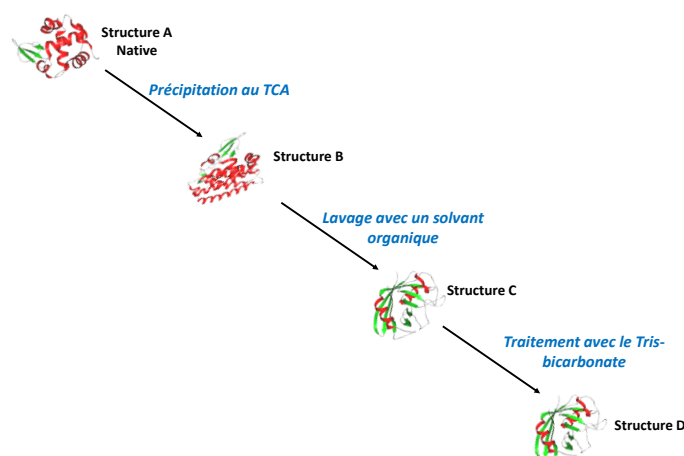


Figure 72. Traitements successifs appliqués à la protéine native

2.3.6. La microscopie à force atomique (AFM)

L'AFM a été utilisée afin d'avoir une image 3D de l'état *molten globule* induit par la précipitation au TCA. Pour cette expérience nous avons utilisé des plaques de Mica fraîchement clivées, sur lesquelles nous avons déposé 30 µL d'une solution de HSA à 10 µg.mL⁻¹ (Marchin and Berrie, 2003). La plaque a été ensuite séchée sous flux d'azote. La HSA a été analysée à l'état natif, à l'état précipité après ajout de 10 µL de TCA (25 % concentration finale) et après lavage à l'éthanol/HCl.

2.4. Comparaison des protocoles de purification sur matrices réelles

Partie expérimentale

L'efficacité de l'agent de lavage éthanol/HCl a été validée sur des matrices complexes de nature et de composition différentes. Les trois matrices ont été dopées avec une solution standard des 4 protéines modèles (HSA, lysozyme, Rnase et exoglucanase-1) à (1 mg.mL^{-1}). L'extraction des protéines a été réalisée suivant les approches publiées par Chourey et al. (2010), Huang et al. (2014) et Song et al. (2016) (figure 73).

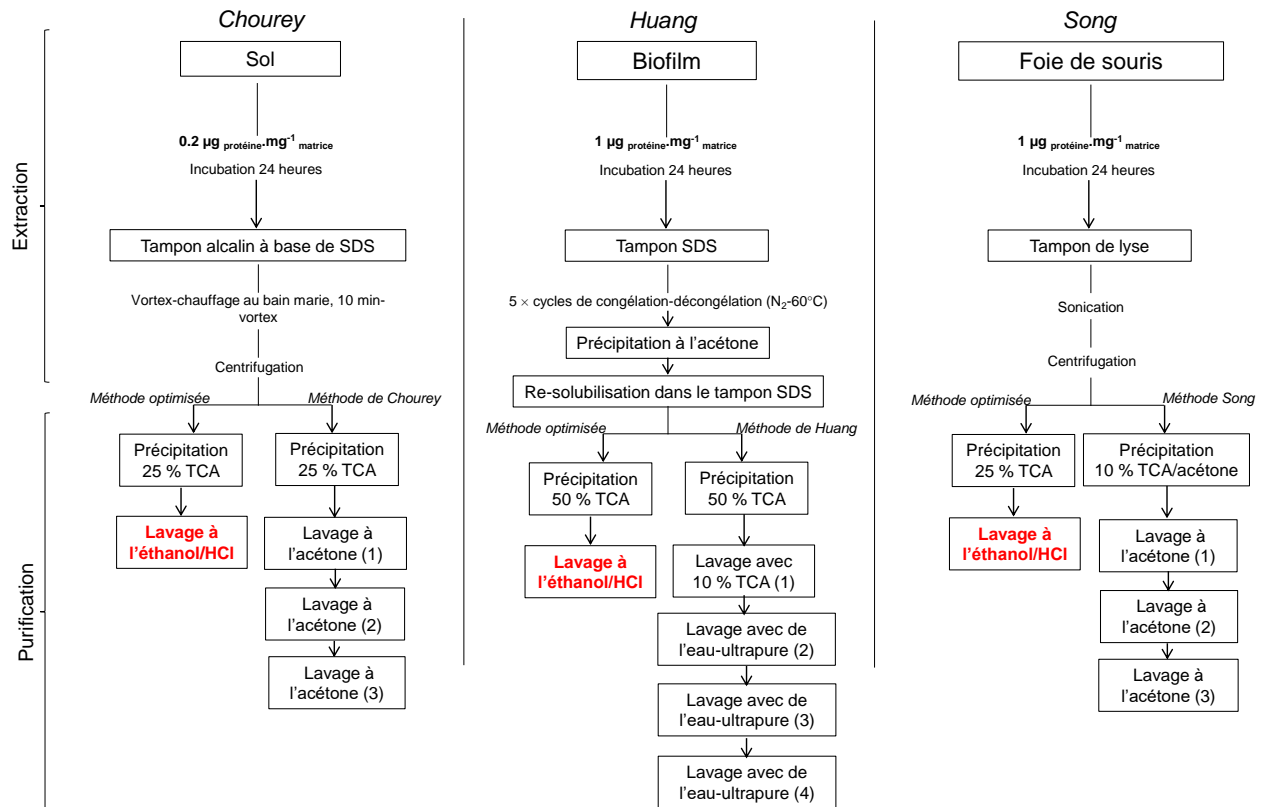


Figure 73. Comparaison de la purification des protéines à partir de matrices complexes

2.4.1. Extraction et purification des protéines du sol

Un échantillon de 500 mg de sol préalablement lyophilisé et tamisé à $500 \mu\text{m}$ a été additionné de $100 \mu\text{L}$ de la solution des 4 protéines standards. Les tubes ont été ensuite laissés pendant 24h00 à température ambiante. L'extraction des protéines a été réalisée avec un tampon de lyse contenant du SDS (5% SDS, 50 mM Tris-HCl, pH 8,5, 0,15 M NaCl, 0,1 mM EDTA, 1 mM MgCl_2 , 50 mM DTT). Le mélange sol/extractant a été alors homogénéisé au vortex pendant 2 à 3 min puis les tubes ont été immergés dans un bain marie $100 \text{ }^\circ\text{C}$ pendant 10 min pour optimiser la lyse cellulaire et inactiver ainsi les protéases libérées lors de cette opération. Les tubes ont été refroidis pendant 5 min, puis mélangés de nouveau pendant 3 min au vortex. Les solutions ont été centrifugées et le surnageant a été alors transféré dans un nouveau tube puis

Partie expérimentale

divisé en deux fractions égales soit 150 μL dans chaque tube. Les protéines ont été ensuite précipitées au TCA (25 % concentration finale) pendant une nuit à $-20\text{ }^\circ\text{C}$. Le culot protéique a été lavé soit 3 fois par l'acétone (protocole de Chourey et al., 2010), soit une seule fois par l'éthanol/HCl. Le solvant résiduel contenu dans chaque tube a été évaporé sous vide à l'aide du concentrateur Speedvac. Les culots protéiques ont été enfin solubilisés dans le tampon d'analyse, 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris (pH 8,5-8,6) après évaporation totale du solvant de lavage, diluées 1/3 en utilisant le même tampon. À partir de 100 μL d'échantillon, 15 μL de DTT à $15,4\text{ mg}\cdot\text{mL}^{-1}$ ont été additionnés. Les échantillons ont été ensuite incubés à $60\text{ }^\circ\text{C}$ pendant 1h00. 15 μL d'iodoacetamide à $44,4\text{ mg}\cdot\text{mL}^{-1}$ ont été ajoutés, puis les échantillons ont été placés dans le noir pendant 45 min à température ambiante. Une fois les 45 min écoulées, 100 μL de 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris ont été ajoutés. Les protéines ont été enfin digérées par la trypsine avec un ratio trypsine/protéine (w/w) égal à 1/6, avant d'être analysées en UPLC-MS/MS. La digestion a été réalisée pendant 24h00 à 37°C puis stoppée par de l'acide formique à 0,4 % (concentration finale).

2.4.2. Extraction et purification des protéines de biofilm

Le biofilm provenant de la vienne a été lyophilisé. Un volume de 10 μL de solution des 4 protéines standards a été ajouté à 10 mg de biofilm, puis laissé 24h00 à température ambiante. Les protéines ont été ensuite extraites suivant le protocole de Huang et al. (2014) avec 10 μL d'un tampon de lyse contenant 1 % (w/v) de SDS à pH 7,4 en présence de DTT et de PMSF (inhibiteur de protéases). L'échantillon a été ensuite soumis à 5 cycles de congélation (N_2 liquide) et décongélation ($60\text{ }^\circ\text{C}$) pour optimiser au mieux la lyse cellulaire. 500 μL d'acétone ont été alors ajoutés et les tubes ont été incubés une nuit à $-20\text{ }^\circ\text{C}$. Le culot a été ensuite solubilisé dans le même tampon de lyse (SDS), puis séparé en deux fractions égales, soit 150 μL dans chaque tube. Les protéines ont été précipitées par une solution de TCA à 50 % pendant 15 min à $4\text{ }^\circ\text{C}$, avant d'être centrifugées. Le culot protéique a ensuite été lavé soit suivant la méthode de Huang et al. (2014), à savoir un lavage au TCA (10 %) suivi de 3 lavages à l'eau, soit avec un lavage à l'éthanol-HCl. Nous avons également évalué l'effet de plusieurs lavages à l'éthanol/HCl (entre 2 et 3) sur la récupération des cibles protéiques. Le solvant résiduel contenu dans chaque tube a été évaporé sous vide à l'aide du concentrateur Speedvac. Les culots protéiques ont été enfin solubilisés dans le tampon d'analyse, 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris (pH 8,5-8,6) après évaporation totale du solvant de lavage, diluées 1/3 en utilisant le même tampon et digérées par la trypsine. Un volume de 1 μL

Partie expérimentale

d'une solution à 1 mg.mL⁻¹ de peptides standards a été ajouté. À partir de 100 µL d'échantillon, 15 µL de DTT à 15,4 mg.mL⁻¹ ont été additionnés. Les échantillons ont été ensuite incubés à 60 °C pendant 1h00. 15 µL d'iodoacetamide à 44,4 mg.mL⁻¹ ont été ajoutés, puis les échantillons ont été placés dans le noir pendant 45min à température ambiante. Une fois les 45 min écoulées, 100 µL de 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris ont été ajoutés. Les protéines ont été digérées avec 50 µL de trypsine 0,1 µg.µL⁻¹ (la trypsine a été additionnée en excès dans cette matrice complexe afin d'avoir une digestion optimale), avant d'être analysées en UPLC-MS/MS. La digestion a été réalisée pendant 24h00 à 37°C puis stoppée par de l'acide formique à 0,4 % (concentration finale).

2.4.3. Extraction et purification des protéines du foie de souris

L'échantillon de foie a été d'abord broyé à l'aide d'un mortier en présence d'azote liquide. La poudre obtenue a été ensuite réparties en quantité égale dans trois tubes eppendorfs de manière à avoir 15 mg d'échantillon dans chaque tube (3 réplicats). Un volume de 15 µL de solution des 4 protéines standards (1 mg.mL⁻¹) a été ensuite additionné. Les protéines ont été incubées en présence de la matrice pendant 24h00 à -20 °C. Les protéines ont été extraites selon le protocole décrit par (Nie et al., 2013; Song et al., 2016). Brièvement, 1 mL de solution d'extraction de lyse contenant (7 M urea, 2 M thiourea, 4% CHAPS, 40 mM Tris-HCl, pH 8.5) contenant 1 mM PMSF et 2 mM EDTA a été additionné. Après 5 min 10 mM de DTT a été alors ajouté (concentration finale). Le tissu a été homogénéisé par ultrasons pendant 15 min dans un bain de glace, puis centrifugé pendant à 16600 g pendant 30 min à 4 ° C. Le surnageant contenant le lysat a été ensuite transféré dans un nouveau tube puis divisé en deux fractions égales, soit 200 µL dans chaque tube. Les protéines ont été alors précipitées soit par TCA (25 % concentration finale), soit par 5 fois de volume d'une solution de 10 % de TCA/acétone (v/v), une nuit à -20 °C avant d'être centrifugés. Le surnageant a été éliminé délicatement et les protéines ont été purifiées soit une fois par d'éthanol/HCl pendant 1 h à -20 °C, soit 3 fois par une solution d'acétone froid (protocole de Song et al., 2016). Le solvant résiduel contenu dans chaque tube a été évaporé sous vide à l'aide du concentrateur Speedvac. Les culots protéiques ont été solubilisés dans le tampon d'analyse, 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris (pH 8,5-8,6) après évaporation totale du solvant de lavage puis dilués 1/3 en utilisant le même tampon. Un volume de 1 µL d'une solution à 1 mg.mL⁻¹ de peptides standards a été ajouté. Les protéines ont été digérées selon les mêmes conditions expérimentales décrites dans le paragraphe 2.4.2.

Partie expérimentale

2.5. Analyse quantitative

2.5.1. Etalonnage externe

Afin de déterminer les quantités des 4 protéines dans les échantillons après précipitation au TCA et lavage à l'acétone, éthanol, éthanol/HCl, deux gammes étalons ont été réalisées pour chaque peptide de la HSA, lysozyme, Rnase et exoglucanase-1, allant de 6-42 $\mu\text{g.mL}^{-1}$ et de 20 à 80 ng.mL^{-1} . Les droites d'étalonnage représentant en ordonnée le ratio des surfaces du pic de l'ion peptide le plus intense par rapport à l'aire du pic de l'ion peptide standard (leu-enképhaline) et en abscisses les concentrations massiques des peptides/protéines, ont été obtenues par régression linéaire *via* le logiciel d'acquisition des données *Xcalibur* et ont été considérées comme valides lorsque le coefficient de détermination (R^2) est supérieur à 0,9. Les limites de détection (LODs) et de quantification (LOQs) ont ensuite été déterminées.

2.5.2. Quantification relative par le logiciel Skyline des protéines issues d'échantillons complexes

Au cours de ce travail, j'ai été amenée à utiliser le logiciel *Skyline* (version 3.7.0.10940 MacCoss Lab) (MacLean et al., 2010) en libre accès pour l'analyse et le traitement des données MS/MS en mode DDA sur le Q-exactive. *Skyline* extrait pour chaque transition (couple ion précurseur/ion fragment) le courant ionique et intègre ensuite l'aire sous le pic chromatographique correspondant à l'ion précurseur associé à cette transition (Egertson et al., 2015; MacLean et al., 2010).

Ainsi, à partir des matrices complexes (sol, biofilm et foie de souris) une quantification relative des protéines endogènes et standards (dopées) a été réalisée, en fonction des spectres de fragmentation MS2.

2.6. Etudes des paramètres influençant l'extraction des protéines

Avant extraction, les matrices ont été dopées avec un mélange de protéines standards (HSA, lysozyme, Rnase et exoglucanase-1) à raison de 0,1 mg/100 mg _{matrice} (0,1 %).

De nombreux essais d'extraction de protéines à partir d'échantillons simples et complexes ont été réalisés dans le but d'étudier les paramètres influençant l'extraction des protéines. Ces essais ont été menés avec différents solvants organiques ou aqueux, des mélanges de solutions aqueuses et organiques (eau/méthanol, eau/isopropanol, DMSO/Tris-HCl, etc.) dans différentes

Partie expérimentale

conditions (temps, température), assistées ou non par des techniques physiques (ultrasons focalisés, agitation), avec ou sans prétraitement au préalable.

2.6.1. Extraction sur une matrice simple : sable de Fontainebleau dopé

2.6.1.1. Effet de la filtration

Les essais préliminaires ont été menés sur une matrice simple, du sable de Fontainebleau. Il a été dopé avec des volumes importants en solvant, soit 10 mL de solvant pour 1 g de sable dopé. La matrice a été ensuite laissée pendant 5 jours à température ambiante. L'extraction a été alors réalisée avec 10 mL d'eau ultrapure ou avec un mélange d'eau ultrapure/méthanol (50/50 : v/v). Les protéines ont été extraites pendant 2h00 sous agitation magnétique à température ambiante. Puis la solution a été filtrée sous vide à 0,45 µm afin d'éliminer le résidu solide de l'extrait. Une fraction du filtrat soit 500 µL a été prélevée puis évaporée à sec au SpeedVac, ensuite purifiée au TCA et à l'éthanol/HCl. L'expérience a également été réalisée sans filtration. Les protéines ont été solubilisées dans 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris, diluées 1/3 en utilisant le même tampon. Un volume de 1 µL d'une solution à 1 mg.mL⁻¹ de peptides standards a été ajouté et les protéines ont été digérées par la trypsine (ratio 1/6). La digestion trypsine a été stoppée par l'acide formique (0,4 % concentration finale).

2.6.1.2. Effet du solvant

L'extraction a été réalisée avec trois solvants différents : le DMSO (99 %), l'eau ultrapure/isopropanol (50/50 : v/v), et le Tris-HCl 25mM (pH 7,5). Une extraction séquentielle a été réalisée afin d'apprécier le rôle de chaque agent sur la récupération des protéines. Pour cela trois combinaisons de solvants ont été réalisées :

- Extraction au DMSO (99 %), suivie d'une extraction avec un mélange eau ultrapure/isopropanol (50/50 : v/v).
- Extraction avec un mélange d'eau ultrapure/isopropanol (50/50 : v/v), suivie d'une extraction avec le Tris-HCl 25mM (pH 7,5).
- Extraction au Tris-HCl 25mM (pH 7,5), suivie d'une extraction au DMSO (99 %).

Partie expérimentale

L'extraction a été réalisée sur 100 mg de sable dopé et auquel un volume de 1 mL de chaque solvant a été ajouté. Les protéines ont été extraites dans un agitateur thermostaté, où la température et l'agitation étaient parfaitement maîtrisées, pendant 2h00 à 25 °C à 2000 rpm avant d'être centrifugées. Une fraction de l'extrait, soit 10 µL, a été récupérée ; le volume a été ajusté à 100 µL avec le tampon 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris (pH 8,5-8,6). Un volume de 1 µL d'une solution à 1 mg.mL⁻¹ de peptides standards a été ajouté et les protéines ont été ensuite digérées par la trypsine (ratio 1/6) avant d'être analysées UPLC-MS.

Nous avons également évalué l'effet du mélange Tris-HCl 25mM (pH 7,5) et DMSO (99 %) sur l'extraction des protéines.

2.6.2. Extraction sur une matrice complexe : Sol du site hydrogéologique du Deffend dopé

L'extraction sur matrice complexe a été réalisée à partir de 50 mg de sol dopé (5 jours adsorption à température ambiante), préalablement lyophilisé et tamisé à 500 µm, auquel nous avons additionné 500 µL de solvant d'extraction.

Plusieurs paramètres ont été testés : température d'agitation, temps d'agitation, méthode d'extraction : extraction classique dans l'agitateur thermostaté ou extraction aux ultrasons focalisés (FULSE), effet de la solution d'extraction.

2.6.2.1. Effet de la température d'extraction

Un volume de 500 µL d'une solution d'extraction de Tris-HCl 25mM (pH 7,5) / DMSO (99 %) a été ajouté à 50 mg de sol dopé. Les protéines ont été extraites par agitation pendant 5h00 à 2000 rpm, à 25, 35, 45 et 55 °C, avant d'être centrifugées.

Une fraction, soit 30 µL, a été récupérée, et le volume a été ajusté à 100 µL avec le tampon 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris (pH 8,5-8,6). Les protéines ont été digérées par la trypsine (ratio 1/6) avant d'être analysées UPLC-HRMS/MS.

Partie expérimentale

2.6.2.2. Effet du temps d'agitation

500 μL d'une solution d'extraction de Tris-HCl 25mM (pH 7,5) / DMSO (99 %) ont été ajoutés à 50 mg de sol dopé. Les protéines ont été extraites par agitation thermostaté à 25 °C, 2000 rpm, pendant 2h00, 3h00, 4h00 ou 5h00 avant d'être centrifugées.

Une fraction, soit 30 μL , a été récupérée. Le volume a été ajusté à 100 μL avec le tampon 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris (pH 8,5-8,6). Les protéines ont été ensuite digérées par la trypsine (ratio 1/6) avant d'être analysées UPLC-MS.

2.6.2.3. Effet de la méthode d'extraction

Deux méthodes d'extraction ont été testées :

- Une extraction dans l'agitateur thermostaté à 25 °C, 2000 rpm, pendant 2h00.
- Une extraction assistée par des ultrasons : temps de sonication 5 min, amplitude de 80 %, temps de pulse et relaxation de 1 sec (FULSE-5) (paramètres optimisés durant le post-doctorat de Pascaline François).

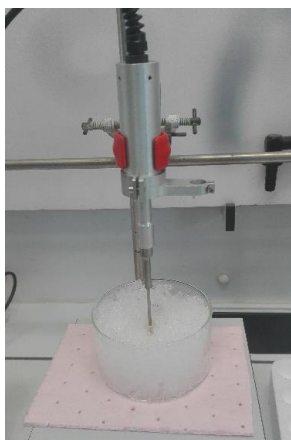


Figure 74. Photo du sonificateur utilisé lors de cette étude (sonde de diamètre de 2 mm)

L'extraction a été réalisée avec 500 μL d'une solution alcaline (DMSO/NaOH). Les protéines ont d'abord été extraites par agitation à 2000 rpm pendant 2h00 à 25 °C, avant d'être centrifugées. Le premier extrait (agitation 2h00, 25 °C) a alors été récupéré pour être analysé par UPLC-HRMS/MS. Une deuxième extraction assistée par ultrasons a été réalisée sur la même matrice. Cette fois-ci, un volume de 500 μL de la même solution a été additionné et les protéines ont ensuite été extraites par sonication (FULSE-5). Le deuxième extrait (FULSE-5) a également analysé par UPLC-HRMS/MS.

Partie expérimentale

2.6.2.4. Effet de la soude

À partir de 50 mg de sol dopé, un volume de 500 µL d'un mélange de solvant contenant NaOH (0,1 M)/DMSO (99 %) /Tris-HCl (25 mM) (v/v/v), a été additionné. Les protéines ont été extraites soit dans l'agitateur thermostaté pendant 2h00, à 25 °C, avant d'être centrifugées à 2000 rpm. Une fraction, soit 30 µL, a été récupérée ; le volume est ajusté à 100 µL avec le tampon 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris (pH 8,5-8,6). Un volume de 1 µL d'une solution à 1 mg.mL⁻¹ de peptides standards a été ajouté. Les protéines ont été ensuite digérées par la trypsine (ratio 1/6) avant d'être analysées UPLC-MS.

2.6.2.5. Effet de nouveaux agents chimiques et biochimiques sur l'extraction des protéines

Trois agents chimiques et biochimiques ont été testés :

- L'hydroxyde de tetraméthylammonium (TMAH) à 66 mM suivie d'une extraction au DMSO.
- Une solution d'acides aminés basiques 1 % (w/v) à pH 5 suivie d'une extraction au DMSO.
- Le MgCl₂.

Brièvement, 50 mg de sol dopé, laissé à température ambiante pendant 5 jours ont été traités soit avec une solution d'acides aminés basiques (1% (w/v)) à pH 5, soit au TMAH (66 mM, soit à l'eau ultrapure (témoin, sans traitement), soit avec 600 mM de MgCl₂. La solution a été alors homogénéisée sous une agitation douce de 300 rpm pendant toute une nuit à 37 °C (Nicora et al., 2013). Les protéines ont été alors extraites au DMSO (99 %) pendant une nuit dans les mêmes conditions opératoires. Une seconde extraction assistée par ultrasons a été également réalisée (FULSE-5).

Les protéines extraites ont été récupérées après centrifugation. Un volume 30 µL de chaque extrait a été prélevé, le volume a été ensuite ajusté à 100 µL avec le tampon 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris (pH 8,5-8,6). Un volume de 1 µL d'une solution à 1 mg.mL⁻¹ de peptides standards a été ajouté. Les protéines ont alors été digérées par la trypsine (ratio 1/6) avant d'être analysées UPLC-MS/MS.

Partie expérimentale

2.6.2.6. Extraction séquentielle : MgCl₂ - acides aminés/DMSO - TMAH/DMSO

L'extraction des protéines cibles a été réalisée en commençant par un traitement au MgCl₂ à 600 mM, suivi d'une extraction aux acides aminés/DMSO et puis d'un traitement TMAH/DMSO. La vitesse d'agitation a été augmentée à 1000 rpm au lieu de 500 rpm car afin d'optimiser au mieux l'extraction ainsi que la solubilisation des protéines à partir de la matrice sol. Dans un deuxième temps, une extraction séquentielle au TMAH-DMSO suivie d'une extraction au MgCl₂ et d'un traitement aux acides aminés/DMSO a été réalisée.

2.6.3. Extraction à partir d'une matrice biologique : xénogreffe de tumeur

L'extraction des protéines a été réalisée sur une petite tumeur (m= 88 mg) d'une souris C57BL/6 porteuse d'une tumeur du carcinome pulmonaire (xénogreffe). Un volume de 100 µL de tampon phosphate salin (PBS 1X : (1X : 137 mM NaCl ; 2.7 mM KCl ; 10 mM Na₂HPO₄ ; 1.7 mM KH₂HPO₄, pH 7,6) a alors été ajouté. La solution a été agitée à une vitesse réduite soit 300 rpm à 37 °C pendant 1h00 avant d'être centrifugée à 2000 rpm, 4°C pendant 2 min. La fraction soluble a été ensuite centrifugée une deuxième fois à grande vitesse, à 13200 rpm pendant 20 min, afin d'éliminer tous les débris cellulaires (Fijneman et al., 2012). Un volume de 60 µL a été alors prélevé puis purifié au TCA (25 % concentration finale) puis à l'éthanol/HCl. Les culots protéiques ont été enfin solubilisés dans le tampon d'analyse, 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris (pH 8,5-8,6) après évaporation totale du solvant de lavage. Un volume de 1 µL d'une solution à 1 mg.mL⁻¹ de peptides standards a été ajouté. À partir de 100 µL d'échantillon, 30 µL de DTT à 15,4 mg.mL⁻¹ ont été additionnés, afin de rompre les ponts disulfures. Les échantillons ont été ensuite incubés à 60 °C pendant 1h00. 30 µL d'iodoacetamide à 44,4 mg.mL⁻¹ ont été ajoutés, puis les échantillons ont été placés dans le noir pendant 45 min à température ambiante. Une fois les 45 min écoulées, 100 µL de 50 mM de bicarbonate d'ammonium contenant 10 mM de Tris ont été ajoutés. Enfin, les protéines ont été digérées avec 90 µL de trypsine à 0,1 µg.µL⁻¹ (la trypsine a été additionnée en excès dans cette matrice complexe afin d'avoir une digestion optimale), avant d'être analysées UPLC-MS/MS. La digestion a été réalisée pendant 24h00 à 37°C puis stoppée par de l'acide formique à 0,4 % (concentration finale).

Une extraction séquentielle (MgCl₂ - acides aminés/DMSO - TMAH/DMSO) a également été réalisée sur une tumeur de la même souris. Les protéines ont alors été purifiées et digérées

Partie expérimentale

à la trypsine suivant la méthode décrite ci-dessous, avant d'être analysées par UPLC-HRMS/MS.

Résultats et discussions

Chapitre I : Développement analytique

Chapitre I : Développement analytique

Une analyse protéomique non ciblée a pour objectif d'identifier et de quantifier un maximum de protéines. En raison de la complexité des échantillons biologiques et environnementaux, une infusion directe en spectrométrie de masse n'est pas suffisante pour analyser les milliers de peptides générés après digestion trypsique. En effet, l'ionisation des peptides peut être limitée lorsqu'elle est confrontée à la présence de molécules de nature non peptidiques. Celles-ci en effet peuvent s'ioniser et complexifier le spectre de masse et ainsi diminuer la détection des composés cibles. Afin de pallier à ce problème, la chromatographie liquide est généralement couplée à la spectrométrie de masse permettant ainsi une séparation puis une identification précise et sensible du « pool » de peptides générés après digestion. Grâce aux développements accomplis en HPLC, les nouvelles techniques séparatives telles que l'UPLC ont permis d'augmenter de manière considérable le pouvoir résolutif et d'améliorer la sensibilité de détection. Actuellement, l'enjeu du développement analytique est d'accroître l'efficacité de séparation, mais également d'amplifier le signal de détection de composés d'intérêt exprimés à l'état de traces.

Actuellement, la mise en œuvre d'approches tels que les systèmes de séparation multidimensionnelle associant plusieurs colonnes chromatographiques (orthogonalité) a permis de réduire de manière significative la complexité des mélanges peptidiques et d'améliorer leur séparation (Valeja et al., 2015; Wu et al., 2012). Cette technique est particulièrement adaptée à l'analyse de composés (métabolites polaires/apolaires, peptides, etc.) présents dans des milieux complexes (Wang et al., 2015; Zhou et al., 2013; Stephan et al., 2016; Wang et al., 2017). Ces systèmes offrent en effet la possibilité d'obtenir une séparation orthogonale et d'augmenter la rétention du nombre et/ou de la quantité d'une classe chimique de composés (Magdeldin et al., 2014b; Nestola et al., 2015). Cependant, la chromatographie multidimensionnelle se heurte à des problèmes techniques, liés à une augmentation du volume mort au niveau des interfaces (vannes, joints de raccordements entre les colonnes, etc.) et à la différence de nature des phases éluantes qui sont rarement compatibles (Koppen et al., 2014).

Le but du développement analytique ici proposé est d'obtenir une détection et une quantification sensible des cibles peptidiques.

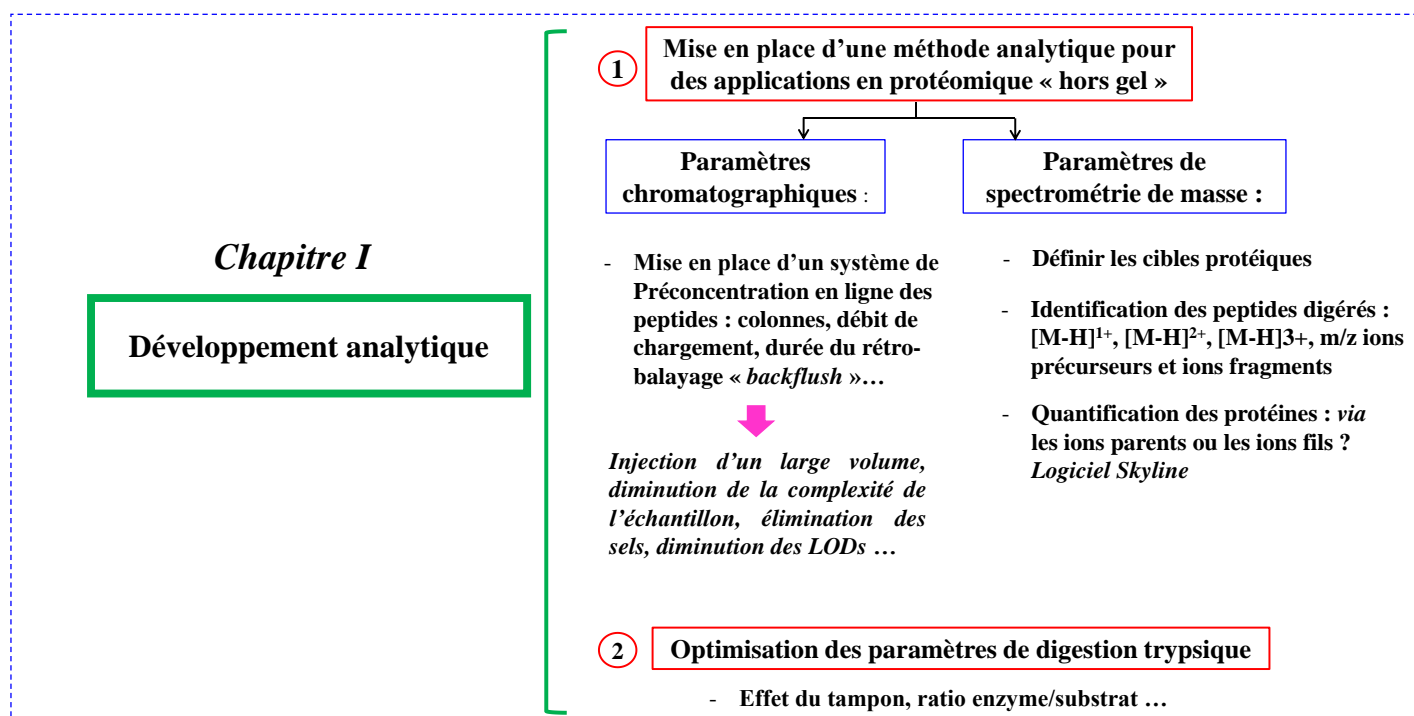


Figure 75. Stratégie expérimentale

1. Optimisation des paramètres chromatographiques : mise en place d'un système de préconcentration en ligne

Actuellement, l'association de 2 (voire 3 ou 4) systèmes chromatographiques est en plein essor. Ces systèmes multidimensionnels formés de plusieurs pompes (1 à 3) associent généralement des colonnes analytiques de phases stationnaires différentes. Ils peuvent être également constitués d'une unité de *trapping*, reliée à une colonne analytique (Minshull et al., 2016).

C'est cette seconde unité que nous avons optimisée dans le cadre de ce travail. En effet, l'idée était d'augmenter le signal des peptides, présents pour certains d'entre eux à de faibles concentrations dans des échantillons complexes. Il s'agissait également de réduire la complexité (élimination des sels, solvants, détergents utilisés pour la préparation d'échantillon, molécules endogènes non protéiques, etc.) sans passer par des étapes off-line supplémentaires (fractionnement ou de purification (SPE), cartouche de filtration, etc.) qui entraînent dans la plupart des cas une perte des composés d'intérêt (Chiesl et al., 2009; Luong et al., 2012). En outre, l'avantage d'un système de préconcentration ou *trapping* réside dans l'injection d'un large volume (ILV) d'échantillon (entre 100 et 5000 μL) ce qui permet de concentrer au maximum les composés tout en améliorant la reproductibilité de l'analyse et en limitant les

Chapitre I : Développement analytique

sources potentielles de contamination de l'échantillon (Boix et al., 2015; Koppen et al., 2014; Lim et al., 2007).

Dans ce chapitre, nous avons développé un système chromatographique, appelé « *dual trap LC unit* » capable de :

- 1) Purifier en ligne l'échantillon afin d'éliminer les impuretés et les interférents non ciblés.
- 2) Amplifier le signal des peptides cibles qui peut être masqué par des peptides plus abondants.
- 3) Séparer et détecter simultanément les peptides des protéines dopées et/ou endogènes à la matrice.

La figure 76 présente le fonctionnement général du système chromatographique développé dans cette étude.

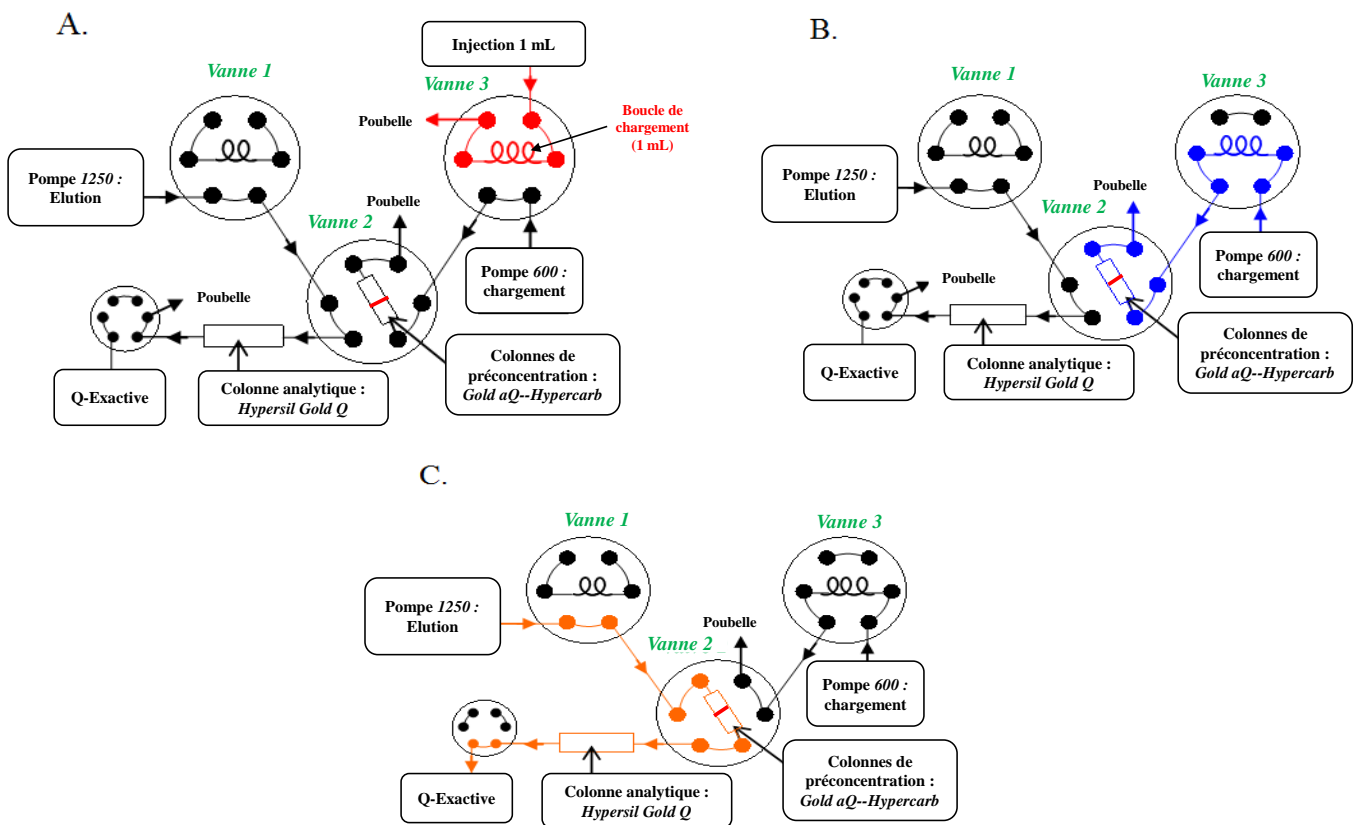


Figure 76. Représentation schématique du système chromatographique ; A. Injection, B.

Préconcentration *trapping*, C. Elution

Chapitre I : Développement analytique

L'analyse par ce système se divise en plusieurs séquences :

- **Injection (A).** Avant l'injection l'unité analytique est découplée, le passeur d'échantillons injecte 1000 μl de l'échantillon dans la boucle d'injection, ce qui déclenche le fonctionnement des deux pompes : la « pompe 600 » équilibre les colonnes de *trapping* avec de l'eau acidifiée (0,1 % acide formique) et la « pompe 1250 » équilibre la colonne analytique selon les conditions précisées dans le paragraphe suivant.

- **Chargement (B).** Les molécules sont chargées en tête des pré-colonnes en basculant la vanne V3.

- **Déchargement (C).** La vanne V2 change de position. L'élution des composés est réalisée par la phase mobile de la pompe 1250 pendant 2 min à 500 $\mu\text{L}\cdot\text{min}^{-1}$. L'élution se fait en sens inverse, on parle alors du mode rétro-balayage ou *backflush*. Durant ce temps, les pré-colonnes ainsi que la colonne analytique sont connectées en série. Les analytes sont séparés puis analysés par spectrométrie de masse.

- **Remise en conditions initiales.** Les pré-colonnes sont alors découplées puis reconditionnées pendant 2 min à 500 $\mu\text{L}\cdot\text{min}^{-1}$ avec une solution d'eau acidifiée (0,1 % acide formique).

Pour cette étude, quatre paramètres analytiques, jouant un rôle important à la fois sur le taux de recouvrement des oligopeptides, ont été étudiés (Magdeldin et al., 2014b; Motoyama et al., 2006; Nestola et al., 2015; Taylor et al., 2009) :

- Nature des phases stationnaires des pré-colonnes (colonne *Hypersil Gold aQ* (semi-polaire) ou *Hypercarb* (polaire)).
- Configuration des pré-colonnes : seules ou en séries
- Débit de chargement entre 125 et 1000 $\mu\text{L}\cdot\text{min}^{-1}$, et temps de chargement entre 60 et 480 s (figure 76.B). Le système de la boucle d'injection permet d'avoir un volume injecté constant de 1 mL.
- Le temps du *backflush* (figure 76.C).

L'optimisation des paramètres chromatographiques de préconcentration a été réalisée sur une solution modèle de cinq peptides standards caractérisés par des poids moléculaires allant

Chapitre I : Développement analytique

de 238 à 1050 g.mol⁻¹, composés de 2 à 8 acides aminés, une polarité (indices d'hydrophathie) variant entre -0,85 et 2,37 (tableau 8).

Tableau 8. Caractéristiques des cibles peptidiques

Peptides	Acides aminés	M (g.mol ⁻¹)	m/z
1. glycine-tyrosine	2	238,10	239,1063 [M + H] ⁺
2. valine-tyrosine-valine	3	379,22	380,2175 [M + H] ⁺
3. méthionine-enképhaline	5	573,23	574,2330 [M + H] ⁺
4. angiotensine II	8	1045,54	523,7745 [M + 2H] ²⁺
5. leucine-enképhaline	5	555,26	556,2766 [M + H] ⁺

Au moyen d'un plan expérimental, nous avons étudié 21 configurations analytiques différentes. Elles ont été comparées en fonction de leur capacité à retenir les oligopeptides par rapport à une méthode sans préconcentration. Les résultats sont présentés dans le tableau 9.

Tableau 9. Analyse multiple de la variance (MANOVA - n = 3 - niveau de confiance à 95%) en fonction du taux de recouvrement des oligopeptides dans les 21 configurations LC.

Chapitre I : Développement analytique

Peptides	Taux de recouvrement (%)	MANOVA (1)			MANOVA (2)									
		Recouvrement ^a	Reproductibilité ^b	Colonnes de <i>trapping</i> ^c			Paramètres de chargement ^d					Temps de <i>backflush</i> ^e (s)		
				<i>Gold aQ</i>	<i>Hypercarb</i>	Colonnes en série	125-480	250-240	375-180	500-120	1000-60	900	240	60
Glycine-tyrosine	[0-142]	< 0,0001	> 0,05	C	A	B	3	2	2	1	2,3	b	a	b
Valine-tyrosine-valine	[24-123]	< 0,0001	> 0,05	A	B	A	2	1,2	2	1	3		*	
Leucine-enképhaline	[5-97]	< 0,0001	> 0,05	B	C	A	*					a	a	b
Méthionine-enképhaline	[3-93]	< 0,0001	> 0,05	B	C	A	*					a,b	a	b
Angiotensine II	[3-118]	< 0,0001	> 0,05	A	C	B	1	1	1	1	2	a,b	a	b

MANOVA (1) : Analyse MANOVA sur la variabilité des taux de recouvrement et de la reproductibilité entre les 21 configurations testées.

MANOVA (2) : Effet de l'arrangement des colonnes de préconcentration, du débit et du temps de chargement et du temps du *backflush*. a : valeurs de *p-value* relatives à la variation du taux de recouvrement des oligopeptides entre les différentes configurations ; b : valeurs de *p-value* associées à la variation du taux de recouvrement entre les 3 répétitions ; c : A, B, C : Comparaison basée sur l'analyse de la plus petite différence significative (LSD) ($\alpha < 5\%$) pour un peptide donné. Les arrangements des colonnes de préconcentration notés avec la même lettre ne sont pas significativement différents ; d : 1, 2, 3 : résultats LSD ($\alpha < 5\%$) pour un peptide donné. Les paramètres de chargements (débit et temps) notés avec la même lettre ne sont significativement pas différents ; e : résultats LSD ($\alpha < 5\%$) pour un peptide donné. Les temps de *backflush* notés avec la même lettre ne sont pas significativement différents. * Absence de significativité ($p\text{-value} \geq 0,05$)

Chapitre I : Développement analytique

Les résultats de la MANOVA montrent que les rendements (aire obtenue du peptide en « *dual trap* » par rapport à l'aire sans « *dual trap* ») sont reproductibles puisque les 3 répétitions donnent des résultats similaires ($p\text{-value} > 0,05$) et ce quelle que soit la configuration analytique étudiée. Par ailleurs, nous pouvons constater que tous les paramètres testés ont un effet significatif sur les rendements analytiques (par exemple entre 5 et 97 % pour la leu-enképhaline). Il est à noter que les rendements de certains peptides notamment les plus petits (gly-tyr et val-tyr-val), sont supérieurs à 100 %. Ceci montre clairement l'intérêt de l'unité de préconcentration, en particulier pour retenir les molécules très polaires. En ce qui concerne l'effet des différents paramètres de préconcentration, nous pouvons remarquer que la nature de la phase stationnaire des colonnes de *trapping* a un effet significatif sur les rendements analytiques des cibles peptidiques. L'utilisation des colonnes en série permet de retenir globalement les 5 peptides cibles. Un débit de chargement de 500 $\mu\text{L}\cdot\text{min}^{-1}$ est favorable pour la rétention des analytes, de même, qu'une durée *backflush* de 240 (s).

Afin de valider la méthode analytique, une analyse en composantes principales (ACP) a été réalisée avec pour variables continues les rendements analytiques (de rétention) de chaque peptide standard et pour individus les différentes configurations chromatographiques testées.

Le cercle des corrélations (figure 77.a) permet de définir les axes factoriels (PC). Ces derniers permettent ensuite de répartir les configurations analytiques suivant les rendements obtenus. Les deux premiers axes factoriels PC1 et PC2 expliquent environ 78 % de l'information totale. L'axe PC1 est corrélé positivement aux rendements des peptides ayant un poids moléculaire intermédiaire (val-tyr-val) ou élevé (met-enképhaline et leu-enképhaline, l'angiotensine II). Tandis que, le rendement du dipeptide polaire (gly-tyr) est corrélé positivement à l'axe PC2.

Le graphique des individus (figure 77.b) montre que les configurations associant les deux colonnes (nommées « *series* ») sont corrélées positivement avec l'axe PC1. Ces configurations sont donc caractérisées par des rendements en leu-enképhaline, met-enképhaline, val-tyr-val et angiotensine II élevés. Les configurations impliquant une colonne *Hypersil Gold aQ* permettent la rétention des peptides apolaires et semi-polaires (angiotensine II et val-tyr-val). À l'opposé, les configurations impliquant la colonne *Hypercarb* sont favorables pour la rétention des peptides de faibles poids moléculaires (gly-tyr).

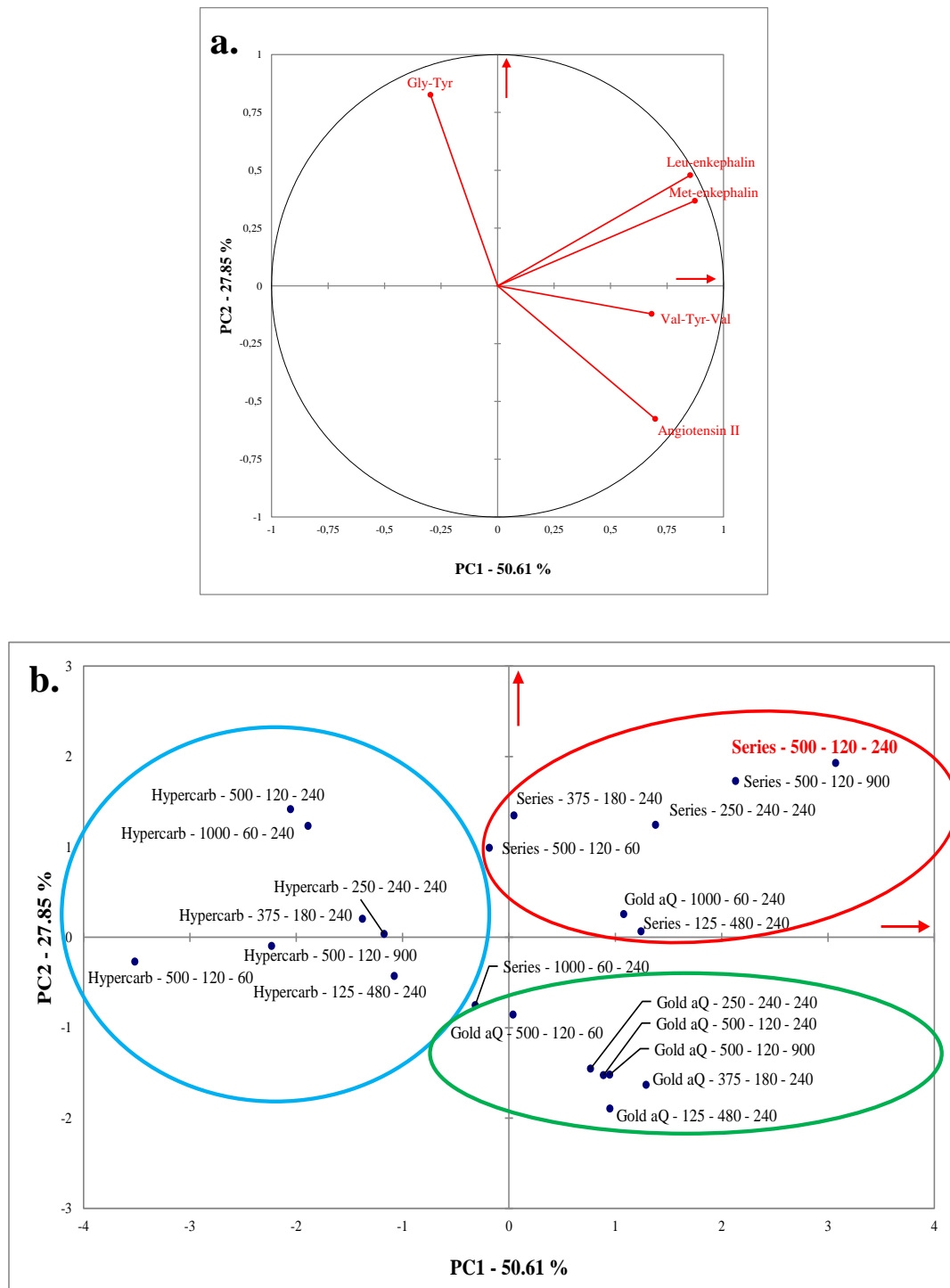


Figure 77. Analyse en Composantes Principales (ACP) basée sur le taux de recouvrement des peptides (ratio d'aire du peptide après préconcentration par rapport à l'aire du peptide sans préconcentration (injection directe)) ; a. Cercle des corrélations ; b. Graphique des observations. Codes des différentes configurations : nom de la colonne-débit de chargement-temps de chargement-temps du *backflush*. Le terme « *series* » désigne l'arrangement en série des colonnes de préconcentration, *Hypersil Gold aQ* (1) et *Hypercarb* (2) connectées à la colonne analytique *Hypersil Gold Q*

Chapitre I : Développement analytique

En outre, le graphique des individus (Figure 77.b), met en évidence que les configurations sont essentiellement réparties suivant l'arrangement des colonnes (*Hypercarb* seule, groupe bleu ; *Hypersil Gold aQ*, groupe vert ; *Hypercarb/Hypersil Gold aQ*, groupe rouge). Le débit, le temps de chargement et de la durée du *backflush*, n'ont donc qu'un effet minime au regard des phases stationnaires utilisées. Suivant ce constat, il semblerait que la configuration optimale corresponde à une configuration orthogonale en série (*Hypersil Gold aQ* puis *Hypercarb*), un débit de chargement de $500 \mu\text{L}\cdot\text{min}^{-1}$, durant 120 secondes et un temps de *backflush* de 240 (s).

L'association des deux colonnes chromatographiques (colonne semi-polaire et carbone graphite) permet ainsi la rétention de l'ensemble des peptides polaires, semi-polaires et apolaires (Lewandrowski and Sickmann, 2010). Une association similaire de phases stationnaires (*Hypercarb*/colonne C_{18}) a été appliquée pour la séparation des différents peptides issus de l'hydrolyse des histones (Minshull et al., 2016). En effet, les composés semi-polaires et apolaires sont adsorbés sur l'*Hypersil Gold aQ* qui est une colonne C_{18} à « *end-capping* » polaire, tandis que l'*Hypercarb*, composée de 100 % de feuillets de carbone poreux, fonctionne de façon complémentaire puisqu'elle assure une forte rétention vis-à-vis des composés très polaires et chargés (Adoubel et al., 2000; Chaimbault et al., 2000; Zhao et al., 2014). Un tel montage en série permet d'améliorer l'efficacité de rétention contrairement à un montage en parallèle impliquant plus de connectiques et donc des volumes morts plus importants. Par ailleurs, l'intérêt de l'association de deux colonnes non ioniques réside dans l'emploi des mêmes phases mobiles (eau ultrapure et acétonitrile (0,1 % acide formique)) pour la préconcentration et la séparation analytique (*Hypersil Gold Q*) des cibles peptidiques et permet de pallier aux problèmes de contrôle du pH imposés par les phases mobiles ioniques.

L'utilisation de ce système de préconcentration en ligne « *dual trap* », connecté à une séparation analytique (colonne analytique C_{18} : *Hypersil Gold Q*), a permis d'amplifier de manière considérable le signal MS des peptides cibles. En effet, pour une même concentration de peptides injectée (soit $10 \text{ ng}\cdot\text{mL}^{-1}$), les signaux des composés étaient environ 200 fois supérieurs, comparés à l'injection directe (sans « *dual trap* ») (figure 78).

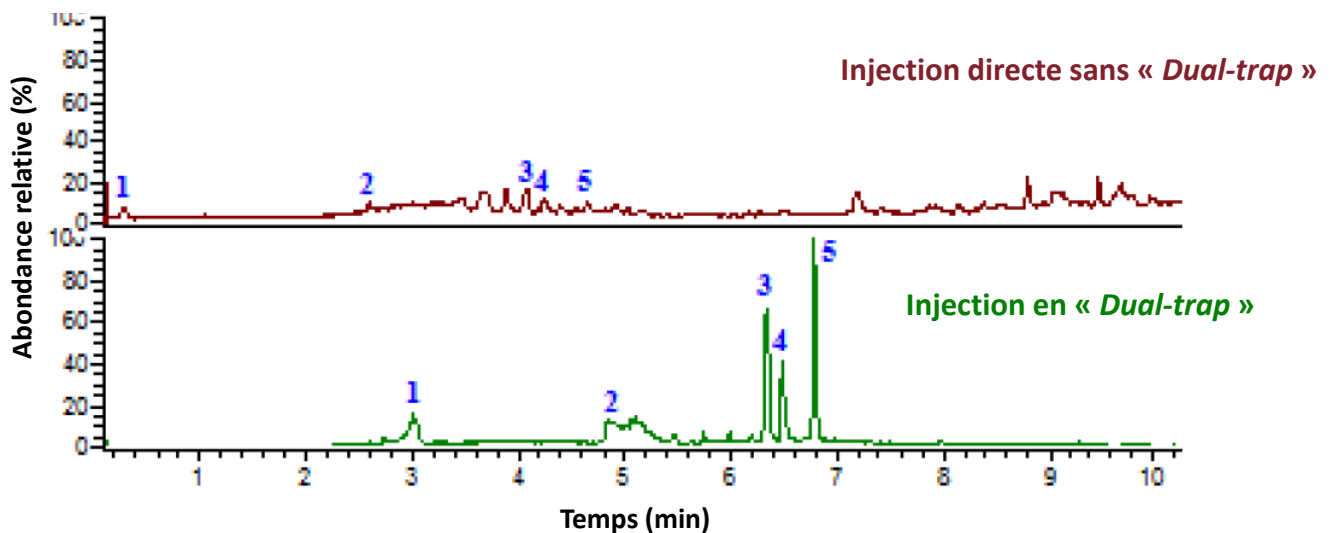


Figure 78. Courant ionique total (m/z 300-2000) d'une solution de peptides standards analysés par UPLC-MS/MS en mode « Full scan » avec une résolution égale à 70 000. La concentration injectée dans les deux systèmes est de 10 ng.mL^{-1} . 1. gly-tyr ; 2. val-tyr-val ; 3. met-enképhaline ; 4. Angiotensine II ; 5. leu-enképhaline

Nous avons par la suite évalué la sensibilité du système dans son ensemble en calculant les limites de détection (LOD) et de quantification (LOQ) des peptides standards. La sensibilité de l'analyse a été considérablement améliorée après préconcentration avec des LODs de l'ordre du $0,11 \text{ ng.mL}^{-1}$ et des LOQs d'environ $0,36 \text{ ng.mL}^{-1}$. Alors qu'elles étaient de $0,5 \text{ }\mu\text{g.mL}^{-1}$ et de $1,5 \text{ }\mu\text{g.mL}^{-1}$ en injection directe. Ce résultat montre clairement le grand intérêt de combiner un système orthogonal (*Hypersil Glod aQ-Hypercarb*) pour piéger une grande variété de cibles peptidiques, à une injection en large volume.

2. Identification des protéines cibles par spectrométrie de masse

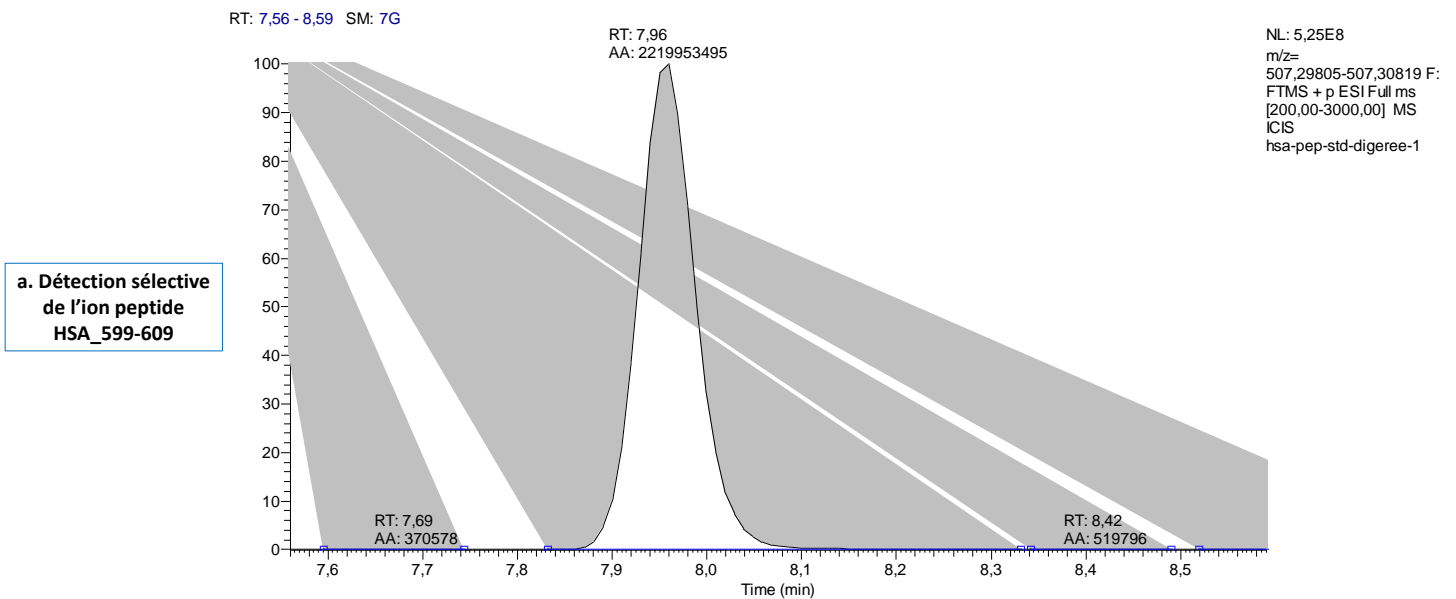
L'analyse protéomique par LC-MS/MS a pour but d'identifier et de déterminer précisément la quantité de protéines dans un échantillon donné. La relation entre les concentrations en protéines et en peptides dépend de l'efficacité de la digestion trypsique des protéines et de la protéotypicité des peptides ; un peptide protéotypique étant un peptide spécifique d'une protéine ciblée (Mallick et al., 2007). Afin d'optimiser les paramètres de masse, une solution standard de 4 protéines (HSA, lysozyme, Rnase et exoglucanase-1) a été utilisée.

L'analyse par LC-MS/MS de cette solution digérée par la trypsine nous a permis d'identifier précisément : 48, 10, 9 et 21 peptides théoriques, suivant la base de données *UniProtKB/Swiss-*

Chapitre I : Développement analytique

Prot, après digestion de la HSA, lysozyme, Rnase et exoglucanase-1, respectivement. Suivant une méthode LC définie préalablement, le temps de rétention, le rapport m/z des ions peptides multiprotonés $[M-H]^{z+}$, appelés aussi ions parents, et le rapport m/z de leurs ions fragments, appelés ions fils ont été combinés pour assurer une identification fiable.

Prenons l'exemple du peptide dont la position au niveau de la séquence de la HSA est 599-609 (peptide allant de l'acide aminé n°599 à l'acide aminé n°609). Sa masse molaire exacte est de $1013,59897 \text{ g.mol}^{-1}$ (figure 79). D'après la figure 79, nous avons une détection sélective du pic chromatographique de l'ion chargé deux fois à la masse m/z égale à 507,30328 ($tr= 7,96 \text{ min}$) (figure 79.a et 79.b). Cet ion étant le plus intense (comparé à l'ion chargé une fois, $m/z= 1013,59741$), nous avons cherché les ions fragments MS/MS (m/z égales à 570,32379 et 641,35962 ; figure 79.c) et nous les avons comparés à ceux trouvés dans la base de données (MS-Product), afin de confirmer l'identité du peptide.



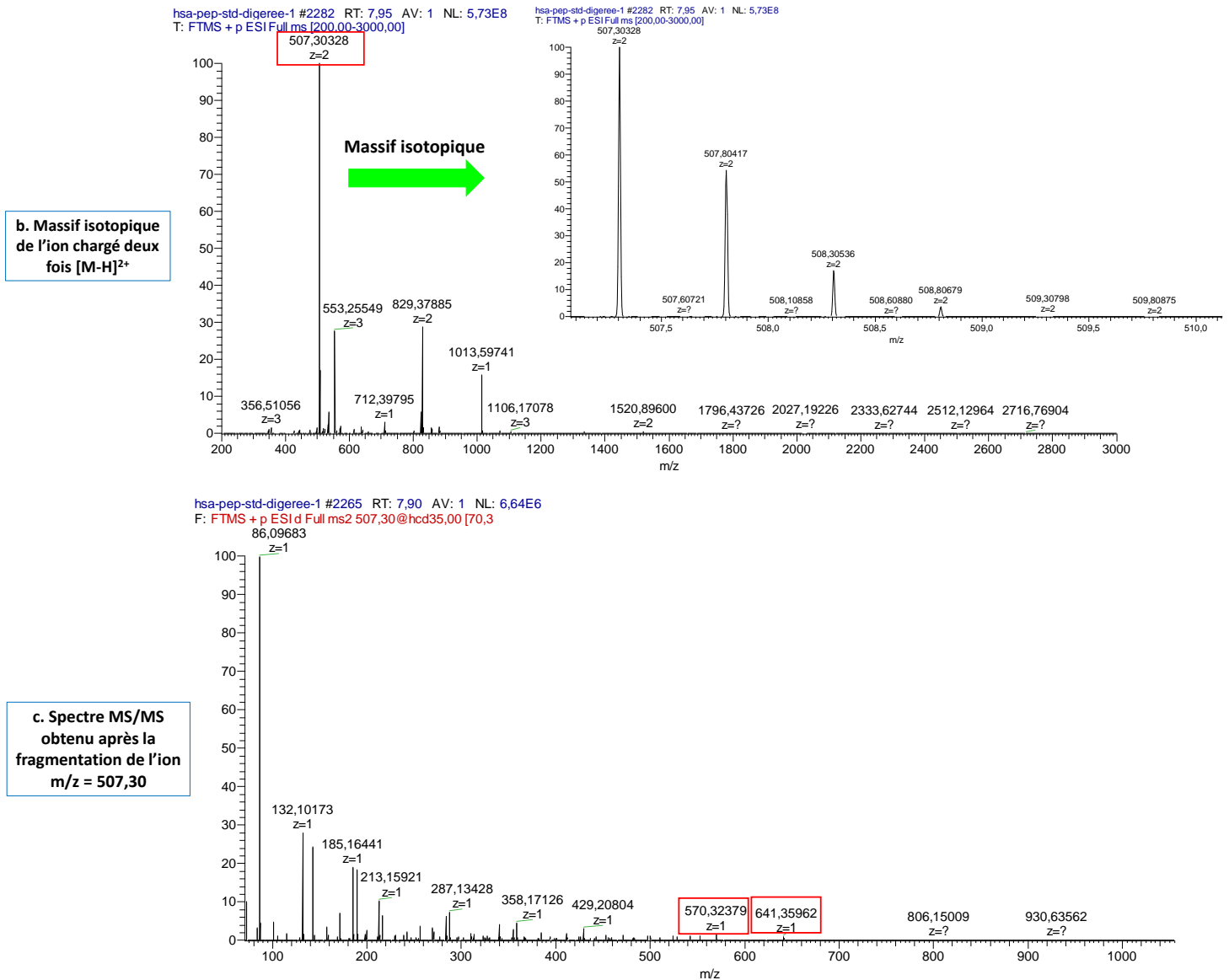


Figure 79. Exemple de spectre de masse MS et MS/MS pour peptide HSA_599-609

Ce travail a été réalisé pour chacune des protéines, ce qui nous a conduit au tableau 10 listant l'ensemble des peptides présentant un signal élevé et suivis au cours de ce travail.

3. Quantification des protéines

3.1. Quantification absolue

Au cours des chapitres II et III, les optimisations des protocoles d'extraction et de purification ont au départ été réalisées sur des solutions standards ou matrices modèles (injection directe sans « *dual trap* ») puis sur matrices réelles et complexes (injection en « *dual trap* »). L'injection d'un large volume en « *dual trap* » a été particulièrement intéressante pour

Chapitre I : Développement analytique

amplifier le signal des protéines présentes à des concentrations variables et/ou faibles dans des matrices complexes.

Afin de calculer les rendements de purification et d'extraction pour chaque protéine standard, nous avons réalisé des droites d'étalonnage pour chacun de leurs peptides cibles (aire peptide / aire étalon interne « Leu-enképhaline ») à partir de solutions standards préparées dans le tampon d'analyse. Les limites de détection (LODs) et de quantification (LOQs) pour chaque peptide ont alors été calculées.

Tableau 10. Liste des peptides identifiés après analyse UPLC-HRMS/MS après digestion tryptique des quatre protéines cibles. Nomenclature : LYSO= lysozyme ; RNASE= Ribonucléase ; HSA= Albumine humaine sérique ; EXO = exoglucanase-1. t_r (min) = temps de rétention en min

Chapitre I : Développement analytique

Protéine	Position peptide	Séquence peptidique	m/z des ions précurseurs				m/z des ions fragments			Injection directe	Injection en « dual trap »	Injection directe	Injection en « dual trap »		
			[M-H] ¹⁺	[M-H] ²⁺	[M-H] ³⁺	[M-H] ⁴⁺				t _r (min)	t _r (min)	R ²	LOD ; LOQ (pmol)	R ²	LOD et LOQ (pmol)
LYSO-1	52-63	FESNFNTQATNR	1428,650 2	714.8288	476.8882	357.9180	1152,5 392	804,39 59	690,35 29	3,28	5,56	0,997 8	0,54 ; 2,19	1	0,40 ; 1,41
LYSO-2	33-39	HGLDNYR	874,4166	437,7119	292,1437	219,3596	737,35 77	567,25 22	452,25 52	1,56	3,60	0,995 3	1,57 ; 5,33	1	1,94 ; 5,77
RNASE-1	66-87	CKPVNTFVHESL ADVQAVCS QK	2517,222 5	1259,114 9	839,7457	630,0611				4,80	6,84	0,997 9	0,22 ; 0,77	1	0,25 ; 1,67
RNASE-2	131-150	HIIVACEGNPYV PVHFDASV	2224,085 6	1112,546 4	742,0334	556,7768				5,27	7,29	0,995 6	0,13 ; 0,94	1	0,40 ; 1,79
RNASE-3	118-124	YPNCAYK	915,4029	458,2051	305,8058	229,6062	655,28 68	541,24 39	310,17 61	2,08	4,20	0,995 7	2,57 ; 8,74	1	- ; 0,05
HSA-1	570-581	AVMDDFAAFVE K	1342,634 8	671,8210	448,2164	336,4142	1041,4 888	926,46 18	811,43 49	5,69	7,70	0,994 7	0,20 ; 0,55	1	0,14 ; 0,36
HSA-2	599-609	LVAASQAALGL	1013,599	507,3031	338,5478	254,1552	570,32 46	641,36 17	213,15 98	5,01	7,06	0,987 5	0,18 ; 0,47	0,9 869	0,08 ; 0,21
HSA-3	162-168	YLYEIAR	927,4934	464,2504	309.8360	232,6288	764,43 01	651,34 61	488,28 27	4,02	6,31	0,983 8	0,17 ; 0,46	1	0,05 ; 0,07
EXO-1	412-432	GSCSTSSGVPAQ VESQSPNA K	2077,945 5	1039,476 4	693,3200	520,2418	218,14 99			3,07	5,34	0,997 5	0,17 ; 0,50	1	0,11 ; 0,30
EXO-2	184-195	YGTGYCDSQCP R	1463,567 8	732.2876	488.5275	366,6474	175,11 90	221,09 21		2,51	4,79	0,991 4	0,37 ; 1,36	0,9 911	0,18 ; 0,69
EXO-3	433-439	VTFSNIK	808,4563	404,7282	270,1570	202,8695	201,12 34	147,11 28		3,54	5,82	0,993 1	1,18 ; 3,80	0,9 993	0,29 ; 0,83

Chapitre I : Développement analytique

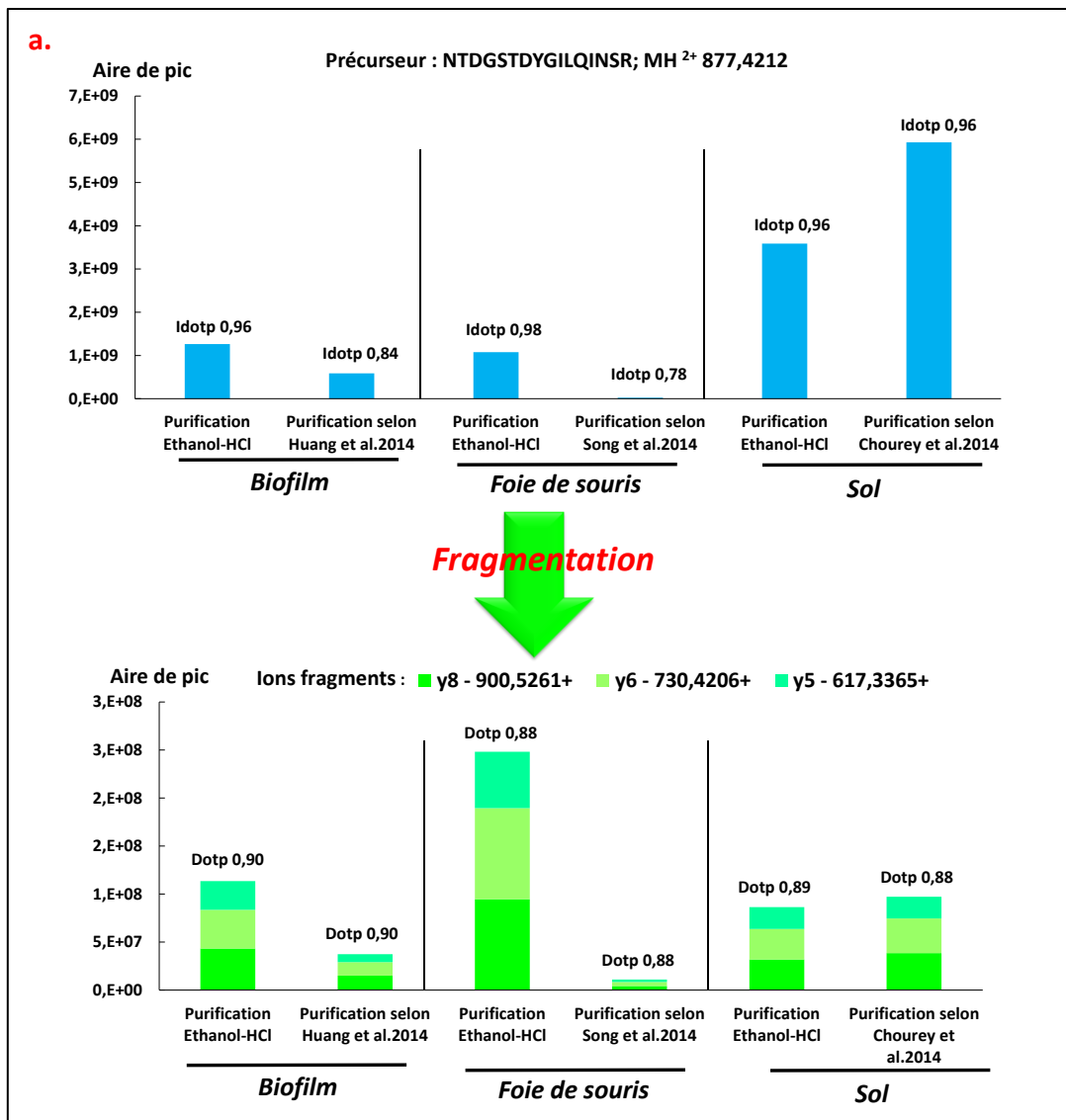
Les limites de détection et de quantification obtenues dans les deux systèmes d'analyse sont de l'ordre de la picomole. En injection directe, les LODs et LOQs varient entre 0,03 et 0,47 $\mu\text{g.mL}^{-1}$ (0,13 et 2,57 pmol) et entre 0,09 et 1,60 $\mu\text{g.mL}^{-1}$ (0,46 et 8,74 pmol), respectivement. En injection « *dual trap* », elles diminuent à 0,05 et 1,86 ng.mL^{-1} (0,05 et 1,94 pmol) et (0,05 et 5,77 pmol), respectivement. Ceci reflète bien la sensibilité du système analytique dans sa totalité (« *dual trap* » + colonne *Hypersil Gold Q+* HRMS)

Les courbes d'étalonnage ont été par la suite utilisées afin de déterminer les rendements d'extraction et de purification des protéines.

3.2. Quantification relative

Afin de valider les protocoles d'extraction et de purification, les quantités des protéines introduites (spikées) sur matrices complexes ont été déterminées avec un logiciel de protéomique dédié à la quantification : *Skyline*. Le prétraitement des données MS/MS par *Skyline* consiste tout d'abord à (1) importer la séquence en acides aminés des protéines cibles, (2) créer une librairie spectrale MS/MS et (3) importer les fichiers bruts (fichiers *raw Xcalibur*) issus de l'acquisition du spectromètre de masse. Il est ensuite nécessaire de paramétrer les données relatives à la détection des ions précurseurs et de leurs différentes transitions (couple ion précurseur/ions fragments). *Skyline* extrait pour chaque transition le courant ionique et intègre ensuite l'aire sous le pic chromatographique correspondant à l'ion précurseur associé à cette transition (Egertson et al., 2015; MacLean et al., 2010). La quantification relative des peptides peut être également réalisée sur la base des aires de pics chromatographiques des ions fragments (y ou b) (figure 80.a et b).

Sur les figures 80, sont présentés également les résultats *Skyline* relatifs à une quantification relative des deux ions précurseurs du lysozyme. L'objectif ici était de comparer les différents protocoles de purification (cf. chapitre II) et de vérifier si la quantification sur la base des précurseurs permettait d'obtenir la même tendance qu'une quantification relative sur la base de leurs fragments.



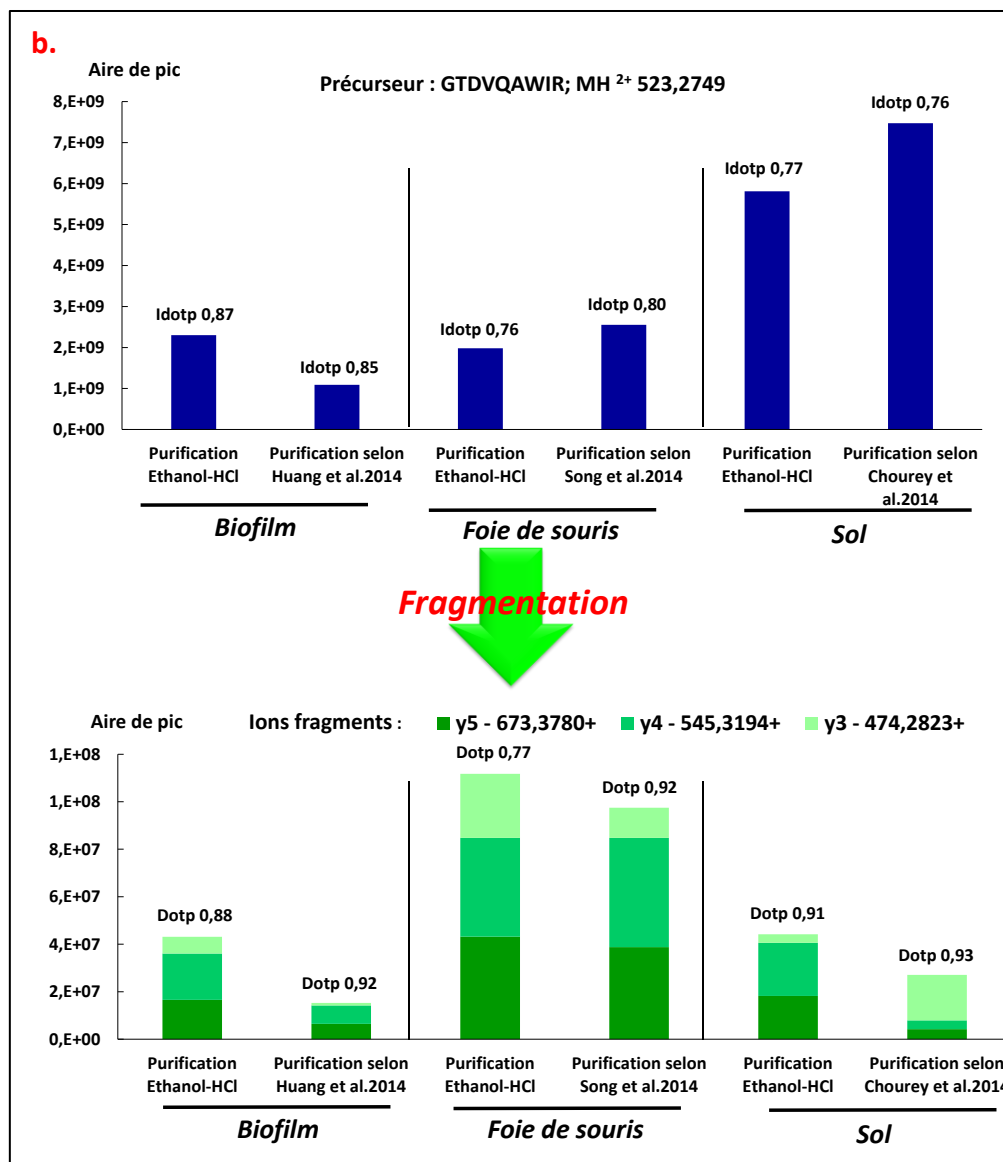


Figure 80. Exemple de quantification relative du lysozyme à partir de matrices complexes, obtenue par le logiciel Skyline. a. Quantification sur l'ion précurseur NTDGSTDYGILQINSR ($m/z=877,4212 [MH]^{2+}$) et sur ses ions fragments y8, y6 et y5 ; b. Quantification sur le précurseur GTDVQAWIR ($m/z=523,2749 [MH]^{2+}$) et sur ses ions fragments y5, y4 et y3

D'après la figure 80, nous pouvons constater que la quantification relative des deux ions précurseurs du lysozyme est légèrement différente de celle obtenue à partir de leurs ions fragments. Ceci est en particulier vrai pour la matrice foie de souris et sol. Dans le cas d'une quantification sur les spectres MS/MS, *Skyline* intègre l'aire sous le pic chromatographique correspondant à chacun des fragments identifiés, ce qui rend la quantification plus sensible et précise (Egertson et al., 2015; Rauniyar, 2015). En effet, la quantification à l'échelle des ions

Chapitre I : Développement analytique

fragments est moins impactée par les interférences, contrairement à la quantification réalisée sur les ions précurseurs qui est généralement biaisée par l'effet matrice (Bruderer et al., 2015). L'identification et la quantification des cibles protéiques à partir des ions fragments seront donc plus informatives et caractéristiques de la séquence du peptide de départ.

Le logiciel *Skyline* calcule également un indice isotopique *isotopic dot product (idopt)*, compris entre 0 et 1. Cette valeur numérique permet de comparer la distribution isotopique observée expérimentalement à celle attendue théoriquement pour un peptide donné. Ainsi, une valeur de *idopt* proche de 1 représente une similitude quasi-totale entre la théorie et l'expérimental (Rauniyar, 2015).

4. Optimisation de la digestion trypsique

L'approche « *bottom-up* » repose sur l'analyse par LC-MS/MS de mélanges peptidiques issus de l'hydrolyse enzymatique des protéines (Aebersold and Mann, 2016; Pragma et al., 2017).

Cette hydrolyse est réalisée le plus souvent en présence d'une enzyme appelée trypsine. Cette enzyme génère des peptides basiques au niveau de leurs extrémités C-terminales. Ces peptides une fois fragmentés dans la cellule de collision du spectromètre de masse conduisent généralement à une série d'ions y (cf. Partie bibliographique) rendant leurs spectres de masse facilement interprétables par les bases de données (Pan et al., 2014). Cette endopeptidase coupe spécifiquement les liaisons peptidiques après les acides aminés chargés positivement, lysine (Lys) et arginine (Arg), sauf s'ils sont suivis d'une proline (Pro). La digestion trypsique est réalisée idéalement en milieu légèrement basique 7,5 et 9 à une température de 37 °C.

Néanmoins, une légère modification du pH (surtout une baisse du pH, égale à 4) ou de la température peut entraîner une altération de son activité catalytique. Ceci diminue par conséquent la qualité de l'échantillon, puisque la totalité des protéines ne sera pas « digérée ».

Nous avons donc optimisé les paramètres de digestion, que sont la composition de la solution tampon, le ratio enzyme/substrat (E/S), la durée de la digestion et la concentration d'acide à ajouter pour stopper la réaction enzymatique. Cette optimisation a été réalisée sur une solution standard d'angiotensine II, peptide de 8 acides aminés, dont une arginine en milieu de séquence.

Le choix du tampon de digestion est primordial pour le contrôle de l'activité enzymatique de la trypsine (Bala et al., 2000). Le bicarbonate d'ammonium (entre 10 et 50 mM, pH 8) est le

Chapitre I : Développement analytique

tampon le plus employé en protéomique (P. Hao et al., 2015; Min et al., 2015). D'autres tampons à base de Tris (entre 50-100 mM, pH 8) (Chen et al., 2007; P. Hao et al., 2015) sont également employés. Dans notre étude, nous avons testé deux tampons, un tampon classique le bicarbonate d'ammonium et le tampon Tris-bicarbonate d'ammonium. Les résultats sont reportés dans la figure 81.

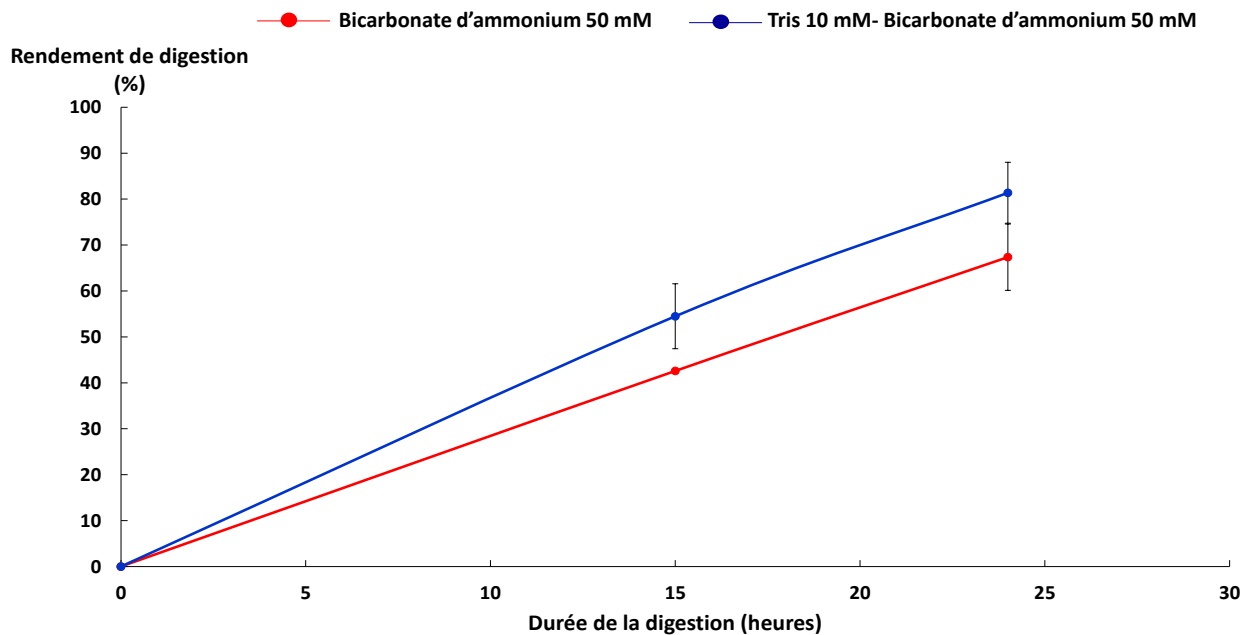


Figure 81. Effet du tampon de digestion sur l'efficacité de la digestion de l'angiotensine II. Le ratio trypsine/substrat est fixé à 1/6 (Au début de la réaction enzymatique (t = 0 h 00), l'activité de la trypsine est nulle (rendement = 0 %))

L'utilisation du tampon Tris 10 mM-Bicarbonate d'ammonium 50 mM (pH 8,5-8,6) permet d'atteindre un rendement de digestion de 80 % après 24 h 00 de réaction. Ce tampon permet donc à la trypsine d'atteindre une activité protéolytique quasi-maximale. Ceci est probablement lié à son pH légèrement plus basique que le bicarbonate d'ammonium 50 mM (pH 8). L'activité optimale de la trypsine se situe entre pH 7,5 et 8,5 (Hustoft et al., 2012). Par ailleurs, selon le fournisseur *Promega*, la trypsine porcine présente une activité maximale dans une gamme de pH allant de 7-9, ce qui lui permet notamment de contourner l'acidité de certaines matrices biologiques.

Le rapport massique enzyme/substrat (E/S) est également un paramètre important. De manière générale, le rapport E/S utilisé est compris entre 1/10 et 1/100 (w/w) (Lebert et al.,

Chapitre I : Développement analytique

2015). Des ratios supérieurs allant jusqu'à 1/3 ont été appliqués notamment lorsqu'il s'agissait d'étudier le protéome total d'un échantillon complexe (Hildonen et al., 2014; Li et al., 2015). Dans notre étude, nous avons évalué l'effet de quatre ratios E/S sur l'efficacité de digestion trypsique (figure 82).

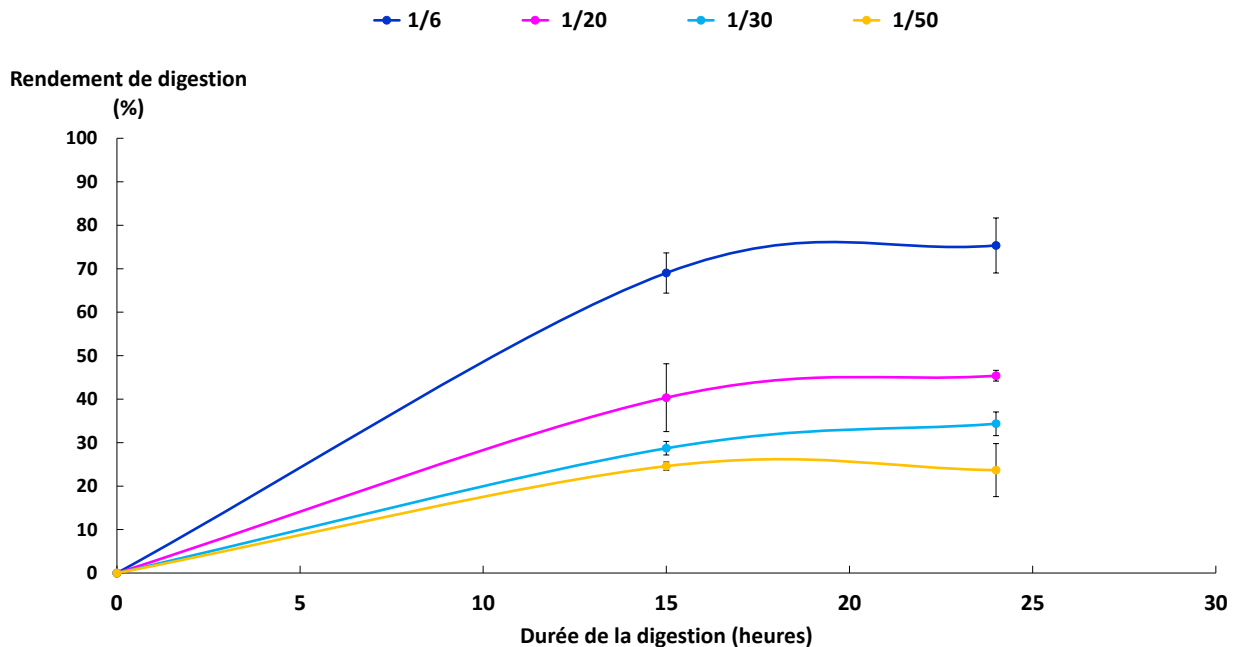


Figure 82. Effet du ratio E/S sur l'efficacité de la digestion de l'angiotensine II

Le rendement de digestion final augmente significativement avec le ratio E/S. Un ratio élevé de 1/6 permet une digestion quasi-totale de l'angiotensine II au bout de 24 h 00 de digestion (Hildonen et al., 2014). Ce ratio a été retenu pour la suite de l'étude afin de garantir une digestion optimale des protéines d'intérêt.

Ce ratio nous permettra d'être en large excès de trypsine, ce qui devrait permettre la digestion de l'ensemble d'un échantillon, quelle que soit la concentration en protéines. Des ratios E/S plus importants ($E/S = 1/3$) peuvent être également envisagés dans le cas des échantillons très complexes.

L'activité de la trypsine est inhibée en milieu acide. Nous avons alors testé deux concentrations d'acide formique (0,4 % et 2,5 % (v/v)) afin de stopper la réaction de digestion. Nous avons constaté qu'une concentration de 0,4 % est suffisante pour inhiber l'activité de la trypsine (Lebert et al., 2015).

Chapitre I : Développement analytique

Ces mêmes paramètres de digestion ont été évalués sur la digestion de la HSA. Les résultats obtenus étaient similaires. Un tampon Tris 10 mM-Bicarbonate d'ammonium 50 mM, une durée de 24 h 00 et un ratio E/S de 1/6 se sont révélés être les conditions optimales.

5. Conclusion

L'analyse des protéines dans des échantillons complexes nécessite une mise au point de méthodes analytiques hautement performantes pour assurer une séparation, une identification et une quantification efficiente de peptides présents à des concentrations variables suivant l'échantillon étudié. Dans ce chapitre, nous avons montré le potentiel d'un système chromatographique en ligne « *dual trap* », pour l'amplification du signal des peptides tout en assurant une séparation et une détection efficace des peptides. Un gain de sensibilité significatif entre le système « *dual trap* » et le système en injection directe a bien été obtenu.

Le système « *dual trap* » développé au cours de cette étude a fait l'objet d'une publication dans *Talanta*.

L'analyse protéomique par LC-MS/MS nous a permis de déterminer expérimentalement les séquences des peptides protéotypiques issus de la digestion des protéines cibles (HSA, lysozyme, Rnase et exoglucanase-1) et de renseigner toutes les informations nécessaires à l'identification de ces peptides (temps de rétention, m/z des ions précurseurs ($[M-H]^+$, $[M-H]^{2+}$, $[M-H]^{3+}$), m/z des ions fragments, etc.).

L'analyse des données au moyen du logiciel *Skyline*, nous a permis d'établir une méthode de quantification relative basée sur les aires des pics chromatographiques correspondant aux ions fragments. Dans ce cas, les spectres MS/MS permettent d'accroître de manière considérable la qualité d'identification, et donc de quantification des peptides détectés.

Enfin, nous avons également optimisé la digestion trypsique des protéines en solution en vue de leur analyse par LC-MS/MS. Plusieurs paramètres influençant l'activité catalytique de la trypsine ont été évalués (choix du tampon de digestion, ratio E/S, etc.) pour permettre une hydrolyse optimale des protéines présentes dans des matrices complexes (ratio E/S élevé (1/6), tampon légèrement basique pour contourner l'acidité de certaines matrices biologiques).

Les conditions optimales de la digestion trypsique ont fait l'objet d'une publication axée sur le ciblage du microenvironnement tumoral pour le traitement de certains cancers. Cette publication est actuellement en ligne dans *Chemical Science*.

*Chapitre II : Développement d'une
approche simple pour la purification des
protéines*

Chapitre II : Développement d'une approche simple pour la purification des protéines

Les progrès instrumentaux réalisés au cours de ces dernières années ont permis la détection par spectrométrie de masse de protéines dans des milieux très complexes. Ils ont conduit à des avancées majeures axées sur la technologie (haute résolution, haute sensibilité) pour l'analyse de traces, sans pour autant améliorer la compréhension des systèmes étudiés (compréhension des interactions des protéines avec leurs milieux, conformation, accessibilité, etc.). Ainsi, de manière générale, les protocoles de purification développés en protéomique reposent sur la méthode essai-erreur, générant des protocoles le plus souvent longs et multi-étapes. Ces approches utilisent de surcroît de nombreux agents chimiques agressifs souvent incompatibles avec l'analyse par spectrométrie de masse. En outre, la réponse en masse des peptides issus de la digestion de protéines est fortement impactée par la présence de ces composés (Zhou et al., 2012). L'obtention d'un échantillon protéique parfaitement purifié pour une analyse protéomique par LC-MS reste donc une étape clé.

L'extraction des protéines d'une matrice donnée conduit le plus souvent à une co-extraction de composés interférents tels que les acides nucléiques, les polysaccharides, les lipides, endogènes à la matrice, mais également à la présence des agents d'extraction tels que des détergents, sels, chaotropes (Cañas et al., 2007). Afin d'éliminer ces interférents, une étape de purification est nécessaire. Dans une grande majorité des protocoles développés en protéomique, la purification inclut une étape de précipitation suivie de multiples lavages et d'une dernière étape de solubilisation des protéines dans un tampon aqueux avant digestion trypsique. Compte tenu des limites liées à l'extraction (utilisation de fortes concentrations de détergents, perte des protéines durant les étapes de filtration, etc.), la purification des protéines se doit être la plus simple et la plus efficace possible afin de concentrer de façon sélective la fraction protéique et ainsi accroître la sensibilité de l'analyse LC-MS.

Dans ce chapitre nous détaillerons la démarche expérimentale que nous avons suivie afin de proposer une précipitation quantitative des protéines et un nouvel agent chimique permettant d'assurer en une seule étape un lavage et une solubilisation optimale du culot protéique sans utiliser de détergents ni de forces mécaniques thermiques. Une étude mécanistique, *via* des analyses spectroscopiques, a permis de valider le choix de cet agent en démontrant au niveau moléculaire son effet sur la conformation et la solubilité des protéines.

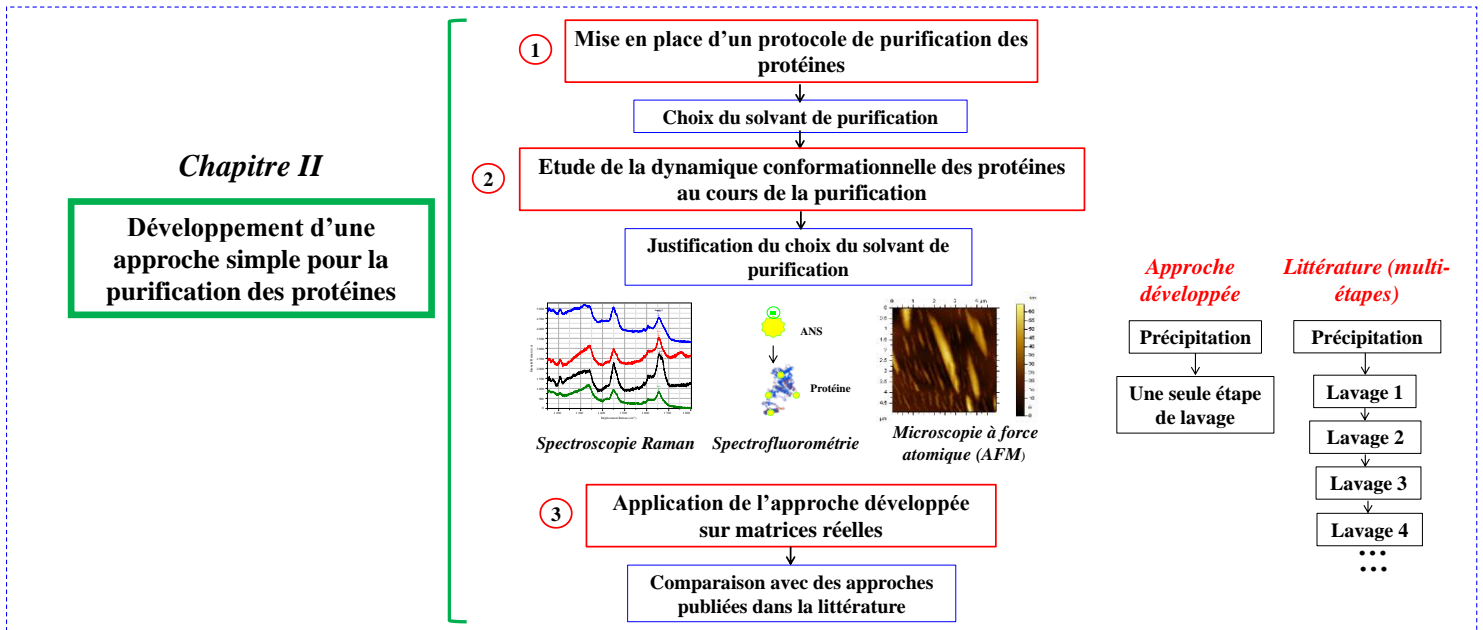


Figure 83. Stratégie expérimentale

1. Étude préliminaire

L'acide trichloroacétique (TCA) (entre 10 et 25 % (w/v)) est un des agents de précipitation les plus utilisés en protéomique (Fic et al., 2010; Jacobs et al., 2001; Jiang et al., 2004; Rajalingam et al., 2009). Il assure une concentration efficace des protéines par rapport à d'autres réactifs chimiques utilisés couramment en protéomique (solvants organiques, sels, etc.) (Feist and Hummon, 2015; Fic et al., 2010; Jiang et al., 2004). Il est généralement dissout dans l'eau ou dans l'acétone.

Nous avons évalué la précipitation des protéines avec une solution aqueuse à 25 % (w/v) (concentration finale) (Chourey et al., 2010) et comparé le taux de protéines précipitées à celui obtenu avec une solution de TCA/acétone à 10 % avec ou sans DTT, solution fréquemment utilisée en protéomique pour purifier les protéines à partir de matrices végétales ou biologiques (Isaacson et al., 2006; Song et al., 2016, 2012b; Wu et al., 2014) (tableau 11).

Chapitre II : Développement d'une approche simple pour la purification des protéines

Tableau 11. Évaluation de la quantité de HSA présente dans le surnageant. Les valeurs n'ayant pas la même lettre sont significativement différentes ; NI : non identifiée

	Agents de précipitation		
	10 % TCA/acétone (0,07 % DTT)	10 % TCA/acétone	25 % TCA/eau
Rendement de la HSA perdue dans le surnageant (%)	30,75 ^a	41,22 ^a	NI ^b

Des pertes significatives de HSA dans le surnageant acétone ont été mises en évidence contrairement aux essais en solution aqueuse. Ce résultat indique que cette protéine se solubilise partiellement dans les solvants organiques. Dans ce cas, les groupements hydrophobes trichlorométhyle du TCA augmentent la solubilité du complexe protéine-TCA dans les solvants organiques (Chen et al., 2005). Pour la suite de notre travail, nous nous placerons donc en milieu aqueux lors de l'étape de précipitation.

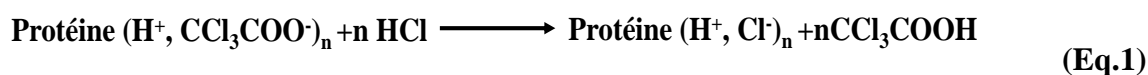
Une fois la précipitation validée, les étapes de lavage et de solubilisation du culot protéique ont été abordées. En effet, plusieurs études ont démontré que le TCA ne permettait pas la resolubilisation totale du culot protéique dans les tampons aqueux utilisés pour l'hydrolyse enzymatique (Nandakumar et al., 2003). Ceci pourrait s'expliquer par le fait qu'en présence de TCA, les protéines adoptent une structure compacte, rigide appelée *molten globule*, stabilisée par le TCA (Kumar et al., 1994; Ptitsyn, 1995; Rajalingam et al., 2009; Schultz, 2000; Sivaraman et al., 1997; Xu et al., 2003)). Afin d'éliminer cet excès de TCA et de faciliter la resolubilisation des protéines, le culot est généralement lavé plusieurs fois avec un solvant organique dont le plus connu est l'acétone (Chourey et al., 2010; Jiang et al., 2004; Nandakumar et al., 2006; Zhang et al., 2015). Bien qu'il permette d'éliminer le TCA en excès, l'acétone peut induire des pertes en protéines lors des différentes étapes de lavage (Crowell et al., 2013) ainsi qu'une mauvaise resolubilisation des culots protéiques dans le tampon d'analyse (Crowell et al., 2013; R. Hao et al., 2015; Jiang et al., 2004; Mesquita et al., 2012; Zhang et al., 2015). Pour pallier à ce problème, les lavages à l'acétone peuvent être suivis par une solubilisation en présence d'agents agressifs chaotropes ou dénaturants (urée, thiourée) et de détergents (SDS). Ces agents permettent de dénaturer complètement la structure *molten globule*, la rendant ainsi plus flexible et propice à la solubilisation (R. Hao et al., 2015; Jacobs et al., 2001; Lin et al., 2013; Manadas et al., 2006; Nandakumar et al., 2003; Rabilloud et al., 2007; Wu et al., 2014;

Chapitre II : Développement d'une approche simple pour la purification des protéines

Zhang et al., 2011, 2015). Toutefois, ceux-ci devant être complètement retirés de la solution pour permettre une digestion et une analyse optimales, une étape supplémentaire de purification est nécessaire.

Il nous fallait donc éliminer les interactions du TCA avec les protéines pour modifier la structure *molten globule* et ainsi faciliter leur solubilisation sans induire d'interférences supplémentaires.

L'efficacité d'un nouvel agent de lavage du précipité, l'éthanol/HCl, a été étudiée à partir de solutions standards de protéines (HSA, lysozyme, Rnase et exoglucanase-1). L'idée était d'échanger le contre-ion (OOC-CCl_3) du TCA par les ions chlorures (Cl^-) du HCl ($\text{CCl}_3\text{COO}^- \rightarrow \text{Cl}^-$) (Eq.1), afin d'extraire le TCA du culot protéique et le transférer dans le solvant organique (Nishii et al., 1995).



Pour que le transfert du TCA ne s'accompagne pas d'une réhydratation des protéines sous leur forme ionique pouvant alors induire leur solubilisation partielle dans la phase aqueuse, nous avons utilisé de l'éthanol/HCl anhydre (HCl gazeux dissout).

L'emploi de l'éthanol (éthanol/HCl) à froid (-20 °C) permettrait ainsi de maintenir la protéine sous forme précipitée et éviter ainsi d'éventuelles pertes au cours de l'étape de lavage.

Nous avons alors testé l'effet de ce nouvel agent en comparaison des agents usuels. Une étude préliminaire par dosage Bradford et électrophorèse 1D nous a permis de confirmer l'efficacité de la combinaison éthanol/HCl (sans eau) de par la quantité totale de protéines récupérée après des lavages à l'acétone, l'éthanol et l'acétone/HCl (figure 84 ; annexe 1).

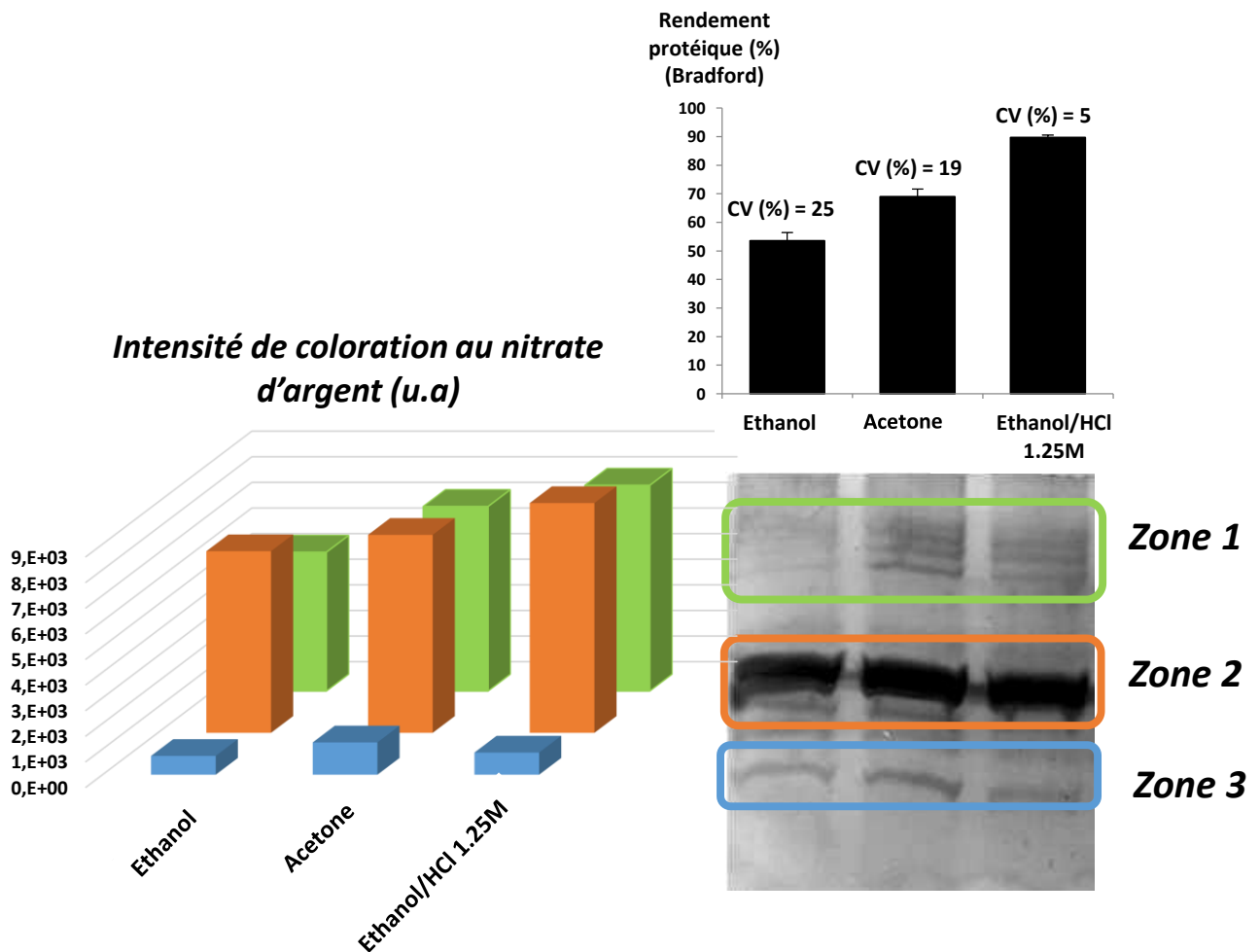


Figure 84. Electrophorèse 1D d'une solution de protéines standards (HSA, lysozyme, Rnase et exoglucanase-1) précipitée au TCA et lavage avec les différents solvants (éthanol, acétone et éthanol/HCl). La révélation des bandes protéiques a été réalisée par coloration au nitrate d'argent. Les protéines ont été également quantifiées par la méthode de Bradford

Les résultats du dosage des protéines par la méthode Bradford ont montré que le lavage à l'éthanol/HCl permet de récupérer environ 90 % des protéines totales (figure 84). Ces résultats sont en concordance avec ceux obtenus par électrophorèse sur gel de polyacrylamide, où les bandes protéiques les plus intenses correspondent bien aux protéines lavées avec ce même agent.

Pour une meilleure compréhension nous avons également évalué l'efficacité de ces agents de lavage pour chaque protéine modèle par UPLC-HRMS/MS. La stratégie d'analyse est représentée figure 85. Les résultats sont présentés dans la figure 86.

Chapitre II : Développement d'une approche simple pour la purification des protéines

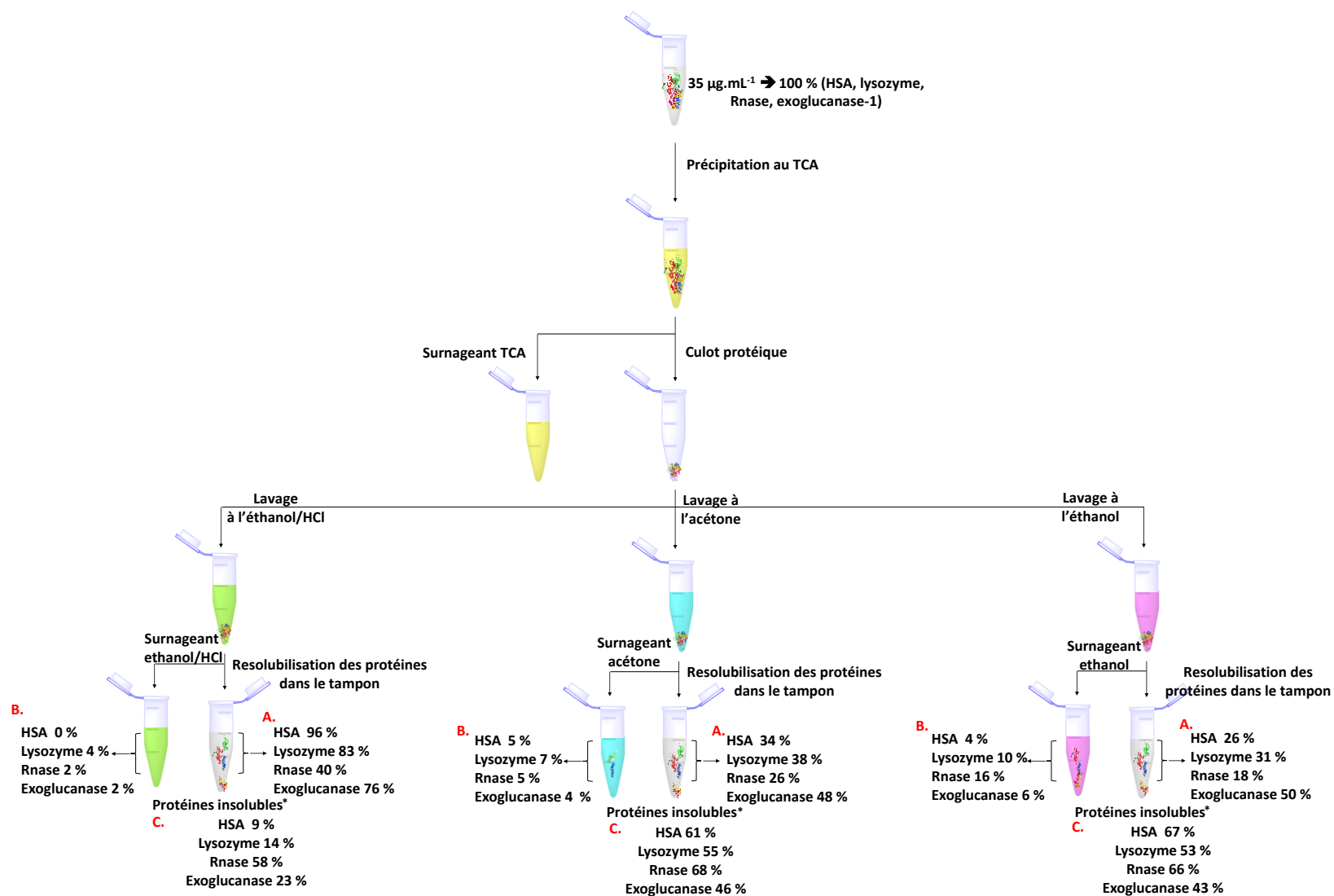


Figure 85. Stratégie d'analyse

Chapitre II : Développement d'une approche simple pour la purification des protéines

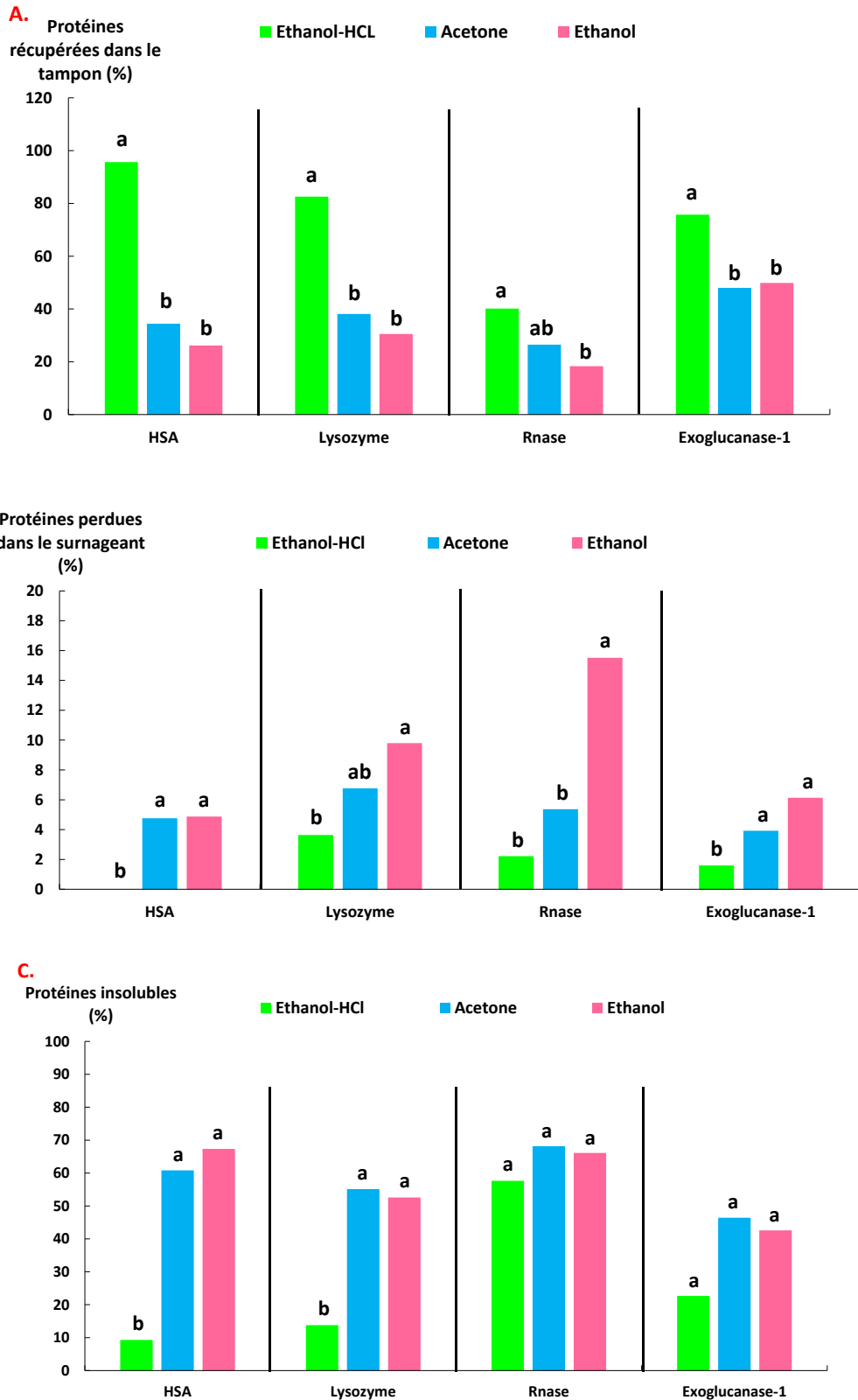


Figure 86. Comparaison des rendements de récupération des protéines après précipitation au TCA et lavage avec les différents solvants. * pourcentage de protéines insolubles = 100 % - % protéines dans le surnageant - % protéines dans le tampon d'analyse

Chapitre II : Développement d'une approche simple pour la purification des protéines

D'après la figure 86.A, nous pouvons constater que le lavage à l'éthanol/HCl a permis d'atteindre les meilleurs rendements de récupération de la HSA, du lysozyme et de l'exoglucanase-1 dans le tampon aqueux utilisé pour la digestion et l'analyse en LC-MS/MS (entre 77 et 96 %). L'acétone et l'éthanol ont quant à eux conduit à une perte des protéines dans le surnageant (entre 4 et 16 %) (Capriotti et al., 2017; Chen et al., 2005; Crowell et al., 2013; KORNER and DEBRO, 1956; Nandakumar et al., 2003; Rodkey, 1965; Zhang et al., 2015). Ces mêmes agents de lavage n'ont permis de solubiliser qu'environ 50 % des protéines précipitées (entre 45 et 70 % de protéines non solubilisées). L'éthanol/HCl a permis une solubilisation de 3 protéines sur 4 (HSA, lysozyme et exoglucanase-1), tandis qu'une resolubilisation partielle de la Rnase a été obtenue.

Afin de comprendre l'efficacité de l'éthanol/HCl comparée aux autres agents de lavage, et ce quelle que soit la protéine ciblée, une étude mécanistique a ensuite été réalisée. Le TCA induit la précipitation des protéines en déstabilisant leur structure native *via* la rupture des liaisons hydrophobes, ioniques et hydrogènes qui la stabilisent. Les ions trichloroacétates (OOC-CCl_3) interagissent avec les groupements positifs de la protéine, au niveau des sites basiques des résidus aminoacides. Le TCA forme également des liaisons hydrogènes ($\text{C-Cl}\dots\text{H-Y}$) entre ses atomes de chlore et les groupements donneurs d'hydrogène du squelette peptidique (N-H et O-H). Les groupements trichlores du TCA peuvent également former des liaisons halogènes avec les groupements carbonyles du squelette peptidique ($\text{C-Cl}\dots\text{O-C}$) (Borozan and Stojanović, 2013). Ces interactions intermoléculaires (hydrogènes et halogènes) participent à la formation de cet état intermédiaire de repliement, le *molten globule*. Cet état est en outre caractérisé par une structure compacte, partiellement dénaturée, qui préserve les mêmes éléments de structures secondaires que l'état natif, mais perd l'intégrité de sa structure tertiaire. Ceci rendrait vraisemblablement les protéines peu enclin à la solubilisation dans des tampons aqueux (Dolgikh et al., 1981; Ptitsyn, 1995).

2. Compréhension moléculaire de l'effet de l'éthanol/HCl sur la déstabilisation du *molten globule*

Le TCA induit donc une précipitation des protéines sous une forme MG. Pour expliquer les premiers résultats obtenus avec l'éthanol/HCl, une hypothèse serait que cet agent déstabiliserait l'état *molten globule* et le convertirait en un autre état plus favorable à la solubilisation. Afin de mieux appréhender l'effet de l'éthanol/HCl sur le MG, nous avons étudié de manière exhaustive les mécanismes moléculaires impliqués dans les changements structuraux des

protéines au moyen de techniques spectroscopiques : la spectrofluorométrie, la spectroscopie Raman et la microscopie à force atomique.

2.1. Spectrofluorométrie

La spectrofluorométrie est une technique couramment employée pour étudier les changements conformationnels des protéines en fonction de leur environnement. L'acide 1-anilino-8-naphthalène sulfonate (ANS) a été utilisé comme fluorophore extrinsèque, car il permet de mettre en évidence l'état *molten globule* (Kumar et al., 1994; Rajalingam et al., 2009; Sivaraman et al., 1997). En effet, il se fixe aux régions hydrophobes du *molten globule* dirigées vers l'extérieur du milieu, et forme ainsi un complexe fluorescent après excitation entre 370 et 390 nm. Cette fixation entraîne une variation de la fluorescence de l'ANS en termes d'intensité et de déplacement de la longueur d'onde d'émission (Kumar et al., 1994). Une modification de cette fluorescence lors de l'ajout d'un agent de lavage serait donc le signe d'un changement de conformation de l'état *molten globule*. L'ANS peut également se fixer *via* son groupement sulfonate sur les groupements chargés positivement des résidus d'acides aminés (histidine, arginine et lysine) (Hirsch et al., 2004; Matulis and Lovrien, 1998), ce qui peut fausser l'intensité de la fluorescence maximale d'émission de l'ANS (Qadeer et al., 2012).

Les données de fluorescence obtenues pour chaque protéine après précipitation au TCA et lavage à l'acétone, l'éthanol et l'éthanol/HCl sont présentées dans le tableau 12.

Tableau 12. Propriétés spectroscopiques de la HSA, lysozyme et Rnase après précipitation au TCA et lavage à l'éthanol/HCl

Liaison de l'ANS	Lysozyme		HSA		Rnase	
	RFI * 400 nm	λ_{\max} (nm)	RFI * 400 nm	λ_{\max} (nm)	RFI * 400 nm	λ_{\max} (nm)
Native	1.14	510	1.25	555	1.04	535
TCA	1	510	1	515	1	540
Acétone	0,40	545	1,04	515	1	540
Ethanol	0,39	545	0,48	505	0,97	540
Ethanol/HCl	0,66	530	1,19	530	1,03	535

*RFI : Intensité de fluorescence relative

Nous pouvons constater qu'un lavage à l'acétone n'entraîne pas une modification de l'intensité de la fluorescence du couple ANS-HSA et ANS-Rnase. Ce résultat indique donc que ces deux protéines conservent leur état *molten globule*. Ceci pourrait expliquer notamment leur faible resolubilisation dans le tampon de digestion (figure 86.B). En revanche, l'acétone tout

Chapitre II : Développement d'une approche simple pour la purification des protéines

comme l'éthanol, induit une modification de l'état MG du lysozyme (déplacement de la longueur d'onde maximale d'émission : 510 nm \rightarrow 545 nm). Alors que l'acétone et l'éthanol ont un effet variable sur les *molten globule* des 3 protéines, l'éthanol/HCl conduit, quelle que soit la protéine, à un déplacement de longueur d'onde maximale d'émission (déplacement entre 15 et 20 nm pour le lysozyme et la HSA), ainsi qu'à une modification de l'intensité de fluorescence de l'ANS couplé à la protéine. Un tel résultat témoigne clairement d'une modification structurale des protéines au cours des différentes étapes de purification. Les spectres d'émission correspondants au tableau 12 sont représentés dans l'annexe 2.

2.2. Spectroscopie Raman

De manière générale, l'étude de la dénaturation et de l'agrégation des protéines implique une modification de leur structure secondaire. Les changements de conformation des protéines précipitées au TCA et lavées à l'éthanol/HCl ont été étudiés par spectroscopie Raman. Cette technique permet de décrire de manière détaillée la dynamique structurale des protéines dans un domaine allant de 1500 à 1800 cm^{-1} . Ce domaine d'étude constitue une empreinte moléculaire pour l'analyse des bandes amides (I, II et III), et plus particulièrement la bande amide I (entre 1600-1690 cm^{-1}) qui correspond principalement à la vibration d'élongation du groupement carbonyle C=O du squelette peptidique. Cette bande sensible aux changements de structures secondaires correspond généralement à une superposition de plusieurs gaussiennes dont chacune est attribuée à une structure secondaire spécifique (structures désordonnées, hélices α orientées, hélices α désorientées et feuilles β).

Les spectres Raman des protéines étudiées au cours des différentes étapes de purification (natives, précipitées au TCA, lavées à l'éthanol/HCl et reprises dans le tampon d'analyse) sont présentés dans les figures 87, 89 et 91. La décomposition de la bande amide I a été réalisée avec le logiciel LabSpec 5 propre au spectromètre Raman et permet de générer 4 contributions principales (numérotées de 1 à 4 pour le lysozyme, figure 88, de 6 à 9 pour la HSA, figure 90 et de 10 à 13 pour la Rnase, figure 92) : les structures désordonnées (*uo*; 1636-1639 cm^{-1}), les hélices α orientées (*ho*; 1656-1657 cm^{-1}), les hélices α désorientées et les feuilles β (*hu + sh*; 1670-1672 cm^{-1}), les structures secondaires de type *turns* (*tu* ; 1684-1686 cm^{-1}) (David et al., 2008; Sane et al., 1999).

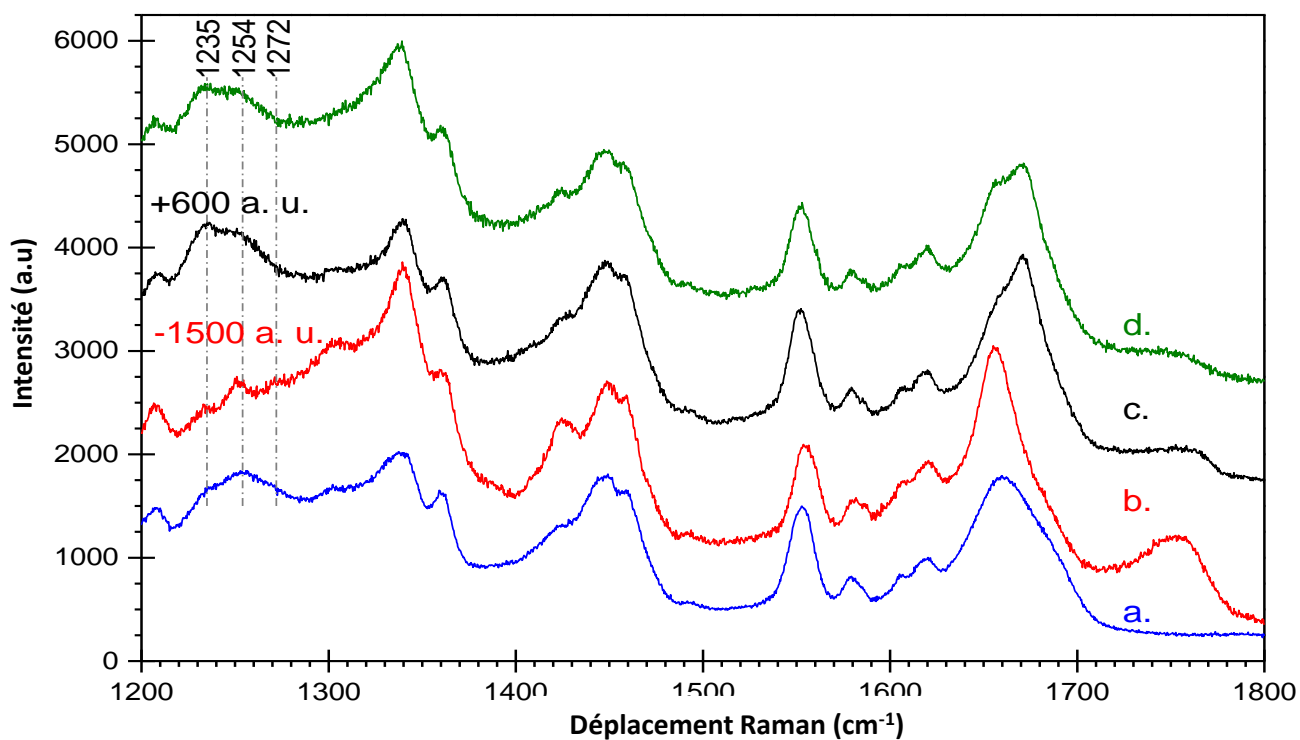


Figure 87. Représentation du spectre Raman du lysozyme à différentes étapes de purification (entre 1200-1800 cm^{-1}). a. Lysozyme natif (bleu); b. Lysozyme précipité au TCA (rouge); c. Lysozyme lavé à l'éthanol/HCl (noir) ; Lysozyme dans le tampon Tris-bicarbonate d'ammonium (vert)

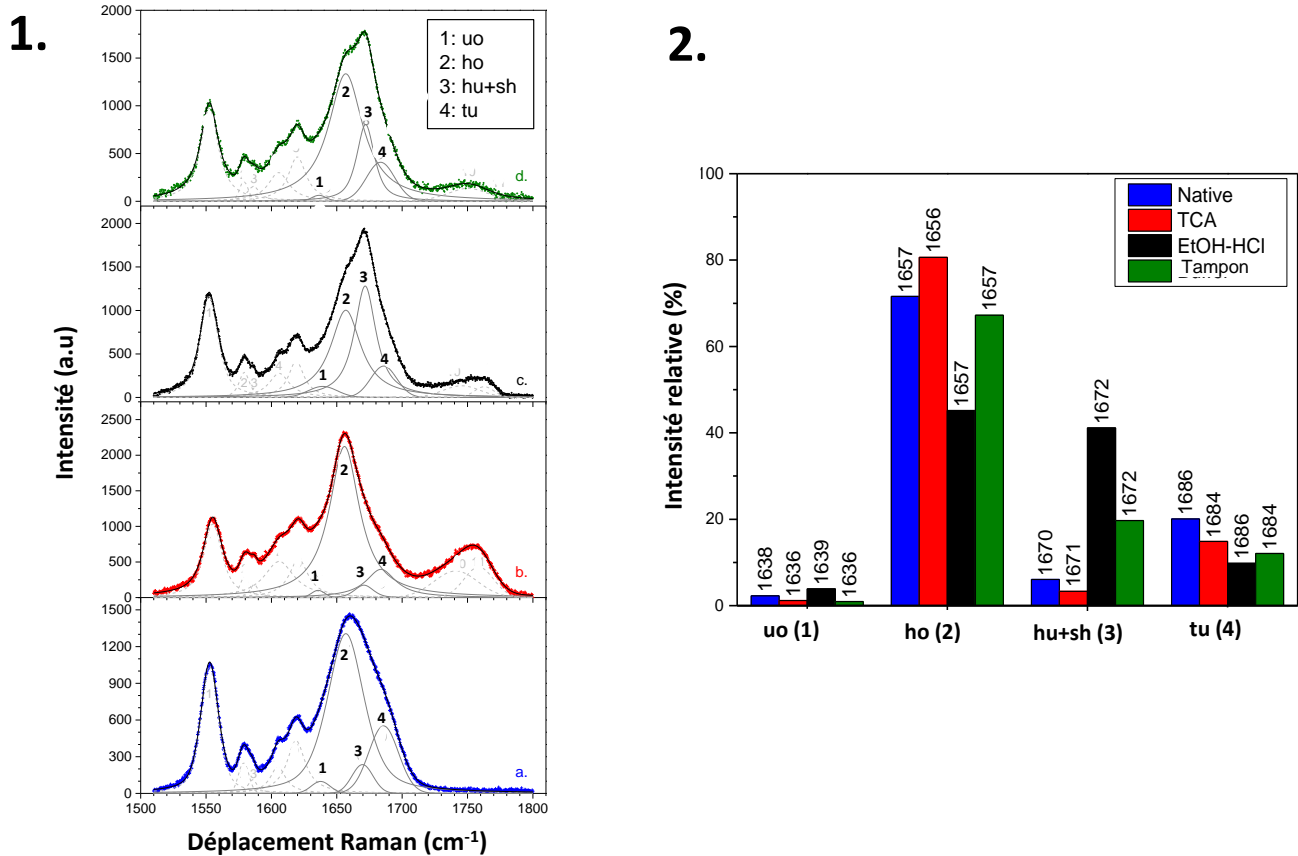


Figure 88. 1. Analyse Raman de la bande amide I (bandes de 1-4) à différentes étapes de purification. a. Lysozyme natif (bleu); b. Lysozyme précipité au TCA (rouge); c. Lysozyme lavé à l'éthanol/HCl (noir) ; d. Lysozyme dans le tampon Tris-bicarbonate d'ammonium (vert); 2. Distribution des différentes contributions (bandes de 1-4) à différentes étapes de purification

Le lysozyme est formé de l'assemblage de 129 acides aminés qui se replie principalement en hélices α (40 à 42%) associées à un feuillet β anti-parallèle et d'un β *turn* (L.-Z. Wu et al., 2008). Le spectre Raman du lysozyme natif (figure 87.a) présente une bande large centrée au niveau de la région caractéristique de la bande amide I à 1659 cm⁻¹ qui correspond principalement à la vibration d'élongation du groupement carbonyle C=O du squelette peptidique, caractéristique des hélices α . La précipitation du lysozyme au TCA entraîne une modification de cette bande amide I avec un glissement vers les basses fréquences (1656 cm⁻¹) et une diminution de sa largeur (figure 87.b). Comme le montre les figures 88.1-b et 88.2, une augmentation des structures hélices α ordonnées (bande 2 ; *ho*) associée à une diminution des structures type β *turns* (bande 4 ; *tu*) et des hélices désorientées (bande 3, *hu + sh*) sont également observées. Ces résultats sont en concordance avec des études antérieures qui

Chapitre II : Développement d'une approche simple pour la purification des protéines

ont démontré qu'en présence de TCA, les protéines sont caractérisées par des structures en hélices α , parfaitement ordonnées, rigides et compactes (Kumar et al., 1994; Rajalingam et al., 2009). Dans cette conformation les protéines exposent leurs groupements hydrophobes vers l'extérieur (milieu environnant), ce qui les rend probablement partiellement solubles dans certains solvants organiques (acétone, éthanol). L'étape de lavage à l'éthanol/HCl (figures 87.c et 88.c) entraîne une diminution de la bande centrée à 1656 cm^{-1} avec un glissement vers les hautes fréquences (1672 cm^{-1}). Ce glissement indique une diminution importante des structures hélices ordonnées (bande 2 ; *ho*) et une transformation d'un état *molten globule* vers un état déplié caractérisé principalement par la présence d'hélices α désorientées et de feuillets β (bande 3 ; *hu + sh*).

En ce qui concerne les deux autres protéines (HSA et Rnase), les changements conformationnels sont comparables à ceux obtenus avec le lysozyme. La HSA est formée exclusivement d'hélices α (figure 89.a et 90.1-a) (Rygula et al., 2013) tandis que la Rnase est caractérisée par une prédominance de feuillets β (figures 91.a et 92.1-a , bande à 1672 cm^{-1}) (Ronald T. Raines, 1998). D'après les figures 90.1-c, 92.1-c l'éthanol/HCl induit une diminution de l'intensité la bande centrée à 1657 cm^{-1} , avec un glissement vers les hautes fréquences. Ce résultat est associé également à une diminution des structures hélices α ordonnées (bande 7 pour la HSA et bande 11 pour la Rnase ; *ho*), converties en feuillets β ou structures désordonnées. Tout comme le lysozyme, l'éthanol/HCl déstructure l'état *molten globule* de la HSA et de la Rnase en une structure riche en structures α désorientées et feuillets β (bande 8 pour la HSA et bande 12 pour la Rnase ; *hu + sh*).

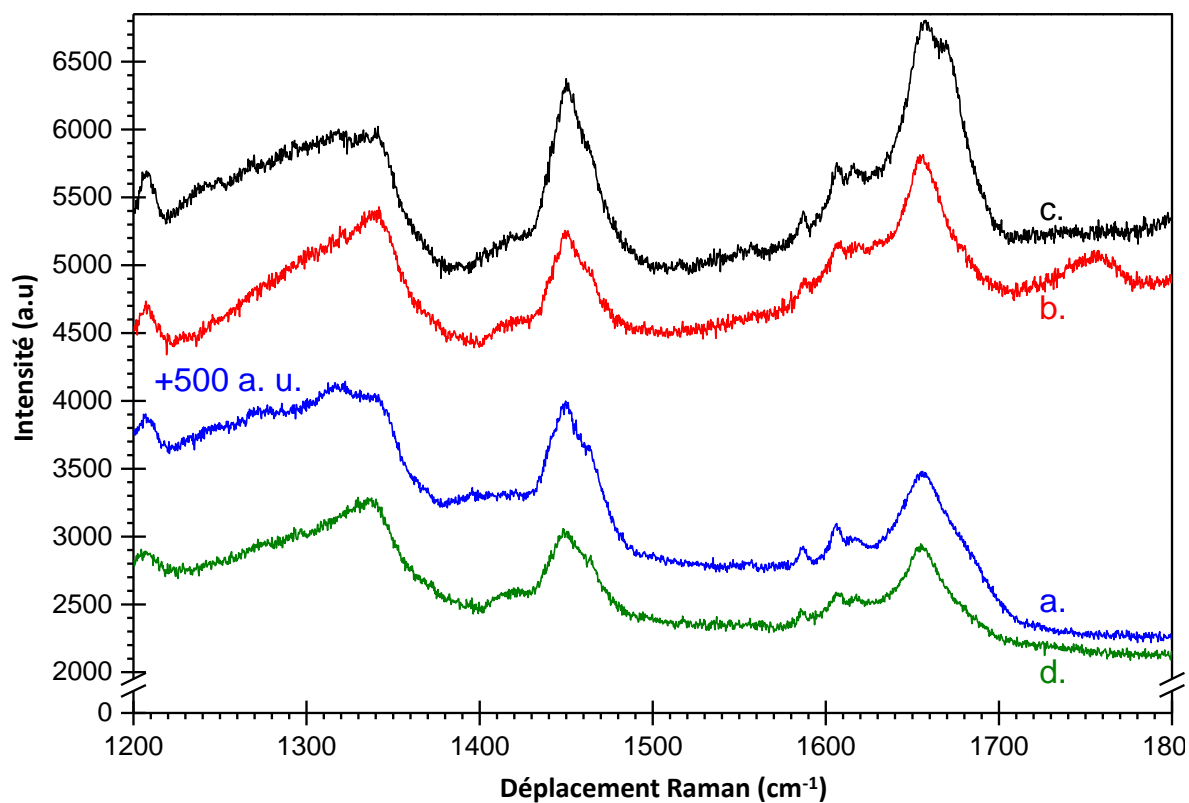


Figure 89. Représentation du spectre Raman du HSA à différentes étapes de purification (entre 1200-1800 cm^{-1}). a. HSA native (bleu); b. HSA précipitée au TCA (rouge); c. HSA lavée à l'éthanol/HCl (noir) ; HSA dans le tampon Tris-bicarbonate d'ammonium (vert)

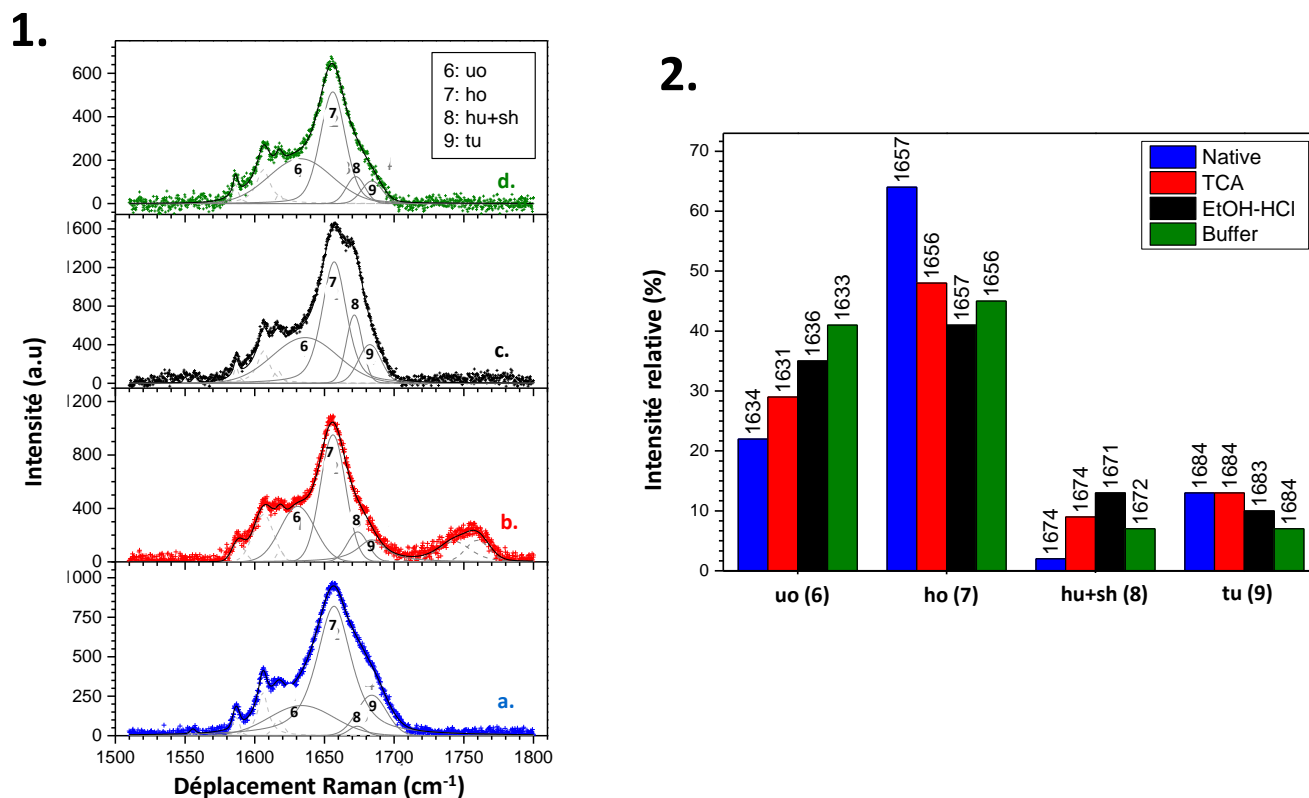


Figure 90. 1. Analyse Raman de la bande amide I (bandes de 6-9) à différentes étapes de purification. a. HSA native (bleu); b. HSA précipitée au TCA (rouge); c. HSA lavée à l'éthanol/HCl (noir); d. HSA dans le tampon Tris-bicarbonate d'ammonium (vert). 2. Distribution des différentes contributions (bandes de 6-9) à différentes étapes de purification

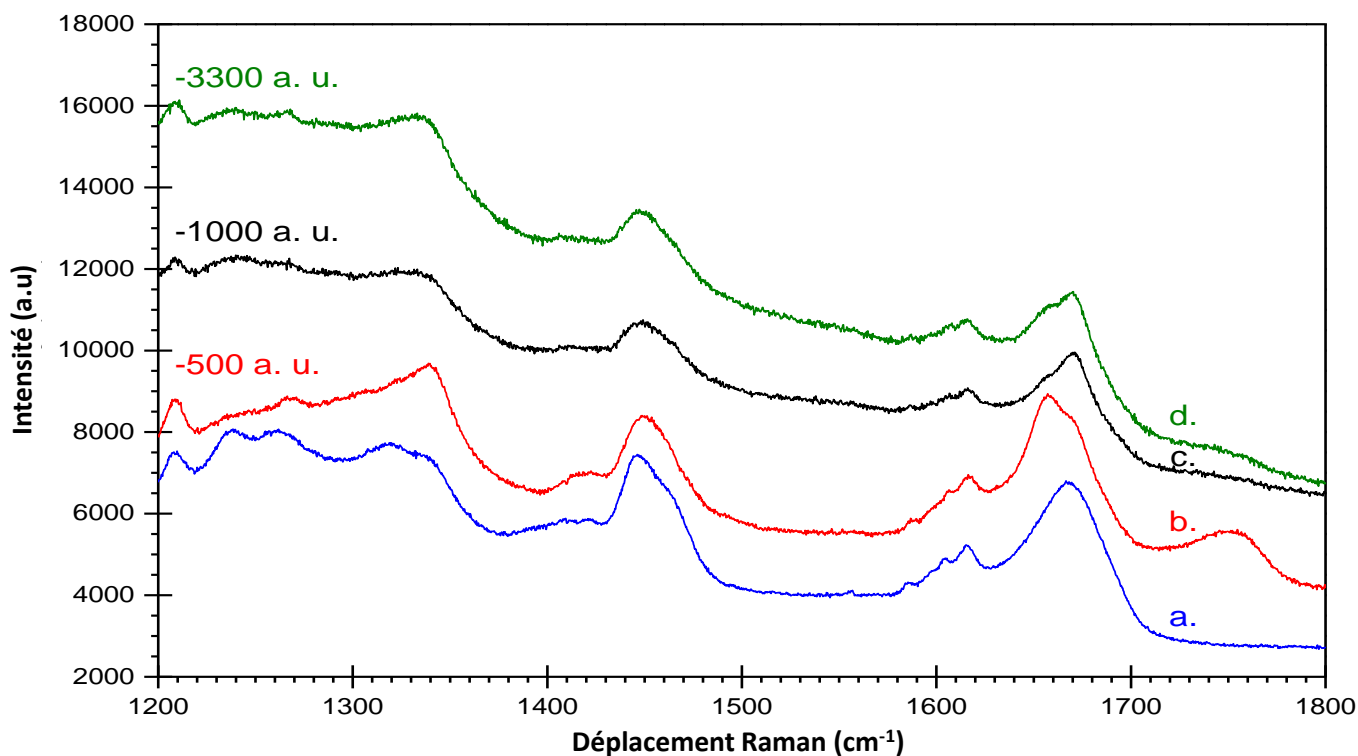


Figure 91. Représentation du spectre Raman la Rnase à différentes étapes de purification (entre 1200-1800 cm^{-1}). a. Rnase native (noir); b. Rnase précipitée au TCA (rouge); c. Rnase lavée à l'éthanol/HCl (bleu) ;d. Rnase dans le tampon Tris-bicarbonate d'ammonium (vert)

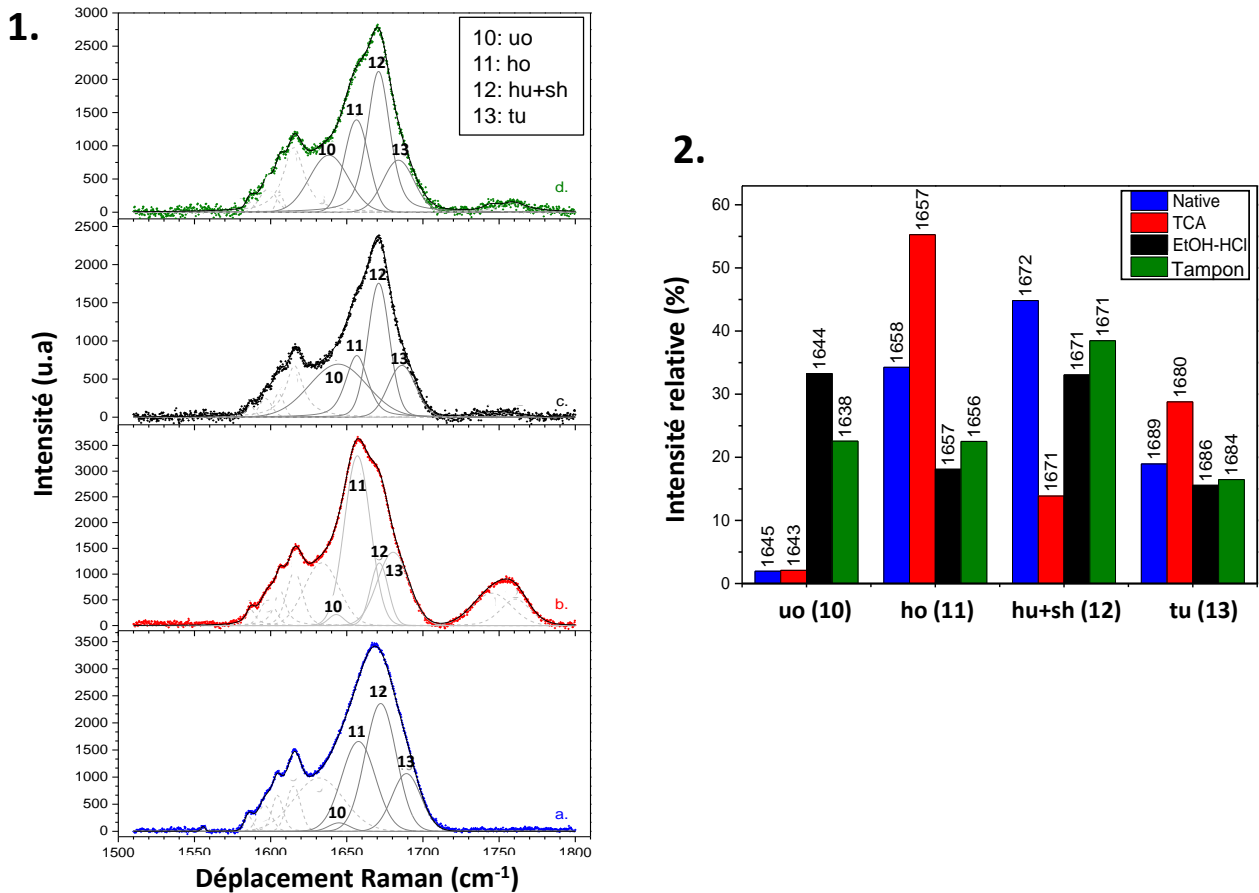


Figure 92. 1. Analyse Raman de la bande amide I (bandes de 10-13) à différentes étapes de purification. a. Rnase native (bleu); b. Rnase précipitée au TCA (rouge); c. Rnase lavée à l'éthanol/HCl (noir); d. Rnase dans le tampon Tris-bicarbonate d'ammonium (vert). 2. Distribution des différentes contributions (bandes de 10-13) à différentes étapes de purification

Afin de valider cette approche éthanol/HCl, les spectres Raman du lysozyme et de la HSA après précipitation au TCA et lavage à l'acétone ont également été étudiés. Le déplacement Raman a été enregistré dans la région de l'amide I. Les résultats sont présentés dans la figure 93.

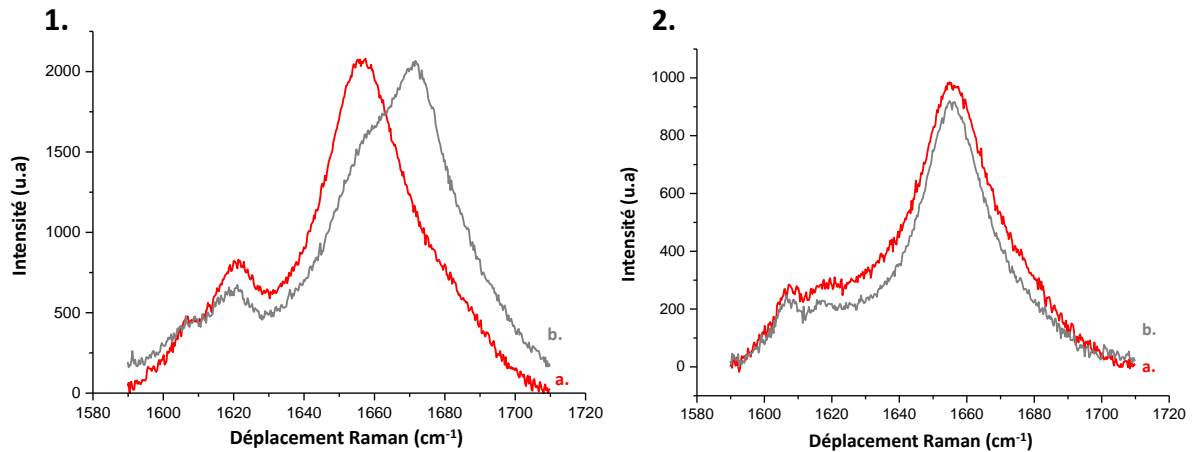


Figure 93. Représentation du spectre Raman (région de l'amide I). 1.a Lysozyme précipitée au TCA (rouge); b. Lysozyme lavée à l'acétone (gris). 2.a HSA précipitée au TCA (rouge); b. HSA lavée à l'acétone (gris)

L'analyse de l'amide I a montré que l'acétone entraîne un changement de la structure secondaire du lysozyme précipité au TCA. Cette modification est similaire à celle observée après lavage à l'éthanol/HCl (apparition d'une bande à 1672 cm^{-1} qui correspond à la présence d'hélices α désorientées et de feuillets β (bande 8 ; $hu + sh$)). Néanmoins, l'acétone ne semble avoir aucun effet sur la HSA. Dans ce cas, la protéine conserve son état MG induit par précipitation au TCA. Contrairement à l'éthanol/HCl qui induit quelle que soit la structure de la protéine une dénaturation de son état MG, l'acétone semble avoir un effet très variable sur les structures protéiques ce qui confirme les rendements protéiques obtenus en solutions standards (figures 86).

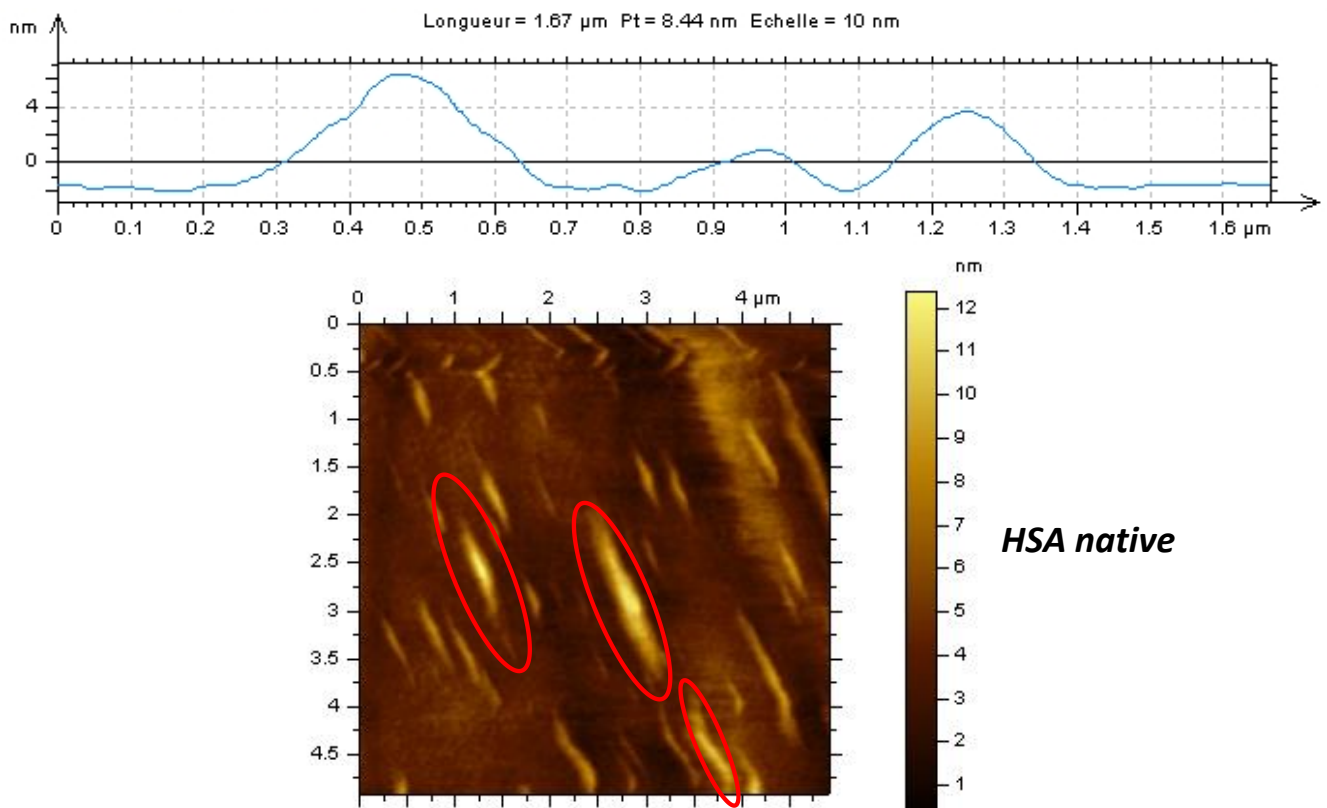
2.3. Microscopie à force atomique (AFM)

L'AFM explore la surface d'échantillons organiques ou inorganiques à l'aide d'une pointe effilée portée par un ressort très souple. La précision des déplacements de la pointe, dans les trois plans de l'espace, permet de visualiser aussi bien des structures biologiques complexes que des molécules uniques sans les déformer au cours d'une analyse AFM. L'échantillon doit être déposé puis adsorbé de façon homogène sur un support spécifique plat. Plusieurs supports sont couramment employés en AFM, les plus connus sont le mica (support hydrophile) et le graphite orienté HOPG (support hydrophobe). L'AFM a déjà été appliquée pour observer les structures en feuillets β caractéristiques des fibres amyloïdes dans la maladie du prion

Chapitre II : Développement d'une approche simple pour la purification des protéines

(Anderson et al., 2006) ou encore pour apprécier la formation de l'état *molten globule* chez la β -lactoglobuline induite sous l'action d'une élévation de la température (Phoon et al., 2013).

Dans notre cas d'étude, l'AFM a permis de visualiser en 3D la formation de l'état *molten globule* après précipitation au TCA et d'évaluer l'effet de l'éthanol/HCl sur cet état. La visualisation des images AFM a été réalisée en mode *tapping* (figure 94).



Chapitre II : Développement d'une approche simple pour la purification des protéines

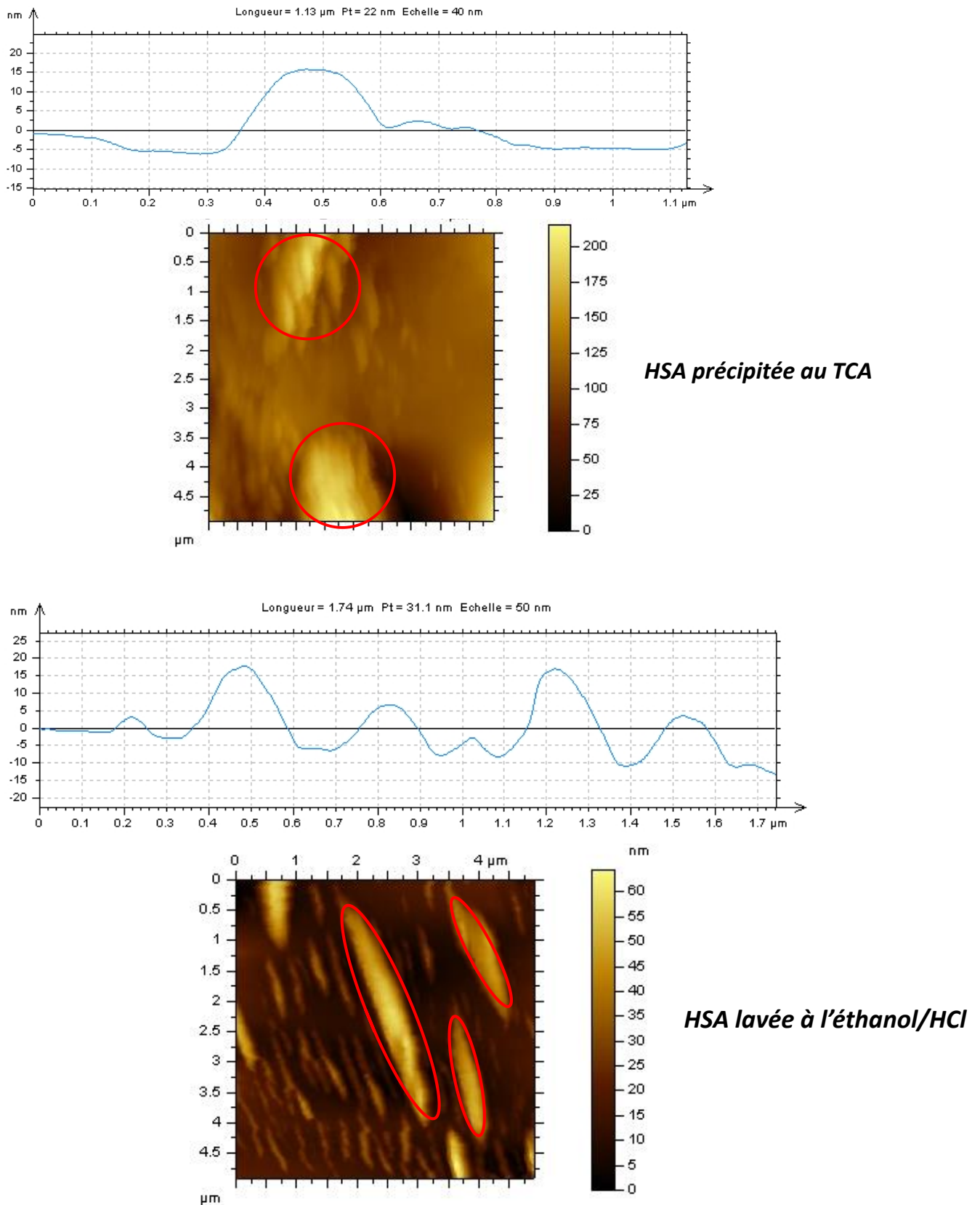


Figure 94. Image AFM de la HSA à différentes étapes de purification

Chapitre II : Développement d'une approche simple pour la purification des protéines

La figure 94 montre que la précipitation de la HSA au TCA entraîne la formation de gros agrégats de diamètre d'environ 250 nm et d'une hauteur de 15 nm. L'image est toutefois mal résolue probablement en raison de la mauvaise adsorption de cet agrégat hydrophobe sur le support hydrophile (Mica). Après lavage à l'éthanol/HCl, l'image AFM est plus nette. Le diamètre de la protéine tend à diminuer, environ 200 nm avec des hauteurs variables (entre 5 et 15 nm). Cette image obtenue après lavage à l'éthanol/HCl est comparable à celle de la protéine native (soluble). Malgré la mauvaise résolution des images AFM, nous pouvons conclure que l'éthanol/HCl entraîne une modification de l'état *molten globule* compact, qui reprend une forme proche de la native comme l'a montré le Raman.

3. Etude de l'efficacité du protocole de purification sur matrices réelles

3.1. Etude sur des protéines standards déposées sur matrice

Après avoir étudié au niveau moléculaire l'effet de l'éthanol/HCl sur les protéines précipitées au TCA, nous avons démontré l'efficacité de cette approche sur des matrices complexes de nature et de composition très différentes (biofilm, foie de souris et sol du Deffend) (cf. matériels et méthodes). Elle a ensuite été comparée à des protocoles de purification « hors gel » publiés (Chourey et al., 2010; Huang et al., 2014; Song et al., 2016). Dans leur ensemble, ces approches utilisent une précipitation au TCA suivie de plusieurs étapes de lavage (en présence d'acétone ou d'eau). Chacun des trois protocoles a démontré une efficacité pour caractériser le protéome des matrices étudiées.

Nous avons alors dopé les 4 protéines cibles (HSA, lysozyme, Rnase et exoglucanase-1) sur les milieux étudiés, puis effectué la purification des protéines avec une précipitation au TCA suivie d'un lavage à l'éthanol/HCl ou suivant les protocoles de référence cités ci-dessus. Les résultats relatifs à la comparaison de la purification en fonction de la matrice sont présentés sur la figure 95.

Chapitre II : Développement d'une approche simple pour la purification des protéines

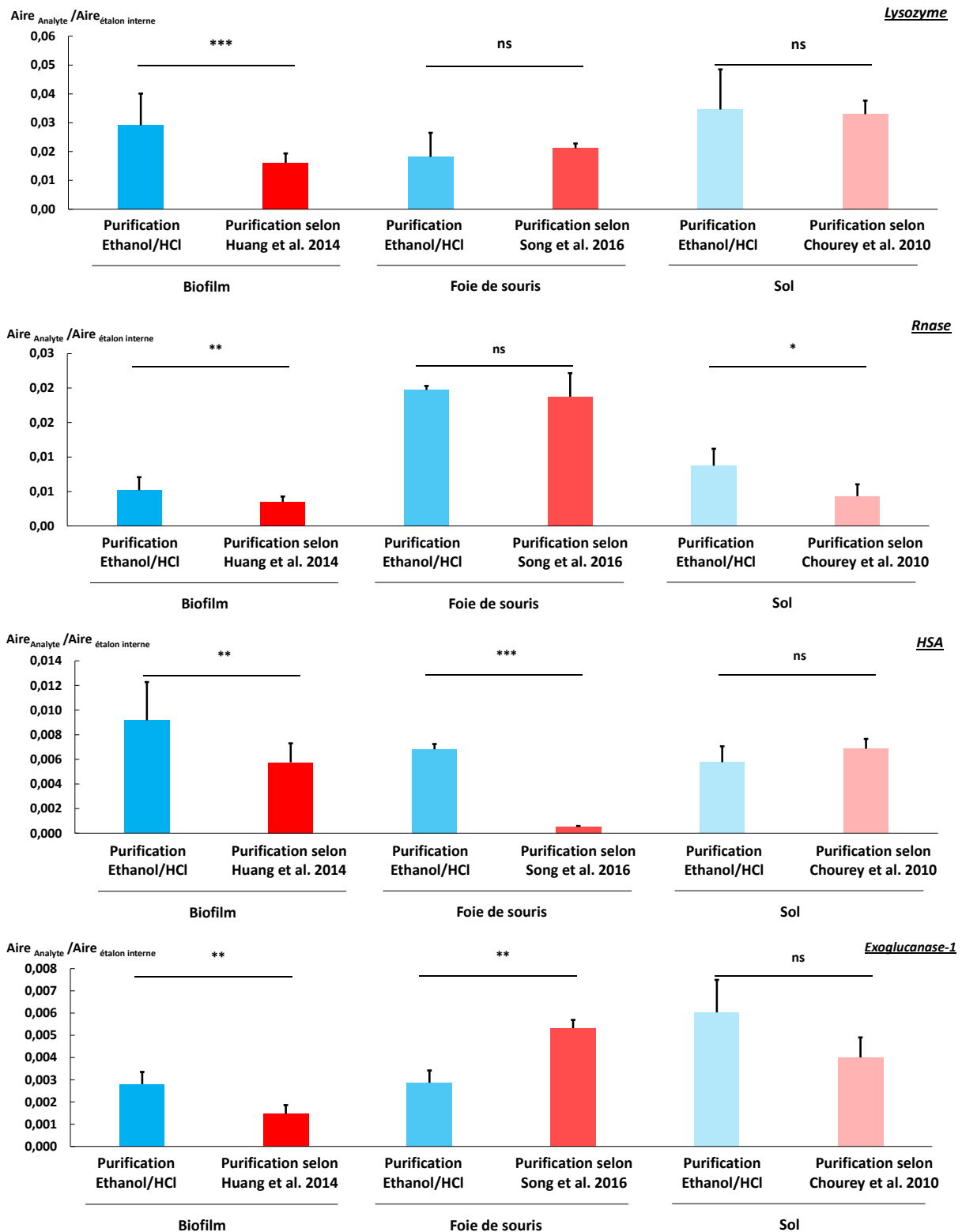


Figure 95. Récupération des protéines dopées en fonction des deux approches de purification : purification à l'éthanol/HCl et purification selon le protocole de la littérature. * $P < 0,1$, ** $P < 0,05$ and *** $P < 0,01$; analyse de variance (ANOVA) avec analyse de la petite différence significative (test LSD, Least Significant Difference) ($\alpha < 5\%$)

Chapitre II : Développement d'une approche simple pour la purification des protéines

De façon générale, nous pouvons constater que les quantités des protéines purifiées sont plus importantes après purification à l'éthanol/HCl. Pour le biofilm, la quantité des protéines récupérées est deux fois supérieure à celle obtenue après purification selon le protocole de Huang et al. (2014). Concernant les protéines dopées sur un foie souris, les quantités récupérées du lysozyme et de la Rnase sont similaires, tandis que la récupération de l'exoglucanase-1 est plus importante dans l'extrait purifié suivant l'approche de Song et al. (2016). Néanmoins, l'éthanol/HCl permet de récupérer la HSA en quantité significativement plus importante par rapport à la méthode publiée. Pour le sol, les deux approches testées aboutissent à des résultats semblables, sauf pour la Rnase, pour laquelle la purification avec l'éthanol/HCl est plus efficace.

Étant donné que les protocoles de la bibliographie incluent généralement deux à trois étapes de lavage, nous avons également voulu vérifier que plusieurs lavages à l'éthanol/HCl n'entraînaient pas de pertes significatives des protéines cibles en cas de purification d'extrait très riche en composés interférents. L'essai a été réalisé sur le biofilm. Comme le montre la figure 96, l'application de deux ou trois lavages n'entraîne pas de pertes significatives des protéines cibles.

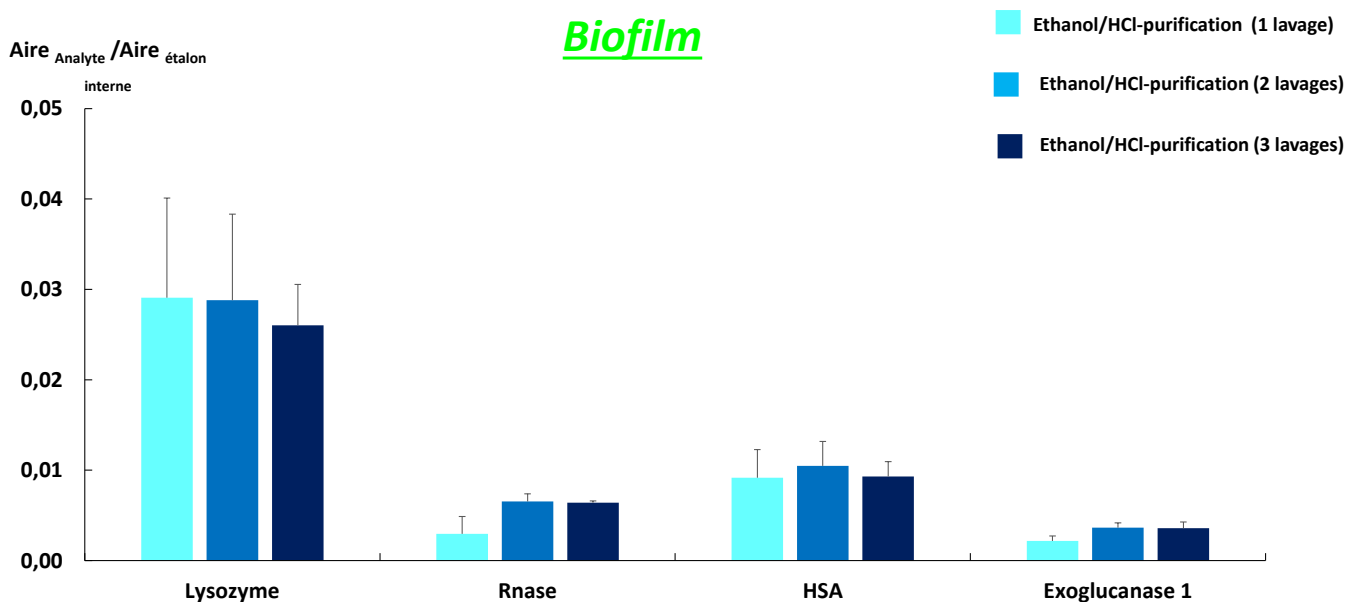


Figure 96. Récupération des protéines dopées en fonction du nombre de lavages à l'éthanol/HCl

Chapitre II : Développement d'une approche simple pour la purification des protéines

Le protocole de purification développé lors de cette étude repose sur une seule et unique étape de lavage et de solubilisation des protéines. A contrario, les approches publiées dans la littérature sont relativement longues et multiplexes (entre 2 et 4 étapes de lavage ajoutées aux étapes de solubilisation au moyen des agents dénaturants (SDS, urée...)). L'utilisation de l'éthanol/HCl a donc permis d'éliminer de manière efficace l'ensemble des agents utilisés pour l'extraction (SDS, CHAPS, urée etc) et la purification (TCA) des protéines tout en facilitant leur solubilisation dans un tampon aqueux avant digestion trypsique.

3.2. Analyse des protéines endogènes aux matrices

Afin d'aller plus loin dans la démonstration de l'efficacité de l'agent de lavage, des recherches à partir des spectres MS2 contre la banque de données *SwissProt* ont été réalisées à partir de ces mêmes échantillons réels (sol, biofilm et foie de souris), en utilisant l'algorithme *Mascot distiller* (v. 2.6.0). Après interrogation des banques de données de protéomique, un classement probabiliste des protéines endogènes extraites et purifiées des matrices est établi en fonction du nombre des peptides attribués à chaque protéine. On obtient ainsi un score qui prend en compte la probabilité que la séquence peptidique ait été identifiée par hasard, ainsi que le pourcentage de couverture de la séquence protéique. Dans un premier temps, nous avons déterminé l'ensemble des protéines identifiées dans chaque matrice, et ce quelle que soit la valeur de leurs scores. D'après les résultats obtenus par *Mascot* (tableau 13), environ 1000 peptides ont été détectés suivant l'approche de purification de la littérature ou le protocole développé (éthanol/HCl). Seules 44 et 48 protéines ont été détectées dans le sol. Ce résultat peut s'expliquer notamment par les quantités limitées de protéines dans les sols (0,2 et 16,3 milligrammes de protéines par grammes de matrice) (Greenfield, 1972; Jan et al., 2009), mais également par le manque de séquences ainsi que par la faible annotation des protéines fonctionnelles dans les bases de données protéiques (Bastida et al., 2009).

Tableau 13. Nombre de peptides et de protéines identifiés dans les matrices complexes

	Nombre de peptides identifiés	Nombre de protéines identifiées
Foie-protocole littérature	1085	153
Foie-protocole développé	952	136
Biofilm-protocole littérature	949	73
Biofilm-protocole développé	948	67
Sol-protocole littérature	913	44
Sol-protocole développé	871	48

Dans un deuxième temps, nous avons évalué l'efficacité de l'éthanol/HCl en étudiant uniquement les protéines communes identifiées dans les deux protocoles de purification. Cette évaluation a été réalisée en comparant les scores *Mascot* obtenus pour chaque protéine. Dans cette étude, un seuil minimal de score *Mascot* supérieur à 67 a été considéré comme significatif pour l'identification des protéines (Perkins et al., 1999). La liste des protéines endogènes communes identifiées dans chacune des trois matrices est présentée dans le tableau 14.

Tableau 14. Liste des protéines endogènes communes identifiées dans chaque matrice après purification selon le protocole développé (éthanol/HCl) et le protocole publié dans la littérature

Chapitre II : Développement d'une approche simple pour la purification des protéines

Matrice	Protéines	Origine phylogénétique	Localisation	Couverture de la séquence protéique (%)		Score		Poids moléculaire (Da)	Indice Gravy
				Protocole développé (éthanol/HCl)	Protocole publié dans la littérature	Protocole développé (éthanol/HCl)	Protocole publié dans la littérature		
Sol	Endoglucanase EG-II	<i>Hypocrea jecorina</i>	Région extracellulaire	18	19	161	251	44883	-0,19
	Xyloglucanase	<i>Hypocrea jecorina</i>	Région extracellulaire	1	1	76	114	87307	-0,21
Biofilm	C-phycoérythrine alpha chaîne	<i>Microchaete diplosiphon</i>	Cytoplasme (thylakoïde)	29	29	269	239	17786	-0,15
	R-phycoérythrine alpha chaîne	<i>Porphyra purpurea</i>	Membrane (chloroplaste)	20	17	168	119	17972	-0,19
	C-phycoeryanine alpha chaîne	<i>Synechococcus sp.</i>	Cytoplasme (thylakoïde)	17	17	181	177	17335	-0,11
	C-phycoérythrine alpha chaîne	<i>Synechocystis sp.</i>	Cytoplasme (thylakoïde)	20	20	209	176	17756	-0,12
	Allophycocyanine alpha chaîne 1	<i>Microchaete diplosiphon</i>	Cytoplasme (thylakoïde)	11	11	76	84	17411	-0,14
	B-phycoérythrine beta chaîne	<i>Porphyridium purpureum</i>	Membrane (chloroplaste)	21	20	117	183	18884	0,25
	C-phycoérythrine beta chaîne	<i>Microchaete diplosiphon</i>	Cytoplasme (thylakoïde)	21	16	138	85	19568	0,21
	R-phycoérythrine beta chaîne	<i>Pyropia haitanensis</i>	Membrane (chloroplaste)	23	28	129	144	18810	0,26
	C-phycoeryanine-1 beta chaîne	<i>Microchaete diplosiphon</i>	Cytoplasme (thylakoïde)	16	12	64	122	18080	0,17
	Allophycocyanine sous-unité alpha 1	<i>Nostoc sp.</i>	Cytoplasme (thylakoïde)	17	19	99	112	17392	-0,09
	C-phycoeryanine beta chaîne	<i>Aglaothamnion neglectum</i>	Membrane (chloroplaste)	11	12	112	111	18290	0,09
	Ribulose bisphosphate carboxylase large chaîne	<i>Trichodesmium erythraeum</i>	NI	5	8	90	122	53615	-0,32

Chapitre II : Développement d'une approche simple pour la purification des protéines

Allophycocyanine alpha chaîne	<i>Anabaena cylindrica</i>	Cytoplasme (thylakoïde)	6	11	84	83	17128	0,01
C-phycoérythrine alpha chaîne	<i>Pseudanabaena tenuis</i>	Cytoplasme (thylakoïde)	18	18	144	126	17780	-0,24
Protéine photosystème II CP47	<i>Odontella sinensis</i>	Membrane (chloroplaste)	8	8	117	114	56436	0,08
Ribulose bisphosphate carboxylase large chaîne	<i>Cyanothece sp.</i>	NI	9	6	94	89	53531	-0,27
Ribulose bisphosphate carboxylase large chaîne (Fragment)	<i>Calyptrosphaera sphaeroidea</i>	Chloroplaste	5	9	90	107	50919	-0,10
Ribulose bisphosphate carboxylase large chaîne	<i>Gracilaria tenuistipitata var. liui</i>	Chloroplaste	8	10	111	132	54442	-0,10
Ribulose bisphosphate carboxylase large chaîne	<i>Cylindrotheca sp.</i>	Chloroplaste	6	6	109	108	54400	-0,12
Allophycocyanine beta chaîne	<i>Cyanidium caldarium</i>	Membrane (chloroplaste)	13	16	94	83	17574	-0,04
Ribulose bisphosphate carboxylase small chain	<i>Antithamnion sp.</i>	Chloroplaste	5	5	72	72	16247	-0,58
Protéine du mécanisme de concentration du CO ₂ (Ccmk)	<i>Synechocystis sp.</i>	NI	18	29	71	72	11128	-0,19
R-phycoerythrin beta chain	<i>Aglaothamnion neglectum</i>	Membrane (chloroplaste)	7	7	100	69	18710	0,27
Ribulose bisphosphate carboxylase large chaîne (Fragment)	<i>Haptolina hirta</i>	Chloroplaste	9	10	141	139	51098	-0,11
Ribulose bisphosphate carboxylase large chaîne	<i>Antithamnion sp.</i>	Chloroplaste	7	7	117	113	54372	-0,12
Allophycocyanine beta chaîne	<i>Thermosynechococcus elongatus</i>	Cytoplasme (thylakoïde)	18	18	103	121	17462	0,10

Chapitre II : Développement d'une approche simple pour la purification des protéines

	Apoprotéine A2 photosystème I P700	<i>Gloeobacter violaceus</i>	Membrane	2	2	78	75	96126	0,15
	Phycobiliprotéine ApcE	<i>Aglaothamnion neglectum</i>	Membrane (chloroplaste)	1	1	73	72	101319	-0,23
	Ribulose bisphosphate carboxylase large chaîne	<i>Synechocystis sp.</i>	NI	6	6	120	117	53084	-0,29
	Allophycocyanine beta chaîne	<i>Galdieria sulphuraria</i>	Membrane (chloroplaste)	16	16	96	73	17536	0,02
Foie de souris	Carbamyl-phosphate synthétase	<i>Mus musculus</i>	Noyau, mitochondrie	39	33	1637	1268	165711	-0,12
	Arginase-1	<i>Mus musculus</i>	Cytoplasme	29	35	300	310	34957	-0,19
	Sélénium-binding protéine	<i>Mus musculus</i>	Cytoplasme, noyau, membrane	31	28	526	405	53147	-0,31
	Argininosuccinate synthétase	<i>Mus musculus</i>	Cytoplasme	32	15	429	191	46840	-0,11
	Glyceraldehyde-3- phosphate déshydrogénase	<i>Mus musculus</i>	Mitochondrie	31	32	321	298	36072	-0,04
	10-formyl- tétrahydrofolate déshydrogénase	<i>Mus musculus</i>	Cytoplasme	9	17	139	361	99502	-0,36
	3-ketoacyl-CoA thiolase, mitochondriale	<i>Mus musculus</i>	Région extracellulaire	10	20	137	216	42260	-0,38
	Albumine sérique	<i>Mus musculus</i>	Noyau, cytosquelette, cytoplasme	15	18	327	349	70700	-0,09
	Alcool déshydrogénase 1	<i>Mus musculus</i>	Cytoplasme	19	29	161	212	40601	0,20
	Aspartate amino- transférase, mitochondrial	<i>Mus musculus</i>	Membrane	15	16	231	215	47780	-0,23
Carboxylesterase 3B	<i>Mus musculus</i>	Réticulum endoplasmique	12	14	201	183	63712	-0,12	
Glycine N-methyl- transférase	<i>Mus musculus</i>	Cytoplasme	29	19	131	127	33110	-0,25	

Chapitre II : Développement d'une approche simple pour la purification des protéines

Cytochrome P450 2D10	<i>Mus musculus</i>	Membrane	9	2	100	123	57539	-0,06
Aspartate amino-transférase, cytoplasmique	<i>Mus musculus</i>	Cytoplasme	7	13	112	115	46504	-0,25
Adénosylhomocystéina se	<i>Mus musculus</i>	Cytoplasme	27	14	335	120	47780	-0,07
Fructose-1,6-bisphosphatase 1	<i>Mus musculus</i>	Cytoplasme	12	16	117	120	37288	-0,12
Carboxylesterase 3A	<i>Mus musculus</i>	Réticulum endoplasmique	13	9	220	139	63677	-0,12
Sarcosine déshydrogénase, mitochondrial	<i>Mus musculus</i>	Mitochondrie	8	6	182	209	102644	-0,25
UDP-glucuronosyltransférase 1-1	<i>Mus musculus</i>	Membrane	4	8	94	141	60749	0,09
Hémoglobine sous-unité beta-1	<i>Mus musculus</i>	Cytoplasme	16	24	111	105	15944	0,08
Enzyme bifonctionnelle, peroxisomal	<i>Mus musculus</i>	Péroxyosome	3	2	98	78	78822	-0,12
Microsomal glutathione S-transférase	<i>Mus musculus</i>	Membrane	17	21	80	87	17597	0,15
Cytochrome P450 2F2	<i>Mus musculus</i>	Membrane	6	7	128	130	56141	-0,13
Pyréthriinoïde hydrolase Ces2a	<i>Mus musculus</i>	NI	9	5	100	76	57539	—
Homogentisate 1,2-dioxygénase	<i>Mus musculus</i>	Région extracellulaire	6	6	81	114	50726	-0,34
Régucalcine	<i>Mus musculus</i>	Cytoplasme	4	13	72	112	33899	-0,28
3-ketoacyl-CoA thiolase B, peroxisomal	<i>Mus musculus</i>	Péroxyosome	13	8	116	84	44481	0,05
Sorbitol déshydrogénase	<i>Mus musculus</i>	Membrane	6	6	90	89	38795	0,06
ATP synthétase sous-unité f, mitochondrial	<i>Mus musculus</i>	Membrane	26	26	70	71	10394	-0,30

Chapitre II : Développement d'une approche simple pour la purification des protéines

ATP synthase sous-unité alpha, mitochondrial	<i>Mus musculus</i>	Membrane	14	10	193	160	59830	-0,10
Urocanate hydratase	<i>Mus musculus</i>	Cytoplasme	2	1	100	76	75227	-0,14
Fumarylacétoacétase	<i>Mus musculus</i>	Région extracellulaire	3	6	75	74	46488	-0,21
Uricase	<i>Mus musculus</i>	Mitochondrie, péroxysome	17	11	157	97	35245	-0,46
Fructose-bisphosphate aldolase B	<i>Mus musculus</i>	Cytosquelette	15	13	180	119	39938	-0,26
UDP-glucuronosyltransférase 2B17	<i>Mus musculus</i>	Membrane	11	6	104	96	61386	-0,03
Pyréthrianoïde hydrolase	<i>Mus musculus</i>	NI	9	7	108	89	62356	-0,08
3-hydroxyanthranilate 3,4-dioxygénase	<i>Mus musculus</i>	Cytoplasme	9	6	90	87	32955	-0,55
Hydroxyméthylglutaryl-CoA synthétase, mitochondriale	<i>Mus musculus</i>	Mitochondrie	7	6	86	70	57300	-0,34
Enzyme trifonctionnelle sous-unité alpha, mitochondriale	<i>Mus musculus</i>	Mitochondrie	9	7	90	81	83302	-0,10
Microsomal triglycéride transfer protéine	<i>Mus musculus</i>	Réticulum endoplasmique	1	1	74	80	99664	-0,16
Cytochrome b-c1 complexe sous-unité 2, mitochondrial	<i>Mus musculus</i>	Membrane	4	4	73	76	48262	-0,06

Chapitre II : Développement d'une approche simple pour la purification des protéines

Les résultats de recherche issus de *Mascot distiller* ont permis d'identifier dans 10 mg de biofilm et 15 mg de foie de souris respectivement, 30 et 41 protéines endogènes communes aux deux protocoles (figure 97). Parmi ces protéines, environ 60 % avaient un score d'identification supérieur lorsque notre protocole de purification a été appliqué.

Seulement deux protéines communes ont été identifiées dans la matrice sol (tableau 14).

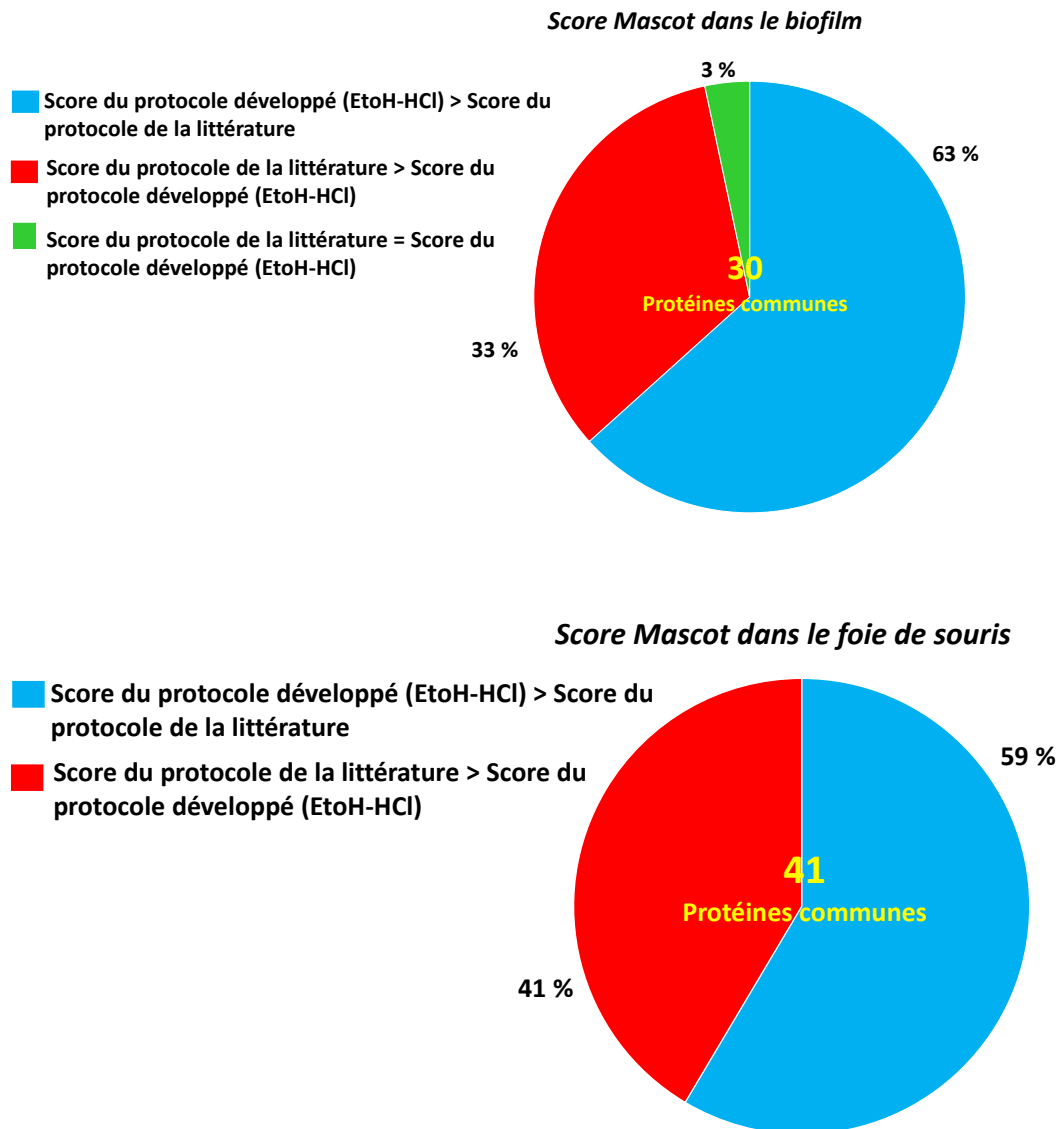
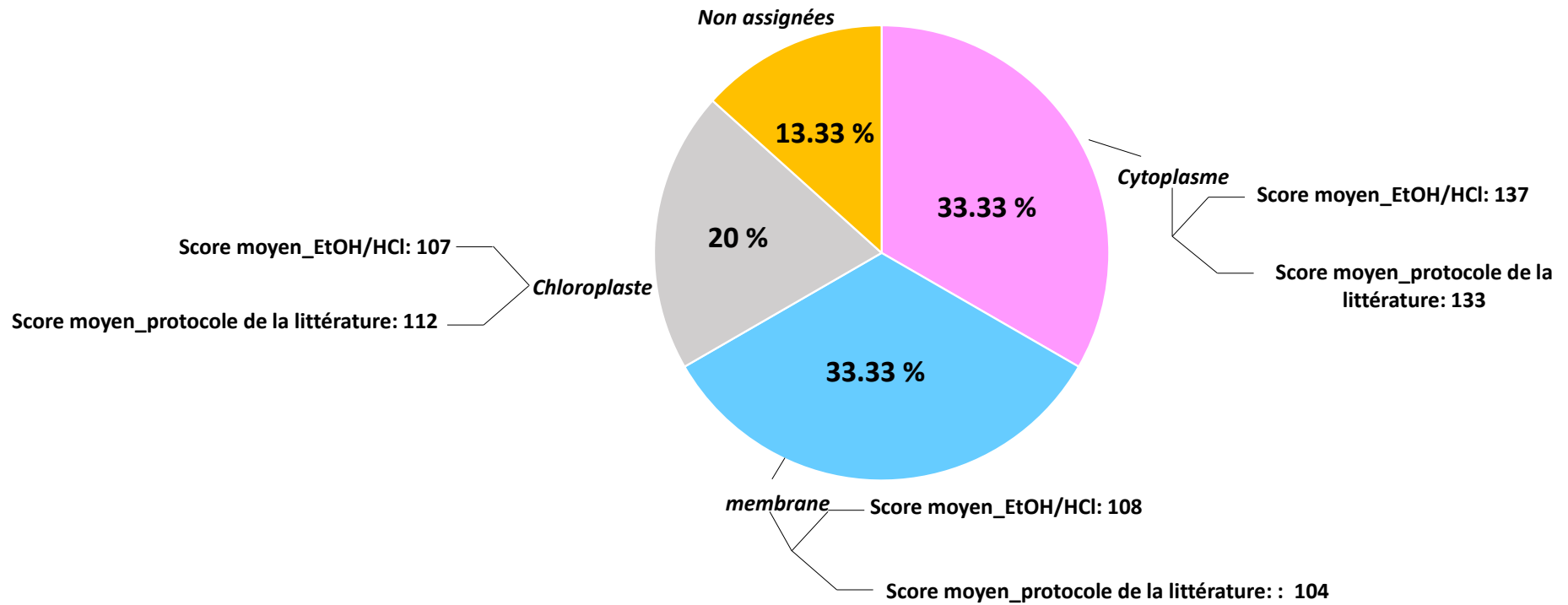


Figure 97. Représentation schématique du pourcentage des protéines communes identifiées dans le biofilm et le foie de souris après purification soit avec l'éthanol/HCl soit avec le protocole publié dans la littérature

Nous sommes par la suite intéressées à la répartition (localisation) des protéines endogènes identifiées par *Mascot*, dans les différents compartiments cellulaires (figure 98).

a.

Matrice biofilm marin



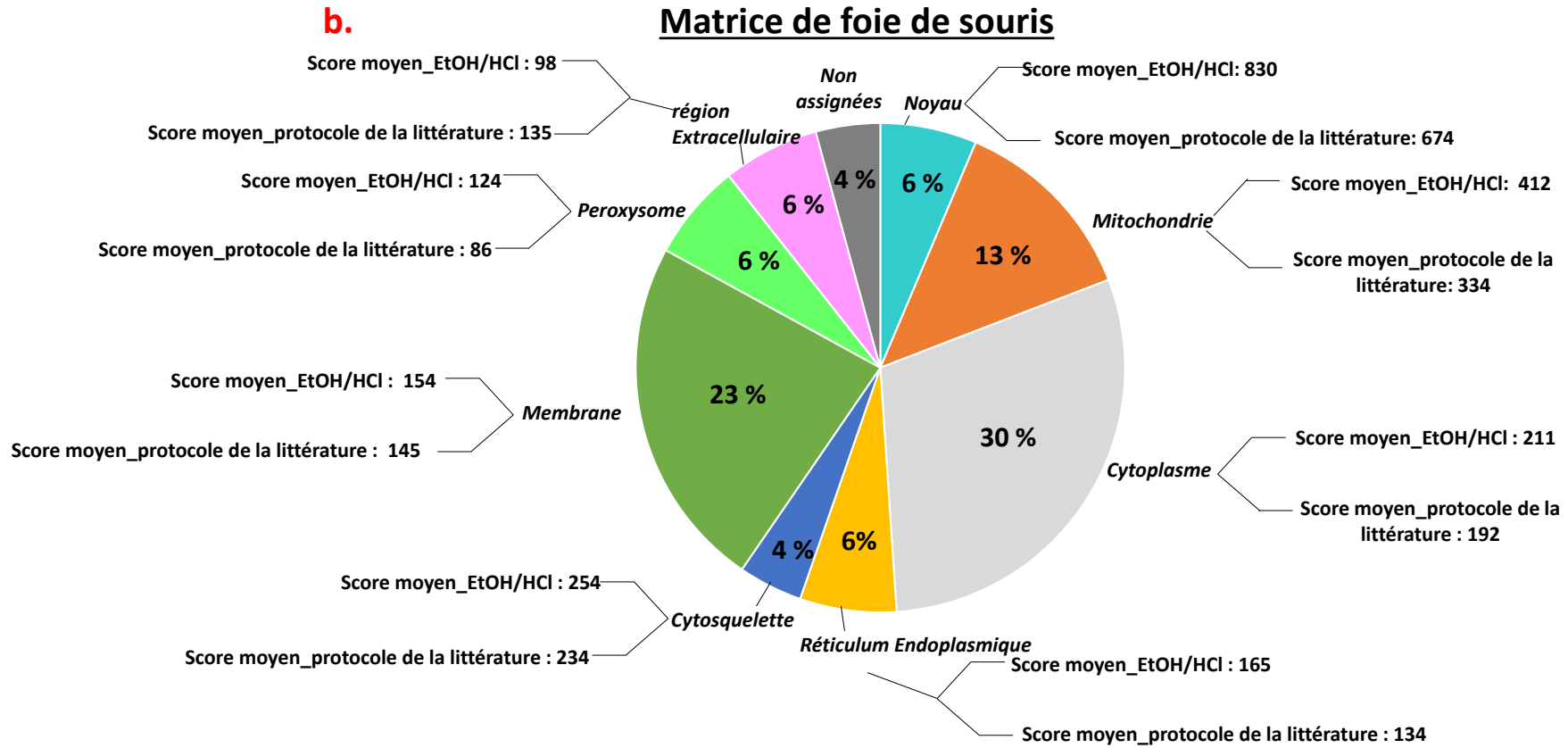


Figure 98. Répartition des protéines endogènes dans les différents compartiments cellulaire (noyau, membrane, mitochondrie, etc.), en fonction du protocole développé (EtOH/HCl : éthanol/HCl) et du protocole publié dans la littérature. a. Répartition cellulaire dans la matrice biofilm marin ; b. Répartition cellulaire dans la matrice de foie de souris

Chapitre II : Développement d'une approche simple pour la purification des protéines

Nous avons calculé pour chaque compartiment cellulaire le score *Mascot* moyen relatif à chaque approche de purification. D'après la figure 98, nous pouvons constater que les protéines purifiées sont localisées majoritairement dans le cytoplasme (environ 30 %) et dans les différents organites cellulaires (entre 20 et 30 %) (Mitochondrie, réticulum endoplasmique, chloroplaste (diatomées marines), etc.), et ce quelle que soit la matrice étudiée (biofilm marin ou foie de souris). Au niveau de ces compartiments cellulaires, le score d'identification *Mascot*, est supérieur lorsque notre approche a été appliquée. En ce qui concerne les protéines membranaires, celles-ci ne représentent que 33 % dans le biofilm (membranes cellulaires et membranes du chloroplaste) et 23 % dans le foie de souris. Ces protéines hydrophobes étant difficiles à solubiliser sont fortement liées à la bicouche lipidique. De nombreuses études utilisent des solutions de détergents ou des agents amphiphiles, après précipitation de ces protéines afin de faciliter leur solubilisation (Doucette et al., 2014; Popot, 2010). En outre, il est intéressant de noter que les scores *Mascot* des protéines membranaires sont meilleurs après purification à l'éthanol/HCl et ce quelle que soit la matrice utilisée.

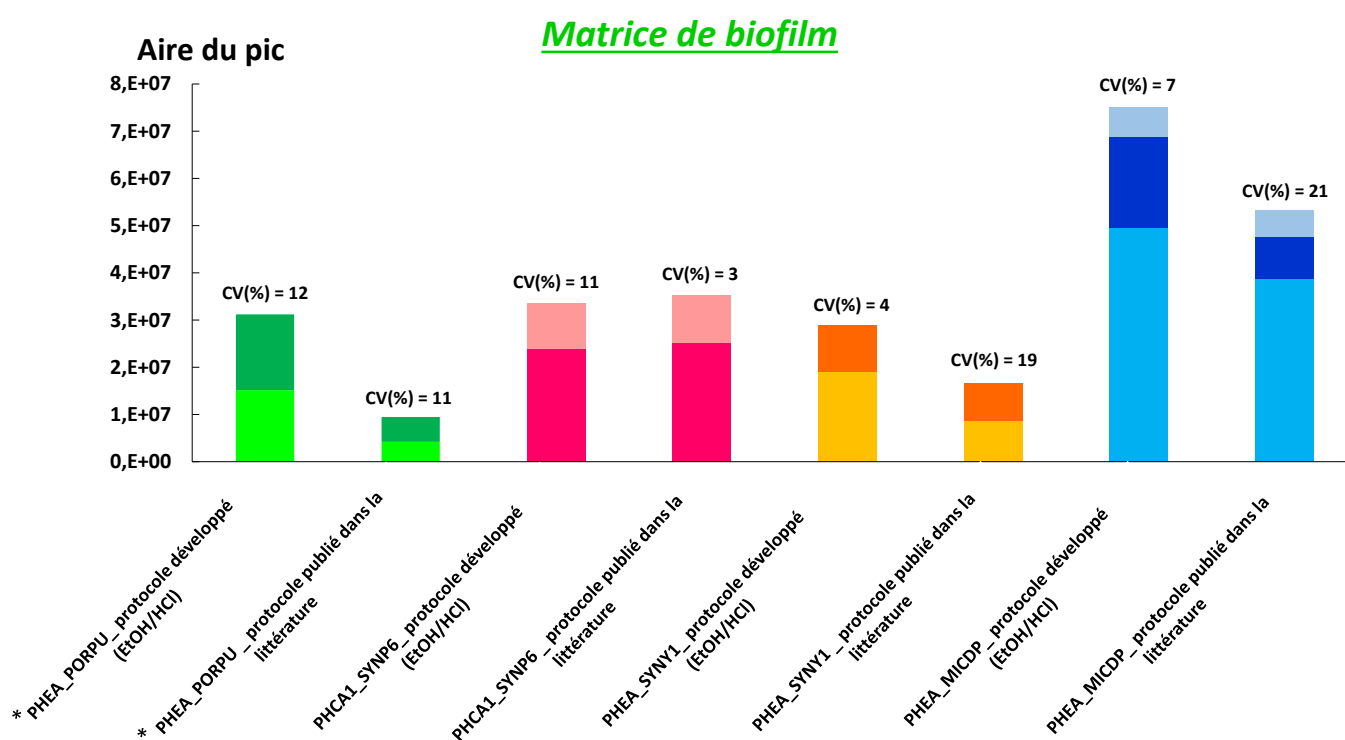
Par ailleurs, nous avons constaté que le nombre de protéines identifiées dans les matrices biofilm et foie est relativement faible au regard des études de protéomique non ciblées (Carvalhais et al., 2015; Huang et al., 2014; Shi et al., 2007; Song et al., 2016). Ceci pourrait être attribué à plusieurs facteurs associés notamment à l'acquisition MS/MS. En effet, hormis le nombre d'ions sélectionnés sur chaque spectre (TopN = 4, nombre relativement réduit), se pose le problème de la durée du gradient de séparation chromatographique qui n'est pas assez longue pour assurer un temps de balayage nécessaire pour acquérir suffisamment de mesures de chaque ion et permettre une d'identification maximale de peptides. En outre, les paramètres en mode d'acquisition données-dépendants (DDA) pour l'analyse en masse ont été choisis de manière à identifier uniquement les ions les plus abondants au détriment des ions faiblement intenses dont le signal a dû probablement être masqué, limitant par conséquent le nombre de peptides identifiés par *Mascot* (Bateman et al., 2014; Johnson et al., 2013a)

3.3. Quantification des protéines endogènes via Skyline

Une quantification relative des protéines endogènes a été réalisée, en fonction des spectres de fragmentation MS2 (ions fils). Cette analyse a été effectuée en comparant la somme des aires sous le pic de chaque ion fragment pour une masse et un temps de rétention donné.

Toutefois, le logiciel *Skyline* nécessite la construction d'une librairie de spectres établis à partir des peptides sélectionnés. De nombreuses librairies ont été créées sous le logiciel *Skyline*. Pour la matrice foie de souris par exemple, il existe des librairies disponibles en ligne et relatives à l'espèce *Mus musculus* (souris). Parmi ces librairies, on peut citer la librairie NIST (*National Institute of Standards and Technologies*) ou *Peptide Libraries*. Concernant les matrices environnementales, biofilm et sol, il a fallu créer une librairie de spectres. Pour cela, les spectres issus de l'analyse en UPLC-MS/MS ont été soumis à un traitement d'identification par l'algorithme *Mascot distiller*. Les spectres obtenus ont été regroupés pour former une librairie spectrale. Ainsi, le logiciel *Skyline* a pu extraire l'aire sous le pic de chaque peptide pour lequel un spectre est présent dans une librairie, et ce à partir des analyses UPLC-MS/MS correspondantes à chaque échantillon.

La quantification sur *Skyline* a été réalisée uniquement sur les protéines les plus identifiées dans les matrices chargées (biofilm et foie de souris), c'est-à-dire celles qui sont caractérisées par des scores *Mascot* les plus élevés (*Top scoring proteins*) (figure 99).



Matrice de foie de souris

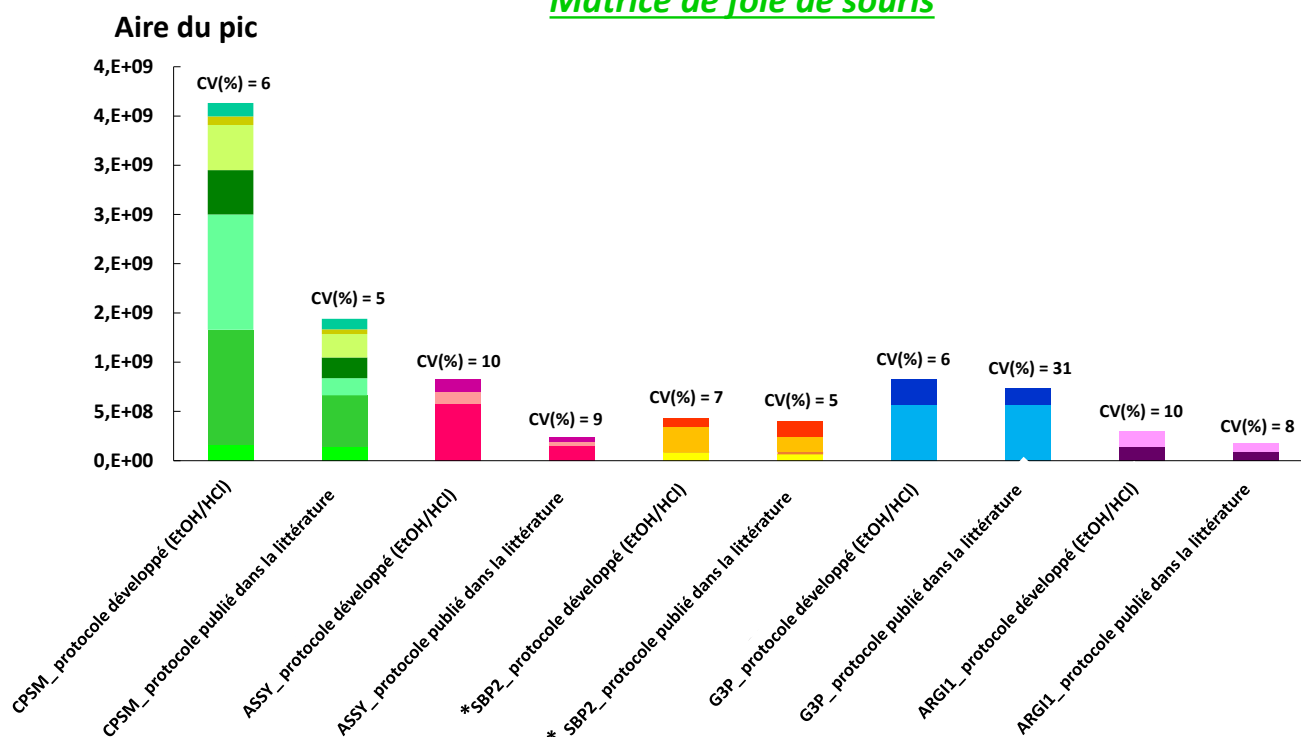


Figure 99. Quantification relative des protéines endogènes en fonction du protocole développé (EtOH/HCl) et du protocole publié dans la littérature. Les aires de pics ont été extraites des spectres MS2 correspondants aux ions fragments puis sommées pour générer les aires des ions précurseurs. La contribution de chaque ion peptide est reportée dans l'histogramme par différentes couleurs. Les protéines marquées par des astérisques (*) représentent les membranaires. La liste des transitions (couple ion précurseur/ions fragments) générée par *Skyline* est détaillée dans les tableaux 16 et 17 (cf. annexe 3), correspondants respectivement à la matrice biofilm, foie de souris

D'après la figure 99, nous pouvons constater que la quantité de protéines présentes dans l'extrait de biofilm et foie, est globalement plus importante après une purification à l'éthanol/HCl que celle obtenue à partir du protocole de la littérature. En outre, les coefficients de variations (CV %) calculés sur la base des ions fragments (Matafora et al., 2017), ne dépassent pas les 12 %, ce qui démontre la robustesse de notre approche de purification (éthanol/HCl), comparés aux CV % obtenus dans les protocoles de la littérature qui peuvent aller jusqu'à 21 et 31 %. L'approche de purification développée dans cette étude semble être intéressante pour éliminer les interférents et permettre une purification ainsi qu'une analyse directe par UPLC-HRMS/MS « hors gel » des protéines. Cette approche assez simple, a permis également d'établir une description qualitative des protéomes d'échantillons de nature et de composition extrêmement variées.

4. Conclusion

La préparation de l'échantillon, notamment le choix des solvants de précipitation, lavage et solubilisation, est une étape clé en protéomique « hors gel ». Dans cette étude le TCA, agent le plus communément employé en protéomique, a permis de déstabiliser la structure native de la protéine, *via* la rupture des interactions de faibles énergies (liaisons hydrophobes, ioniques et hydrogènes). Néanmoins, la précipitation au TCA conduit à la formation d'un état compact riche en structures secondaires, le *molten globule*. Cet intermédiaire de repliement est difficile à solubiliser même après plusieurs lavages du culot protéique au moyen de solvants organiques (acétone). Or il est primordial d'assurer une bonne re-suspension des protéines et de les maintenir en solution en amont des étapes de digestion et d'analyse par LC-MS/MS.

Nous avons démontré, grâce notamment aux techniques spectroscopiques, que l'utilisation de l'éthanol/HCl entraîne une déstabilisation de l'état MG qui se traduit par une élimination du TCA (résiduel en particulier) dans l'éthanol et une solubilisation des protéines dans le tampon d'analyse aqueux avant digestion trypsique. En outre, le lavage à l'éthanol/HCl a permis de récupérer jusqu'à 96 % des protéines cibles comparée aux solvants organiques actuels (acétone, éthanol) (jusqu'à 75 % de protéines perdues dans le surnageant). Cette approche s'est révélée être efficace quelles que soient les protéines ciblées (spikées ou endogènes), et le milieu dont elles sont issues (environnemental : sol, biofilm ou biologique : tissus).

Cette étude a fait l'objet d'une publication actuellement en ligne dans *Journal of Proteomics*.

La figure 100 illustrée ci-dessous résume le rôle de l'éthanol/HCl dans la déstabilisation de l'état *molten globule*.

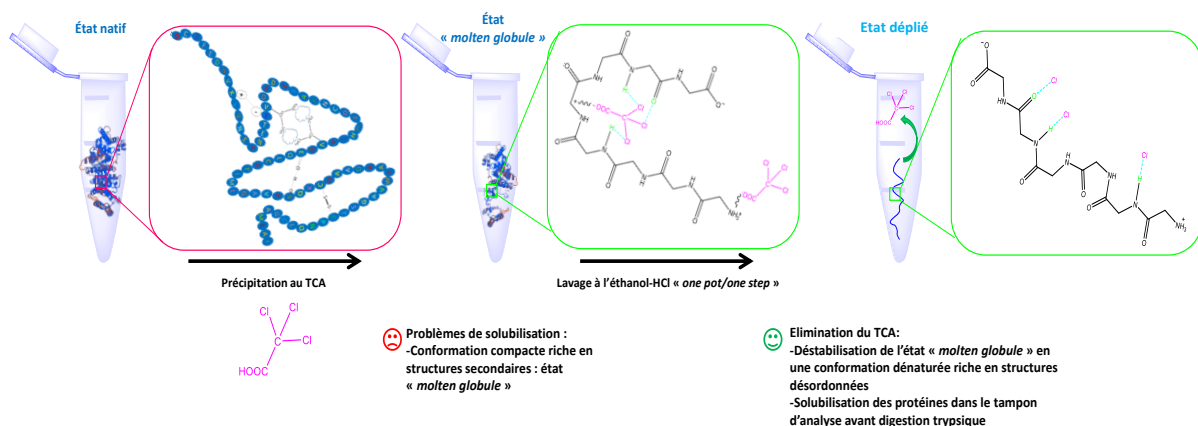


Figure 100. Rôle de l'éthanol/HCl dans la déstabilisation de l'état *molten globule*

***Chapitre III : Etude des paramètres
influençant l'extraction des protéines***

Chapitre III : Etude des paramètres influençant l'extraction des protéines

L'analyse du protéome à partir d'un milieu complexe fait face à de nombreux défis parmi lesquels, le développement de stratégies efficaces pour l'extraction des protéines peut être considéré comme un véritable talon d'Achille de toute approche protéomique. En effet, malgré le développement de techniques d'analyse de pointe, les protocoles d'extraction aujourd'hui développés aboutissent généralement à des rendements rarement satisfaisants. Ils sont par exemple de l'ordre de 7 % de protéines extraites à partir de sols (Bastida et al., 2014) ou de 1 à 15 % à partir de biofilms marins (Leary et al., 2012).

Dans ce chapitre, nous avons étudié différents facteurs pouvant influencer voire améliorer l'extraction des protéines afin de pouvoir proposer à terme un protocole à la carte compatible avec une analyse directe des protéines par LC-HRMS/MS. Nous avons donc proscrit l'utilisation de tous les interférents à l'ESI dans les solutions d'extraction (agents chaotropes/dénaturants, détergents, etc.).

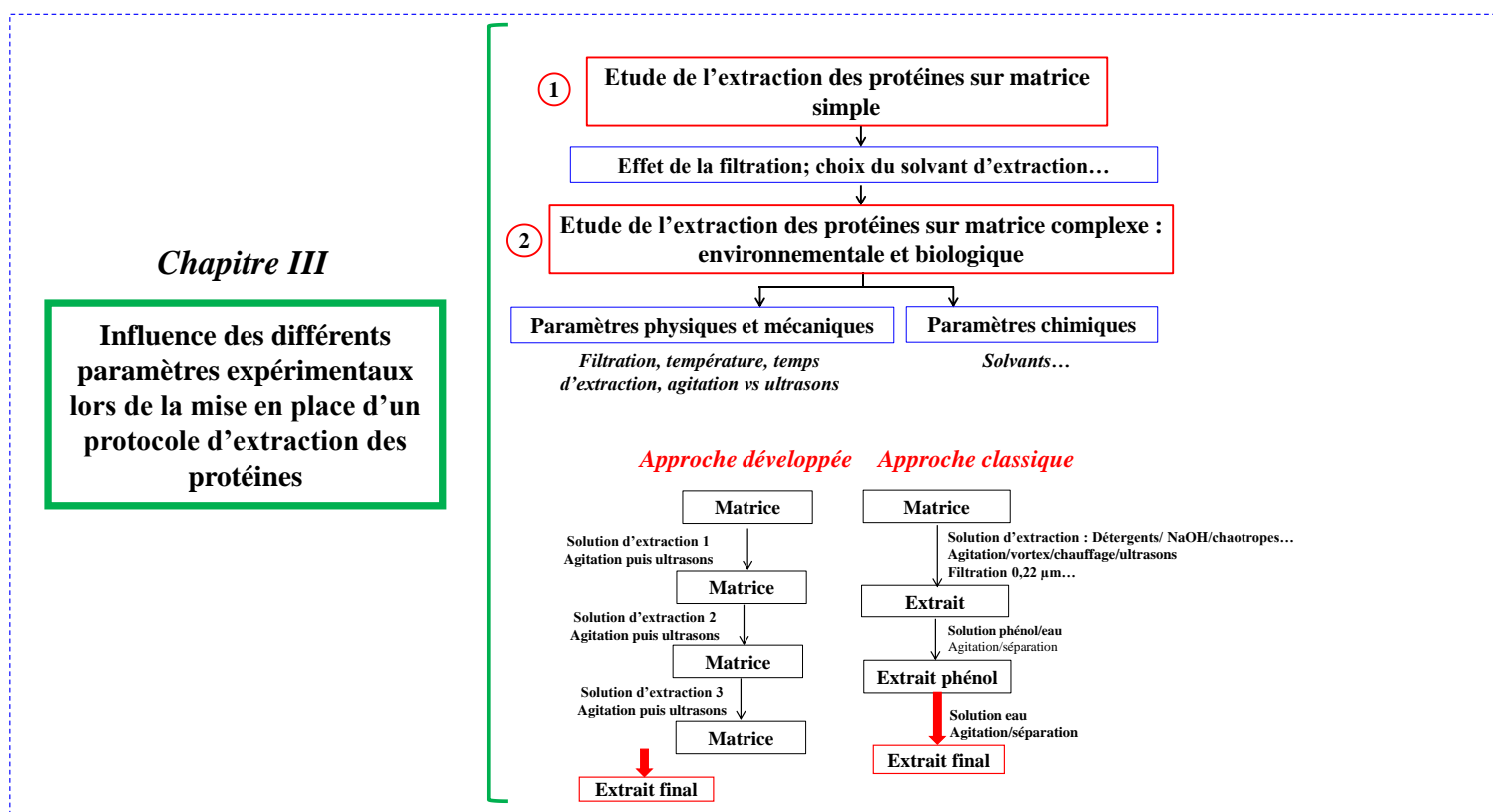


Figure 101. Stratégie expérimentale

Chapitre III : Etude des paramètres influençant l'extraction des protéines

Nous détaillerons la démarche expérimentale que nous avons suivie. Elle a tout d'abord débuté par un screening de différentes conditions d'extraction, puis une optimisation de ces conditions sur matrices complexes. Comme pour l'étape de purification, nous avons travaillé sur les 4 protéines cibles (HSA, lysozyme, Rnase et exoglucanase-1) déposées sur matrices « simples » puis complexes.

1. Essais sur matrices dopées

1.1. Développement sur matrice modèle : sable

Les résultats présentés dans cette première partie ont été obtenus à partir d'un système modèle, le sable, afin de minimiser la présence d'interférents retrouvés dans des matrices plus complexes tel que le sol (matières organiques : substances humiques, matières argileuses et organismes vivants, etc.). Toutefois, le sable n'est pas une matrice inerte vis-à-vis des protéines. En effet, Ma et al. (2007) ont démontré la présence d'interactions entre les protéines prions (présentes dans les sols et responsables des encéphalopathies spongiformes transmissibles) et une matrice modèle de sable (quartz) (Ma et al., 2007). Ceci peut donc expliquer le fait que certains rendements d'extraction ne soient pas toujours égaux à 100 %.

1.1.1. Optimisation de l'étape de filtration post-extraction

Avant de développer le protocole d'extraction proprement dit, nous nous sommes tout d'abord intéressées à l'étape de filtration qui suit généralement cette phase d'extraction. Bien qu'elle permette la séparation entre la phase solide (matrice) et la phase liquide (solvant d'extraction), cette étape, laborieuse et lente, conduit généralement à un colmatage des filtres (Jiang et al., 2004). Ceci peut entraîner une diminution significative de la quantité de protéines d'intérêts dans le filtrat (Hagen, 2011; Owusu-Apenten, 2004).

Dans un premier temps nous avons utilisé une solution de protéines standards et constaté une perte d'environ 82 % de la HSA après filtration sur 0,45 μm . Ces pertes ne peuvent pas être attribuées à la faible taille des pores, car, à titre d'exemple, les tailles de la HSA (60 Å) et du lysozyme (34 Å) sont largement inférieures à celle des pores. Des résultats similaires ont été observés dans des solutions de fluides biologiques. Ruhl et al. (Ruhl et al., 2011) ont mis en évidence que 50 % des protéines présentes dans la salive étaient retenus sur des filtres à seringues dont le diamètre des pores était de 0,45 μm

Pour être plus représentatifs des échantillons que nous souhaitons extraire, nous avons également testé l'effet d'une filtration sur des filtres de 0,45 μm de diamètre sur la récupération

Chapitre III : Etude des paramètres influençant l'extraction des protéines

des 4 protéines standards mélangées à 1 g de sable. L'extraction a été réalisée avec 10 mL d'un mélange eau-méthanol (50/50) à température ambiante (Bertrand, 2009). Ce solvant a été auparavant employé pour l'extraction des petites molécules, acides aminés, dipeptide ou encore des polyphénols à partir des plantes (Bertrand, 2009; Bertrand et al., 2012; Liu et al., 2013). Nous avons réalisé la même expérience sans étape de filtration (figure 102).

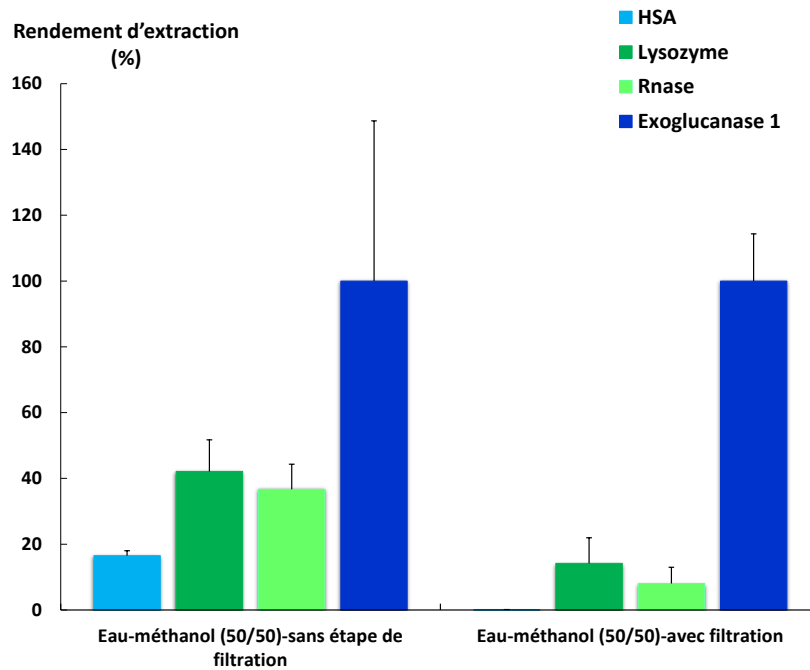


Figure 102. Effet de la filtration sur la récupération des protéines cibles

Les rendements d'extraction sont relativement faibles (sauf pour l'exoglucanase-1), avec le mélange eau/méthanol. En outre l'extraction des protéines à l'aide de solvants organiques ou de mélanges hydro-organiques est susceptible d'induire leur précipitation. Pour cette raison, ce solvant n'a pas été retenu pour la suite de l'étude.

Le rendement de récupération de 3 protéines sur 4 (HSA, lysozyme et Rnase) diminue après filtration. Cette chute est de 30 % pour le lysozyme et la Rnase, et avoisine les 100 % pour la HSA. En outre, nous avons également remarqué au cours de l'expérience, un colmatage rapide du filtre causé par le dépôt de la matrice sur la membrane, ceci est susceptible d'accentuer le phénomène d'adsorption des protéines sur les filtres (Kelly et al., 1993). D'autres types de filtres auraient pu être envisagés tels que les filtres à membrane acétate présentant un faible taux d'adsorption des protéines (Schulze et al., 2005). Néanmoins, en raison de la perte significative des protéines au cours de la filtration (de solutions standards et d'extraits protéiques), nous avons décidé de supprimer cette étape et d'inclure uniquement une séparation solide-liquide par centrifugation.

Chapitre III : Etude des paramètres influençant l'extraction des protéines

1.1.2. Optimisation du solvant d'extraction

L'extraction des protéines par solvant utilise généralement des solutions tampons additionnées de détergents ou chaotropes. L'emploi de ces solutions ne permet néanmoins pas d'obtenir des rendements satisfaisants (de l'ordre de 7 % de protéines extraites à partir de sols (Bastida et al., 2014)) et peut altérer la digestion trypsique ainsi que l'analyse par LC-MS/MS. Nous avons étudié la possibilité d'utiliser le mélange eau/isopropanol ou encore un deuxième solvant organique très polaire, le diméthylsulfoxyde (DMSO), jusqu'alors non utilisé en biochimie, mais pouvant s'avérer efficace pour l'extraction des protéines. Ces solvants ont été associés à un tampon ionique régulièrement employé en protéomique, le Tris-HCl. Aucun des deux agents n'a d'effet sur la digestion ou sur l'analyse en masse. Le DMSO peut même à contrario augmenter le signal en masse des peptides (Hahne et al., 2013).

Le choix du DMSO s'est fait de par sa capacité à solubiliser un grand nombre de molécules organiques, y compris les protéines (lysozyme, ovalbumine, trypsine, etc.), les acides nucléiques (Chang et al., 1991). Cette propriété est due à la liaison S-O qui permet au DMSO de former des liaisons hydrogènes responsables de sa solubilité dans l'eau (Georgia and Xavier, 2014). Il peut également avoir un effet dénaturant sur les protéines lorsqu'il est présent en fortes concentrations (10-70 % v/v dans l'eau) (Tjernberg et al., 2006).

Le Tris-HCl ($C_4H_{11}NO_3 \cdot HCl$) est un tampon ionique couramment utilisé en biochimie pour solubiliser les protéines et les acides nucléiques. Il forme des liaisons hydrogènes entre ses groupements hydroxyles (-OH) et amine (-NH₂) et les acides aminés protéiques (Quan et al., 2008; Taha and Lee, 2010). À la différence des agents dénaturants (urée, par exemple), le Tris aboutit à une solubilisation quasi-totale des tri et tétra-peptides (Taha and Lee, 2010). Dans notre étude, une concentration de 25 mM (pH 7,5) a été choisie. Cette concentration a auparavant été utilisée pour extraire les protéines d'échantillons très divers tels que des phytoplanctons, cellules végétales, ou tissus biologiques. Le Tris est également utilisé pour extraire les protéines membranaires, et ce en augmentant la perméabilité des membranes en interagissant avec les lipopolysaccharides qui la constituent (García-Otero et al., 2013; Rao and Hsieh, 2014; Witzel et al., 2011). En effet, les protéines membranaires présentent généralement une très faible solubilité dans l'eau puisqu'elles sont fortement ancrées dans la bicouche lipidique (Blonder et al., 2007). Des solutions de détergents sont souvent employées pour déstabiliser les membranes biologiques (Seddon et al., 2004).

L'isopropanol a également été testé pour extraire les protéines. De façon générale, les mélanges hydro-alcooliques favorisent la dénaturation des protéines (Chong et al., 2015). Avec

Chapitre III : Etude des paramètres influençant l'extraction des protéines

leur chaîne carbonée et leur groupement hydroxyle (caractère amphiphile), les alcools agissent comme des détergents (Singh et al., 2012), *via* des interactions hydrophobes avec les résidus hydrophobes des protéines et des liaisons hydrogènes avec le groupement carbonyle et amine du squelette peptidique (C=O ; NH₂) (Dwyer and Bradley, 2000). Singh et al. (2012) ont démontré par exemple, que le n-propanol, à une concentration d'environ 45 %, favorise la solubilisation des corps d'inclusion (protéine biologiquement inactive présente sous forme agrégée), comme le feraient des dénaturants forts (urée à 2 M).

Une extraction séquentielle a alors été réalisée, afin d'évaluer l'efficacité de chaque solvant (DMSO, eau/isopropanol ou Tris-HCl) et de démontrer l'intérêt de combiner les propriétés de ces différents solvants pour extraire une quantité maximale des protéines cibles.

Nous représentons sur la figure 103 les résultats les plus probants, où chaque solvant a été utilisé une fois en première extraction (effet de chaque solvant) et une fois en deuxième extraction :

- Extraction au DMSO (99 %), puis extraction avec un mélange d'eau ultrapure/isopropanol (50/50 : v/v).
- Extraction avec un mélange d'eau ultrapure/isopropanol (50/50: v/v), puis extraction avec du Tris-HCl 25mM (pH 7,5).
- Extraction au Tris-HCl 25mM (pH 7,5), puis extraction au DMSO (99 %).

Chapitre III : Etude des paramètres influençant l'extraction des protéines

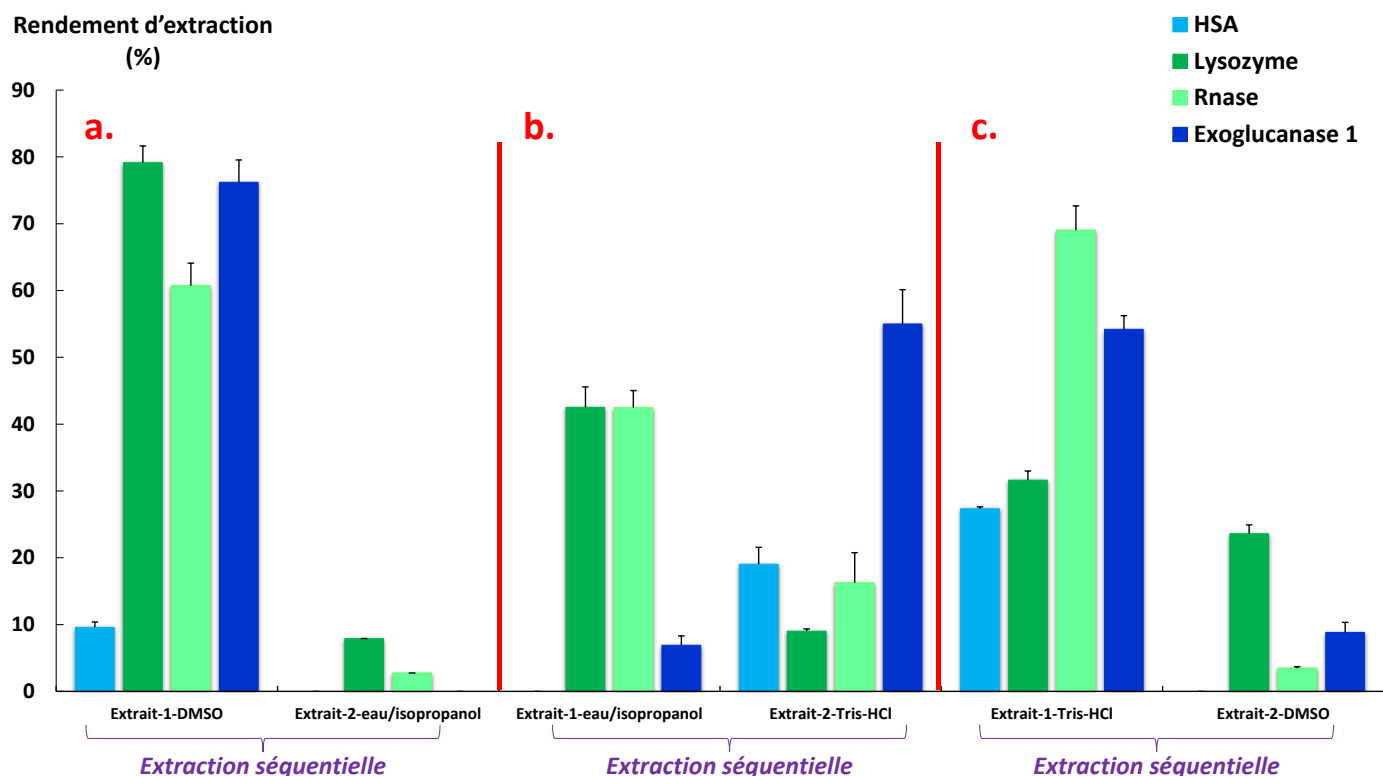


Figure 103. Effet du solvant sur la récupération des protéines pour une extraction séquentielle

Comme présentés sur la figure 103, les rendements obtenus sur la matrice sable sont fonction des caractéristiques physicochimiques des protéines et du solvant utilisé. Trois protéines sur 4 ont été récupérées dès la première extraction avec le DMSO (figure 103.a). Dans ce cas, une deuxième extraction sur la matrice avec un mélange eau ultrapure/isopropanol n'a permis de récupérer qu'une très faible quantité de protéines (7 et 3 % de lysozyme et Rnase, respectivement). Utilisé en tant que premier agent d'extraction, ce même mélange (eau ultrapure/isopropanol) conduit à des rendements ne dépassant pas les 40 % (Figure 103.b). Le reste étant extrait avec le Tris-HCl. Comme le DMSO, ce dernier a permis d'obtenir des rendements de 60 à 80 % pour Rnase et l'exoglucanase-1 (Figure 103.c). Il est par contre moins efficace pour extraire le lysozyme et la HSA. Cette étude montre que l'extraction de protéines introduites sur un échantillon d'espèce minéralogique connue et unique est déjà source de discrimination (les rendements variant d'une protéine à l'autre).

À la vue de ces résultats, nous avons retenu le DMSO et le Tris-HCl pour les expériences d'extraction des protéines cibles à partir d'un sol, matrice environnementale plus complexe de par sa minéralogie et sa matière organique propre. Dans ce cas, un mélange de DMSO/Tris-HCl a été choisi (les rendements d'extraction sur sable obtenus avec ce mélange sont supérieurs

Chapitre III : Etude des paramètres influençant l'extraction des protéines

à 45 % pour 3 protéines sur 4 ; cf. annexe 4) afin de réduire le temps nécessaire à d'extraction (2h00 au lieu de 4h00).

1.2. Extraction à partir d'une matrice environnementale : d'un échantillon de sol dopé (site hydrogéologique du Deffend)

Le sol est défini comme la couche externe de la croûte terrestre. Il est le siège d'un échange intense de matière et d'énergie entre l'air, l'eau et les roches et occupe une position clé dans les cycles globaux (cycle du carbone, azote, phosphore, etc.). Il est constitué de trois phases : solide, liquide et gazeuse. La phase solide comprend la matière organique et la fraction minérale; la fraction liquide issue de l'altération de la roche mère sert de vecteur de matière; et l'interface gazeuse apporte l'oxygène nécessaire au développement des microorganismes (Gobat et al., 2010). La formation du sol résulte de l'altération de la roche mère, ainsi que de la biodégradation de la matière organique fraîche d'origine biologique (matière végétale et animale). Ceci entraîne la formation de résidus dont la transformation, *via* des processus d'humification, aboutit aux substances « humiques ». Les protéines se trouvent généralement encapsulées/adsorbées aux substances humiques et/ou particules argileuses du sol, ce qui rend leur extraction extrêmement difficile (Kiss et al., 1975; Schmidt and Martínez, 2016; Tomaszewski et al., 2011; Yu et al., 2013a).

Les expériences d'extraction des protéines ont été menées sur un sol limono-argileux échantillonné sur le site du Deffend (cf. matériels et méthodes). Les données globales relatives à la caractérisation de la matière organique de ce sol sont regroupées dans le tableau 15 (Célérier, 2008).

Tableau 15. Caractérisation du sol du Deffend (Célérier, 2008)

	Horizon A₁ (0-5 cm)	Horizon A₂ (5-20 cm)
C organique (g.Kg⁻¹)	13,70	9,51
N total (g.Kg⁻¹)	1,18	0,92
MO (%)	2,36	1,65
Cendres (%)	96,1	96,9
C/N	11,60	10,30
pH	7,7	6,4

Chapitre III : Etude des paramètres influençant l'extraction des protéines

Les teneurs en carbone, azote (dont la matière azotée de nature protéique) et matière organique sont faibles et diminuent avec la profondeur. La valeur du pH est relativement proche de la neutralité.

Toutes les extractions ont été réalisées sur 50 mg de sol, préalablement lyophilisé et tamisé à 500 µm puis dopé avec les 4 protéines cibles (0,1 % = 0,05 mg/50 mg_{matrice}). Les conditions du milieu (température, humidité, activité biologique et temps d'adsorption) influent sur l'abondance et la disponibilité des protéines dans les sols. Kanerva et al. (2013) stipulent que la récupération des protéines dopées sur un sol dépend fortement du temps laissé pour favoriser leur adsorption. Dans notre étude, nous avons choisi d'incuber les protéines au contact du sol à température ambiante afin de maximiser leurs chances d'adsorption sur la matrice. L'extraction a été réalisée en présence de 500 µL du mélange DMSO/Tris-HCl (50/50 ; v/v).

1.2.1. Études des paramètres influençant l'extraction des protéines

1.2.1.1. Paramètres physiques et mécaniques

La mise en œuvre de techniques d'extraction nécessite une optimisation des conditions opératoires (température, durée d'extraction, choix de la méthode, etc.) qui agissent en l'occurrence sur la diffusion de la solution d'extraction dans la matrice solide, mais également sur la solubilité des solutés (ici protéines) dans la phase liquide.

L'objectif de ces expériences était de faire un screening rapide (une seule répétition) de l'influence de plusieurs paramètres physiques et mécaniques et d'en déduire leur apport en termes d'efficacité d'extraction. L'idée était d'avoir une tendance sur la pertinence de leur utilisation, afin de limiter le nombre d'étapes et d'obtenir un protocole simplifié. Ces derniers ont été combinés à une extraction solide/liquide avec le mélange DMSO/Tris-HCl. Dans le cadre de ce travail, la température, la durée de l'extraction ainsi que la méthode utilisée (extraction par agitation ou extraction assistée par ultrasons), ont été particulièrement étudiées.

1.2.1.1.1. Effet de la température d'extraction

Quatre températures d'extraction ont été testées (25, 35, 45 et 55 °C). Les résultats obtenus ont permis de montrer une diminution des rendements protéiques avec l'augmentation de la température d'extraction. En effet, bien qu'une température modérée favorise l'extraction en rompant les liaisons de faibles énergies (Cui et al., 2017; Varzakas and Tzia, 2014), elle peut également induire une dénaturation et/ou une agrégation des protéines directement sur la

Chapitre III : Etude des paramètres influençant l'extraction des protéines

matrice (Zink et al., 2016) lorsqu'elle est trop élevée. Une température d'extraction de 25 °C a été donc choisie. (cf. annexe 5)

1.2.1.1.2. Effet du temps d'extraction

Quatre durées d'extraction ont été évaluées (2h00, 3h00, 4h00 et 5h00, à 25 °C). Les résultats ont montré que l'extraction des protéines cibles augmente légèrement au cours de temps pour 3 protéines sur 4. Dans ce cas, la récupération de la HSA, du lysozyme et de la Rnase augmente de 1 à 3 % après 5h00 d'extraction, alors que le rendement de l'exoglucanase-1 tend à diminuer au cours du temps (cf. annexe 6). Un temps d'extraction de 2h00 a été sélectionné pour les expériences suivantes.

1.2.1.1.3. Effet de la méthode d'extraction

Outre la température et le temps d'extraction, nous avons également étudié une extraction agitation mécanique ou assistée par ultrasons (sonication). L'extraction assistée par ultrasons s'est notamment révélée efficace pour améliorer l'extraction des protéines à partir de sols ou de plantes (Keiblinger et al., 2012; Starke et al., 2016). De plus, elle facilite la solubilisation des protéines dans la phase extractante et réduit ainsi l'utilisation de détergents et/ou chaotropes (Wang et al., 2011).

Le mécanisme d'action des ultrasons repose sur la génération d'ondes sonores qui engendrent des vibrations mécaniques au sein même d'un solide, un liquide ou un gaz. La propagation de ces ondes (fréquences entre 20 KHz et 100 KHz) entraîne des séries de compression (pression positive) et de raréfaction (pression négative) au sein du milieu. Cette action permet de générer des bulles de cavitation qui vont se développer puis s'effondrer (imploser), générant ainsi des conditions de pression (50-1000 atm) et de températures locales extrêmes (> 5000 °C) (Chemat et al., 2017). À proximité d'une surface solide, les bulles de cavitation formées deviennent asymétriques et l'implosion de ces bulles produit un jet de liquide à grande vitesse dirigé vers la surface solide ce qui permet de la perforer (figure 104).

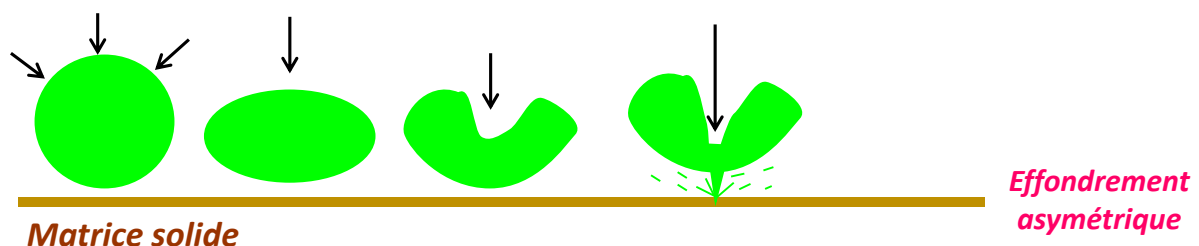


Figure 104. Développement d'une bulle de cavitation à la surface d'un solide

Chapitre III : Etude des paramètres influençant l'extraction des protéines

Chemat et al. (2017) ont décrit les effets sonochimiques sur l'extraction de composés naturels de plantes par plusieurs mécanismes. Seuls les mécanismes pouvant être extrapolés à l'amélioration de l'extraction des protéines seront décrits ci-après :

a) La fragmentation de la matrice qui est engendrée par des phénomènes de collisions inter-particules et des ondes de choc violentes créées à partir de jets de liquide (Kadam et al., 2015). Une des principales conséquences de ce mécanisme est la réduction des tailles des particules solides de la matrice et une augmentation du nombre des particules individuelles, ce qui conduit à une augmentation de la surface spécifique du solide et ainsi à la diffusion du solvant d'extraction dans la matrice (transfert de matière plus important) (Wang et al., 2011).

b) La sonoporation qui engendre des dommages aux cellules vivantes ou à la matière solide agrégée, facilitant ainsi la libération des protéines intracellulaires (Saleh et al., 2016).

c) Un cisaillement est également généré à proximité de la matière solide créant ainsi une meilleure homogénéisation du milieu réactionnel. Ces forces de cisaillements sont suffisamment fortes pour rompre les liaisons chimiques et solubiliser les protéines dans la phase liquide.

d) La production d'espèces extrêmement réactives radicalaires $H\bullet$ et $OH\bullet$, issues de la sonolyse de l'eau et associée aux forces de cisaillement sont responsables de la dénaturation des protéines (Stathopoulos et al., 2004; Zink et al., 2016). En effet, les radicaux libres peuvent s'associer aux résidus d'acides aminés des protéines et ainsi augmenter leur solubilité. À titre d'exemple, le traitement par ultrasons des protéines du lactosérum améliore leur solubilité dans les solutions aqueuses (Kentish and Ashokkumar, 2011).

Dans le cadre de cette étude, la sonication a été comparée à l'agitation classique pour l'extraction des 4 cibles protéiques. Les résultats (cf. annexe 7) ont montré que l'extraction assistée par ultrasons, avec un solvant organo-alcalin, aboutit à une amélioration très faible du rendement d'extraction des 4 protéines (augmentation de 2 % de leurs rendements) comparé à l'agitation classique (2h00 à 25 °C). Ceci peut s'expliquer par la richesse organique du milieu qui peut conduire à une atténuation des ondes ultrasonores dans la proximité de l'émetteur ultrasonique. Cet effet peut alors entraîner une diminution de la récupération des composés d'intérêt (Wang and Weller, 2006).

À la vue de ces résultats, nous avons décidé de combiner l'extraction aux ultrasons à l'agitation classique dans les expériences qui seront présentées dans le paragraphe 1.2.1.2.2. En

Chapitre III : Etude des paramètres influençant l'extraction des protéines

effet, malgré la légère augmentation du rendement d'extraction après sonication et les difficultés liées à extraire les protéines du sol (rendement protéique d'environ 7 %, (Bastida et al., 2014)), les 2 % permettront peut-être de détecter des protéines jusqu'alors jamais identifiées dans ces matrices.

1.2.1.2. Paramètres chimiques

Dans cette partie, nous nous sommes intéressées aux rôles de plusieurs paramètres chimiques sur l'efficacité de l'extraction des protéines. L'effet d'une extraction à la soude, d'un traitement au tetraméthylammonium hydroxyde (TMAH) et aux acides aminés ont été évalués.

1.2.1.2.1. Effet de la soude

De nombreuses études de méta-protéomique (cf. Partie bibliographique.(3.2.2)) utilisent la soude ou des tampons alcalins (pyrophosphate de sodium) pour faciliter l'extraction des protéines des sols (Masciandaro et al., 2008; Benndorf et al., 2007). Nous avons étudié, l'effet de la soude sur le rendement extraction des protéines cibles (figure 105).

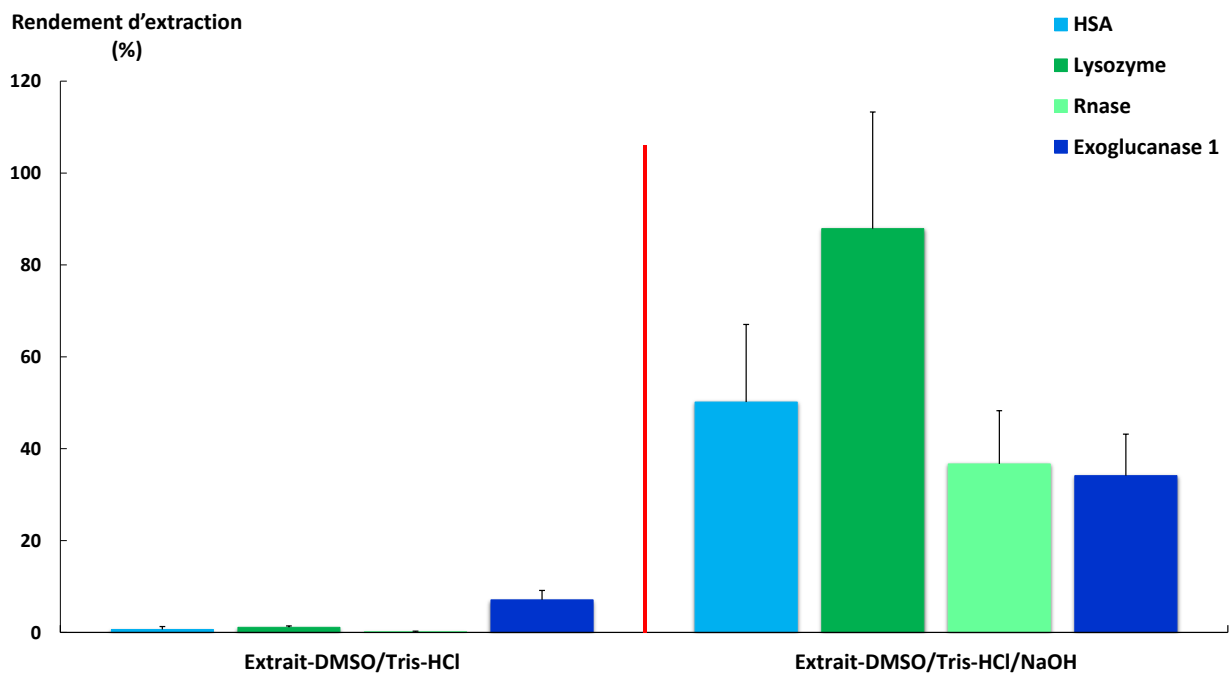


Figure 105. Effet de la soude sur le rendement de récupération des protéines cibles (agitation 2h00, à 25 °C)

L'ajout de soude dans le solvant d'extraction (DMSO/Tris-HCl) a permis en effet d'améliorer de manière considérable la récupération des protéines cibles (jusqu'à 88 % de récupération du lysozyme). Néanmoins, l'un des principaux inconvénients connus de la soude

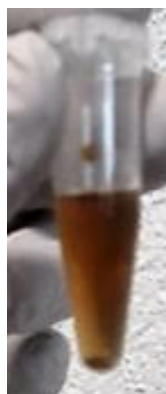
Chapitre III : Etude des paramètres influençant l'extraction des protéines

est qu'elle entraîne une co-extraction de nombreux interférents et notamment, des substances humiques (SH) auxquelles les protéines sont très fortement liées (Benndorf et al., 2009; Tomaszewski et al., 2011; Zampieri et al., 2016; Benndorf et al., 2007). Les interactions protéines-SH sont difficiles à rompre ce qui implique plusieurs lavages de l'extrait au phénol afin d'isoler spécifiquement la fraction protéique (Keiblinger et al., 2016). De ce fait, la combinaison de la soude et du phénol se traduit généralement par des rendements d'extraction très faibles, qui peuvent être attribués à une perte des protéines au cours des lavages successifs au phénol (environ 1 à 7 %) (Bastida et al., 2014; Benndorf et al., 2007). Ainsi dans leur ensemble, les extractions en milieu alcalin se heurtent à la présence de nombreux interférents, réduisant de façon considérable la pureté de l'extrait protéique ce qui a pour conséquence de fausser l'analyse qualitative et quantitative des protéines. Ceci est particulièrement vrai lorsque des méthodes d'analyse spectroscopie UV-visible sont employées (dosage à 280 nm ou dosage colorimétrique (Bradford)) (Bastida et al., 2014; Roberts and Jones, 2008; Benndorf et al., 2007). En effet, les noyaux aromatiques des résidus d'acides aminés (Tyrosine, tryptophane et phénylalanine) absorbent à des longueurs d'ondes situées aux alentours de 280 nm, comme les noyaux aromatiques des substances humiques (Roberts and Jones, 2008). De même pour l'analyse directe en LC-MS/MS, les extraits protéiques obtenus avec la soude conduisent le plus souvent à un colmatage des préfiltres des colonnes analytiques, et à un bruit de fond conséquent en spectrométrie de masse (élution de composés non protéiques).

La figure 106 représentée ci-dessous témoigne de la quantité de matière organique extraite en présence de soude. La teinte brun-rougeâtre confère au sol sa couleur foncée.



Extrait DMSO/Tris-HCl



Extrait NaOH/DMSO/Tris-HCl



Extrait NaOH

Figure 106. Photos de trois tubes eppendorfs après extraction des protéines cibles à partir d'un

Chapitre III : Etude des paramètres influençant l'extraction des protéines

Pour conclure, la méthode d'extraction combinée à la soude a permis certes d'améliorer la récupération des protéines, mais elle reste tout de même limitée par des problèmes de co-extraction des substances humiques. En outre, cette méthode est systématiquement suivie par une extraction au phénol. Or actuellement l'utilisation du phénol est controversée par de nombreux auteurs en raison de sa forte toxicité (Keiblinger et al., 2016).

1.2.1.2.2. Effet de nouveaux agents chimiques ou biochimiques sur l'extraction des protéines

Afin de s'affranchir de l'utilisation de cette base forte, nous avons étudié l'effet de nouveaux agents chimiques pour désorber les protéines :

- Un agent basique de solubilisation : le tetraméthylammonium hydroxyde (TMAH) couplé à une extraction au DMSO.

- Les acides aminés basiques : d'après Nicora et al. (2013) la présence d'acides aminés augmente l'efficacité d'extraction des protéines. L'idée est ici de transférer les interactions initialement établies entre les acides aminés protéiques et la matrice vers des acides aminés libres. Dans notre étude, l'échantillon de sol a été mis en contact avec une solution d'acides aminés basiques à une concentration de 1 % (w/v) à pH 5 pendant toute une nuit à 37 °C sous agitation douce. À ce pH, les aminoacides libres possèdent une charge nette positive et vont pouvoir entrer en compétition avec les acides aminés protéiques, initialement adsorbés sur la matière minérale/organique du sol, qui elle est chargée négativement pour un pH entre 3 et 7 (Bowles, 1984). Les protéines seront alors libérées dans la phase aqueuse. Alors que dans l'étude menée par Nicora et al. (2013), un mélange de sels et de détergents (SDS) a été utilisé pour extraire les protéines après traitement aux acides aminés (Nicora et al., 2013), nous avons couplé ce traitement aux acides aminés à une extraction au DMSO.

- Le chlorure de magnésium ($MgCl_2$) (Rao and Hsieh, 2014).

Les échantillons de sol ont été laissés au contact de ces différents agents chimiques et biochimiques pendant toute une nuit à 37 °C, puis les 4 protéines ont été extraites avec le DMSO, dans le cas d'un traitement au TMAH et acides aminés basiques. Au cours de cette expérience, l'extraction a été réalisée sous une agitation douce afin de favoriser les phénomènes de transfert, puis comparée à un échantillon témoin (Extrait-eau-DMSO). Les résultats sont présentés sur la figure 107.

Chapitre III : Etude des paramètres influençant l'extraction des protéines

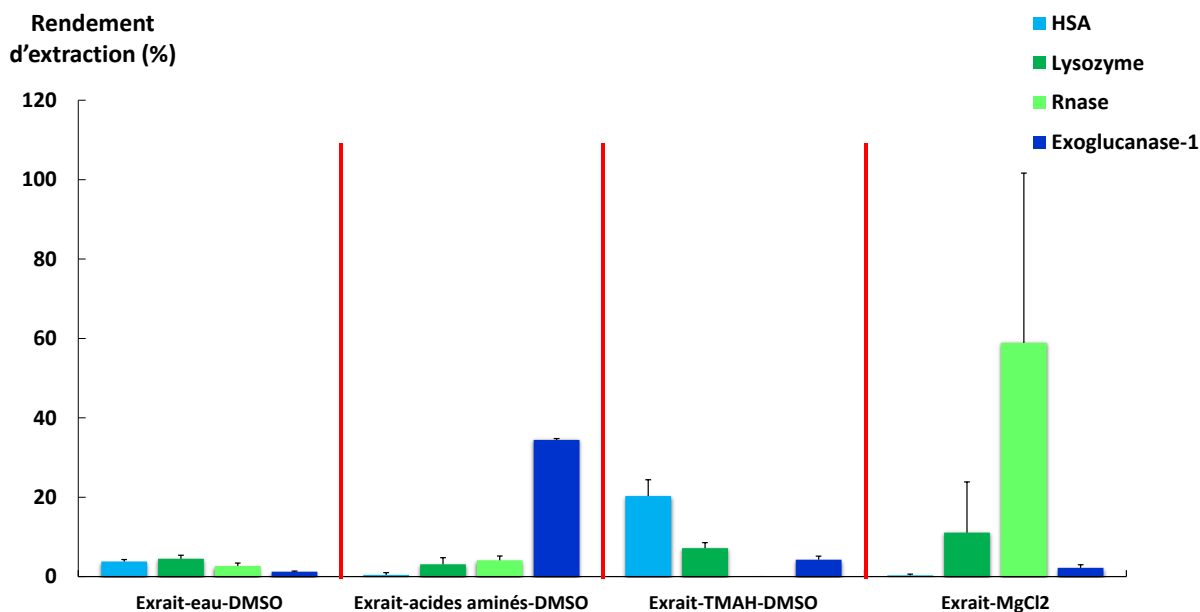


Figure 107. Effet des différents agents chimiques (TMAH, $MgCl_2$) et biochimiques (acides aminés basiques) sur le rendement d'extraction des protéines

Alors qu'un traitement aux acides aminés suivi d'une extraction au DMSO entraîne une augmentation considérable de la quantité de l'exoglucanase-1 avec un rendement final de 35 %. Un traitement au TMAH semble être intéressant pour la récupération des protéines dont le pH_i est acide telle que la HSA (pH_i 4,8). Néanmoins comme la soude, un des inconvénients majeurs du TMAH est son fort pouvoir extractant des matières organiques des sols. L'extrait TMAH était en effet caractérisé par une couleur marron très foncé. Ce traitement peut donc être effectué en fin de séquence d'extraction. Une extraction au $MgCl_2$ semble être favorable pour extraire les petites protéines (en particulier la Rnase).

À l'issue de ces résultats, nous avons réalisé une extraction séquentielle en commençant par une étape de traitement au TMAH. Les résultats présentés sur la figure 108, indiquent qu'une extraction incluant une étape de traitement au TMAH en début du processus est défavorable pour extraire l'ensemble des cibles protéiques. Dans le cas des protéines, dont le pH_i est basique (lysozyme et de la Rnase), ce traitement agressif semble induire leur précipitation au sein de la matrice.

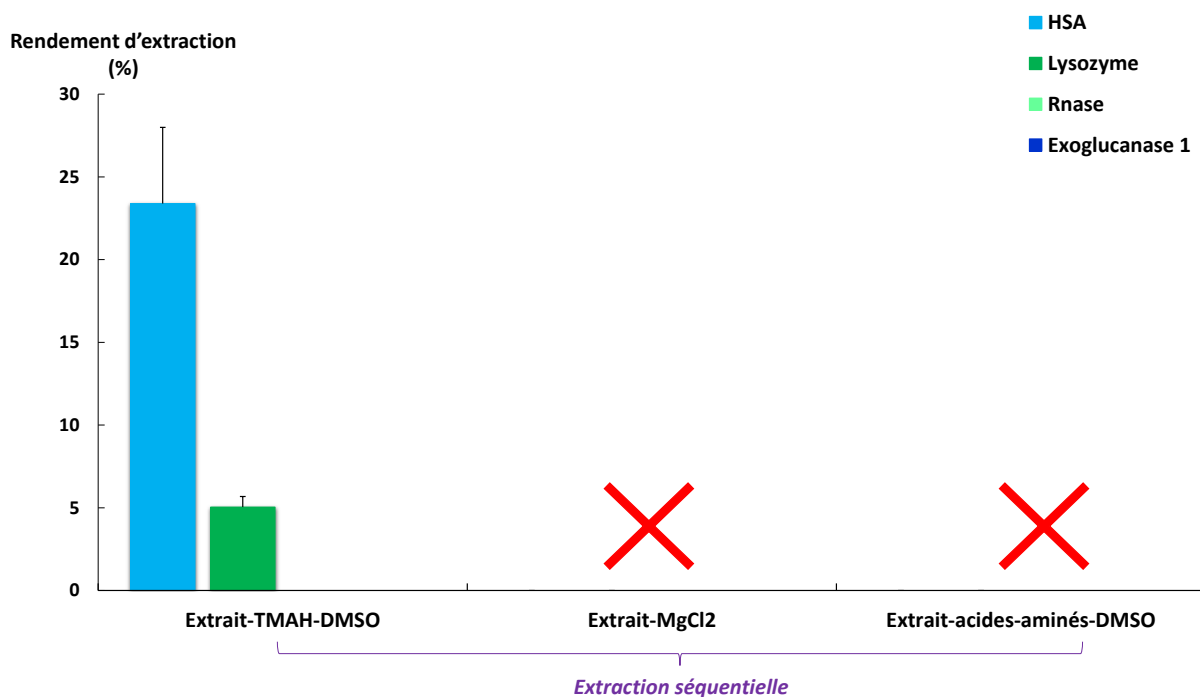


Figure 108.Extraction séquentielle : effet du TMAH

Ainsi, nous avons envisagé de réaliser une approche séquentielle en incluant d'abord une extraction en présence de $MgCl_2$, suivie par un traitement aux acides aminés/DMSO, et puis par un traitement alcalin TMAH/DMSO (le plus agressif). Ce dernier a pour objectif de désorber les protéines qui seraient en forte interaction avec la matière organique de la matrice.

1.2.1.3. Extraction séquentielle : Effets combinés

Les protéines ont été extraites au moyen de divers réactifs (1. $MgCl_2$; 2. Acides aminés/DMSO et 3. TMAH/DMSO), afin de solubiliser de manière séquentielle les différentes fractions protéiques et de limiter les interférents de l'extrait dus à la présence d'autres constituants de la matière organique du sol (comme c'est le cas d'une extraction simple au DMSO/Tris-HCl/NaOH). Au cours de cette expérience, une extraction par agitation classique a été réalisée suivie d'un traitement aux ultrasons (5 min, 80 %), afin d'optimiser la solubilisation des cibles protéiques. Les résultats de l'extraction séquentielle sont présentés dans la figure 109.

Chapitre III : Etude des paramètres influençant l'extraction des protéines

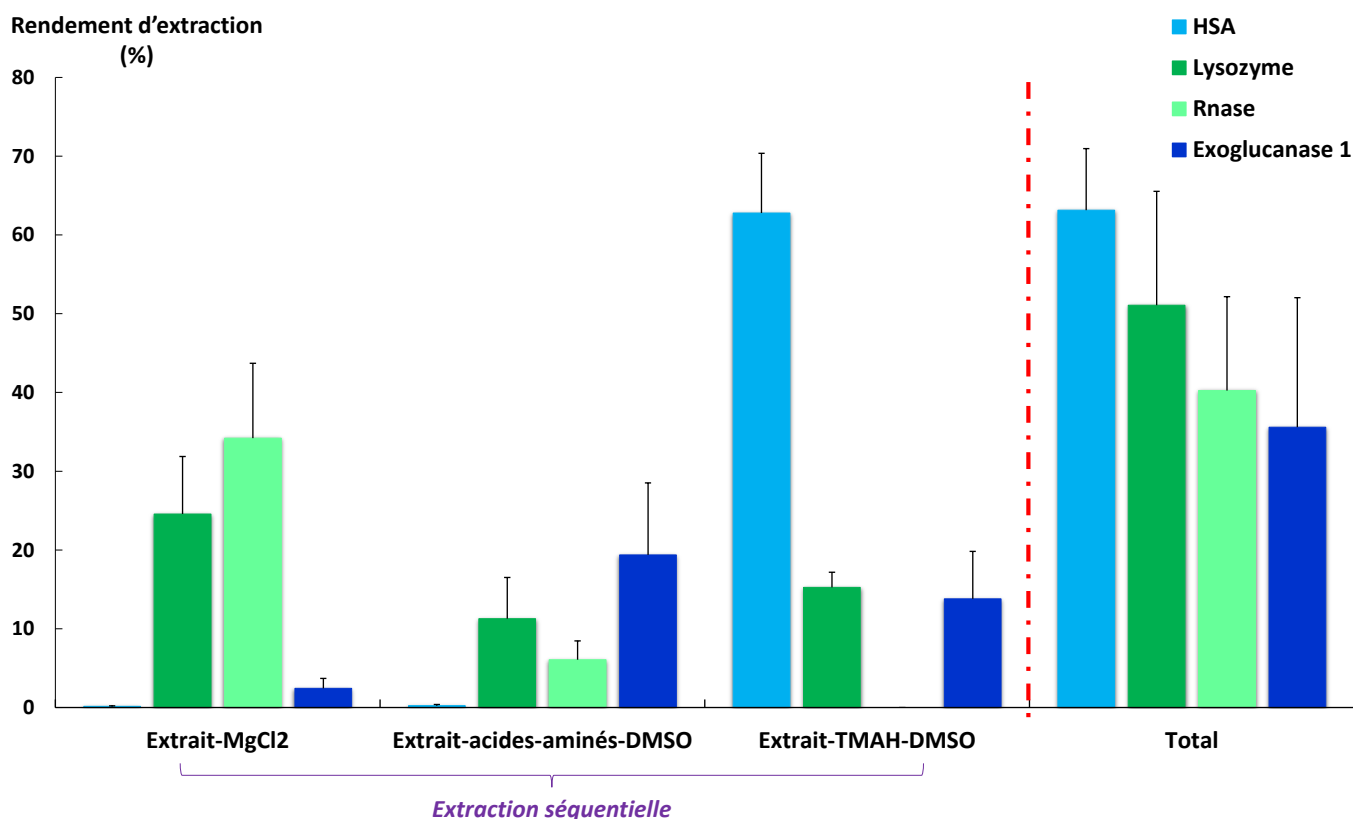


Figure 109. Extraction séquentielle : MgCl₂-Acides aminés/DMSO-TMAH-DMSO

Nous pouvons remarquer que l'extraction des protéines cibles est fonction du solvant utilisé au cours de la séquence. Ainsi, l'utilisation du MgCl₂, a permis d'extraire préférentiellement les petites protéines, lysozyme et Rnase avec des rendements de 25 % et 35 %, respectivement. Les protéines initialement adsorbées sur la matrice minérale du sol peuvent s'échanger avec les cations Mg²⁺ apportés par la solution MgCl₂ (Bastida et al., 2009). Le traitement aux acides aminés suivi d'une extraction au DMSO permet de récupérer 20 % de l'exoglucanase-1, 12 % de lysozyme et 6 % de Rnase. Enfin, un prétraitement au TMAH suivi d'une extraction au DMSO permet de récupérer plutôt la HSA (63 %), le lysozyme (15 %) et l'exoglucanase-1 (14 %).

Nous pouvons constater que les rendements obtenus après cette extraction séquentielle sont comparables à ceux obtenus après extraction avec le mélange DMSO/Tris-HCl/NaOH, sauf pour la HSA. Néanmoins, cette nouvelle approche utilise des solvants moins agressifs en début de séquence et permet de récupérer des extraits purifiés, et donc directement analysables par LC-MS/MS (et ce sans un traitement au phénol ni aux détergents ou aux chaotropes ou une étape de filtration). Par ailleurs, une série d'extractions successives basées sur le choix de différents réactifs chimiques permet d'extraire des protéines aux caractéristiques

Chapitre III : Etude des paramètres influençant l'extraction des protéines

physicochimiques différentes. Pour cette raison nous avons décidé de retenir l'extraction séquentielle MgCl₂-Acides aminés/DMSO-TMAH-DMSO.

2. Extraction des protéines endogènes dans des matrices environnementales et biologiques

Nous avons évalué l'efficacité de ce nouveau protocole pour l'extraction de protéines endogènes présentes dans le sol et dans un tissu biologique.

2.1. Matrice environnementale : sol du Deffend

Suivant le protocole ci-avant, nous avons étudié les protéines endogènes du sol après extraction au DMSO/Tris-HCl/NaOH et MgCl₂-Acides aminés/DMSO-TMAH-DMSO. Dans les deux cas, aucune protéine endogène n'a été identifiée (plus de 700 et 1200 peptides non assignés dans les deux extraits). Parmi ces peptides non assignés par la base de données *Mascot*, il y a certes, des interférents (contaminants protéiques), mais probablement des peptides « potentiels » non attribués par *Mascot* en raison du manque d'informations dans les bases de données protéiques (Renard et al., 2012). Les difficultés rencontrées lors de l'extraction des protéines endogènes résident en partie de leur faible concentration, mais également de la minéralogie du sol. En outre, la présence de molécules interférentes non protéiques (substances humiques et argileuses) peut masquer de manière significative le signal MS et induire une sous-estimation de la quantité réelle des protéines (Bastida et al., 2009; Keiblinger et al., 2012, 2016).

Nous avons initié une étude concernant la rétention des protéines par les particules d'argiles qu'il faudra par la suite poursuivre pour approfondir nos connaissances sur les différents mécanismes de rétention de ces composés dans les sols argileux (cf. annexe 8).

Pour valider notre protocole, nous avons alors étudié des échantillons riches en protéines et dont les protéines sont référencées dans les bases de données.

2.2. Matrice biologique : tumeur de souris

Dans cette expérience, nous avons étudié les protéines exprimées au sein d'une tumeur de carcinome pulmonaire (xénogreffe) chez la souris. L'extraction a été réalisée suivant le protocole développé au cours de ce travail et comparée, à une extraction classique avec un tampon phosphate salin (PBS) (pH 7,6). Cette solution est très utilisée pour extraire les marqueurs tumoraux extracellulaires (Haslène-Hox et al., 2011).

Chapitre III : Etude des paramètres influençant l'extraction des protéines

Après extraction, purification et analyse par LC-HRMS/MS, les recherches par similarité de séquences ont permis d'identifier 54 protéines après extraction au PBS et 33 protéines après extraction séquentielle. 19 protéines étaient communes aux deux protocoles (figure 110). Aucune protéine n'a été détectée dans le dernier extrait TMAH/DMSO du protocole développé.

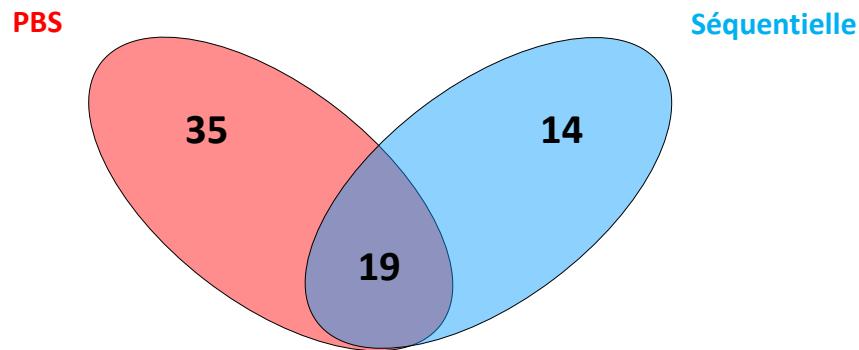


Figure 110. Diagramme de Venn représentant le nombre de protéines communes et différentes identifiées dans chaque extrait

Parmi les protéines communes, on trouve des protéines dites majoritaires, à savoir l'albumine (ALBU_MOUSE), et les hémoglobines (HBB1_MOUSE et HBB2_MOUSE). La présence de ces protéines majoritaires rend difficile la détection des marqueurs protéiques faiblement abondants (Frantzi et al., 2014; Schiess et al., 2009).

Il serait donc nécessaire pour cibler ces protéines de réaliser des traitements supplémentaires afin de dépléter ces protéines abondantes (approche d'immuno-capture en utilisant de kits de déplétion par exemple).

Parmi les protéines identifiées et pouvant être associées à l'état inflammatoire tumoral, nous pouvons citer, la protéine de choc thermique (Hsp7C), protéine de stress surexprimée dans de nombreux cancers y compris celui du poumon. Elle a été détectée dans les deux extraits (PBS et séquentiel).

Plus spécifiquement, dans l'extrait PBS, l'annexine A5 ainsi que l'apolipoprotéine AI (Apo-AI) ont été extraites. L'annexine joue un rôle capital dans de nombreux processus physiologiques et pathologiques (progression tumorale, invasion, métastase, résistance aux médicaments, etc.) (Peng et al., 2014). L'Apo-AI, est quant à elle un constituant majeur des lipoprotéines de haute densité (HDL) (métabolisme des lipides), module le système

Chapitre III : Etude des paramètres influençant l'extraction des protéines

immunitaire, et contrôle le développement des cancers (Zamanian-Daryoush et al., 2013; Zamanian-Daryoush and DiDonato, 2015).

Dans l'extrait séquentiel, des protéines de protection contre le stress oxydant ont été détectées (Peroxyrédoxine-1 et la superoxyde dismutase (SODC_Cu/Zn)) (Marikovsky et al., 2003).

De façon globale, l'approche d'extraction séquentielle n'a pas permis de récupérer une liste aussi longue de marqueurs potentiels comparée à l'approche classique (extraction au PBS). Nous pouvons envisager en perspectives de réadapter ce protocole à une matrice biologique. En effet, contrairement à la matrice « sol », le traitement alcalin (TMAH-DMSO) ne semble pas ici être nécessaire.

3. Conclusion

L'extraction des protéines à partir des milieux complexes (environnementaux et biologiques) constitue une étape très délicate en raison de la présence de molécules non protéiques ou de protéines majoritaires qui diminuent de manière significative la qualité des extraits protéiques. En outre, au sein des milieux environnementaux tel que le sol, les protéines sont fortement adsorbées à la matrice organo- minérale, ce qui rend leur extraction extrêmement difficile.

Dans ce chapitre, nous avons évalué l'efficacité de plusieurs paramètres physiques et chimiques sur l'extraction des protéines à partir d'une matrice environnementale, et ce dans le but de développer un protocole expérimental capable d'extraire des protéines aux propriétés physicochimiques variées. Ainsi, l'approche développée repose sur une extraction séquentielle au MgCl₂-Acides aminés/DMSO-TMAH-DMSO. L'originalité de ce protocole repose sur l'utilisation de solutions d'extraction exempte de détergents et d'agents chimiques agressifs, dont la présence pose un certain nombre de problèmes en cas d'une analyse directe par UPLC-HRMS/MS.

Conclusion générale et perspectives

Conclusion générale et perspectives

L'étude des biomarqueurs protéiques dans les milieux complexes (biologiques et environnementaux), sujet émergeant depuis quelques décennies, est développée afin de caractériser et d'appréhender la complexité du vivant dans son ensemble. Elle a toutefois mis en évidence des verrous analytiques liés notamment à la complexité des protéomes en termes de nombre de protéines, de diversité moléculaire et d'abondance relative.

Dans ce contexte, ce projet de thèse avait pour objectif d'améliorer l'identification et la quantification non ciblée de ces molécules dans des échantillons environnementaux et biologiques, en suivant une démarche intégrée au cours de laquelle chacune des étapes du processus global, allant du traitement d'échantillon (extraction/purification) à l'analyse UPLC-HRMS/MS, a été optimisée.

L'étape d'analyse par UPLC-HRMS/MS a dans un premier temps été étudiée. Elle a conduit au développement d'un système « *dual trap* » capable d'amplifier et d'améliorer de manière considérable la sensibilité du signal de MS de peptides. Dans un premier, les paramètres d'analyse HRMS/MS ont dû être fixés afin d'avoir une sensibilité et spécificité de détection vis-à-vis de peptides issus de la digestion de protéines modèles. Celles-ci ont en effet servi au développement des étapes amont de traitement d'échantillon.

Nous avons développé dans un deuxième temps une phase de purification incluant une précipitation des protéines suivie d'une seule et unique étape de lavage. La stratégie utilisée repose sur l'utilisation d'un nouvel agent, l'éthanol/HCl, qui permet la déstabilisation de l'état *molten globule* induit par la précipitation au TCA. L'augmentation de solubilité du culot protéique dans les tampons aqueux utilisés pour la digestion et l'analyse, induite par l'éthanol/HCl a été étudiée par une combinaison d'outils et de méthodes chimiques et biochimiques (Bradford, électrophorèse, AFM, Raman, spectrofluorométrie et UPLC-HRMS/MS). Comparé aux protocoles publiés (allant jusqu'à 4 étapes de lavage), ce protocole s'est révélé être pertinent pour l'analyse qualitative et quantitative de protéines issues de différents échantillons.

Enfin, dans une dernière partie nous avons étudié les différents paramètres expérimentaux influençant l'extraction des protéines à partir de matrices environnementales modèles puis complexes. Ce développement méthodologique repose sur l'emploi de solvants d'extraction compatibles avec l'analyse MS. Les résultats préliminaires ont montré qu'une approche séquentielle permettrait à la fois d'avoir une vue cartographie globale d'un protéome, tout en

Conclusion générale et perspectives

offrant la possibilité de cibler une fraction donnée en fonction de la solubilité des protéines dans chacun des extractants.

Si la purification nous semble aujourd'hui une étape maîtrisée, les étapes amont et aval sont encore à optimiser. En effet, le choix des extractants ainsi que le nombre de séquences devront être modulés et de nouvelles combinaisons devront être envisagées. Tel que souligné par nos résultats sur la purification, l'efficacité d'un traitement est directement associée à la structure, l'hydrophobicité et la localisation cellulaire des protéines. À titre d'exemple, dans le cas d'une analyse ciblée sur les protéines transmembranaires il pourrait être intéressant de combiner le lavage à l'éthanol/HCl à une étape de solubilisation utilisant l'acide formique ou des surfactants originaux tels que des polymères amphiphiles. L'objectif à terme, serait de proposer un protocole à la carte suivant la matrice étudiée et le questionnement scientifique associé. En outre, pour une analyse non ciblée, les paramètres de méthodes de masse devront être modifiés afin de détecter dans un même échantillon les peptides faiblement abondants dont les signaux peuvent être masqués. Alors que l'identification des protéines dans les échantillons biologiques peut s'appuyer sur des bases de données exhaustives, ce travail a mis en évidence qu'il n'en était de même pour les échantillons environnementaux, le sol en étant un exemple significatif.

Au-delà de son application proprement dite, ce travail a souligné l'intérêt fondamental d'étudier au niveau moléculaire les mécanismes chimiques mis en jeu entre les protéines et les agents utilisés lors des traitements. Le développement de l'étape d'extraction bénéficierait de cette même démarche scientifique, afin d'aboutir à un protocole efficace, dans le cadre d'une étude ciblée ou non ciblée.

Annexes

Annexes

Annexe 1: Détermination de la quantité totale des protéines par électrophorèse 1D et dosage Bradford

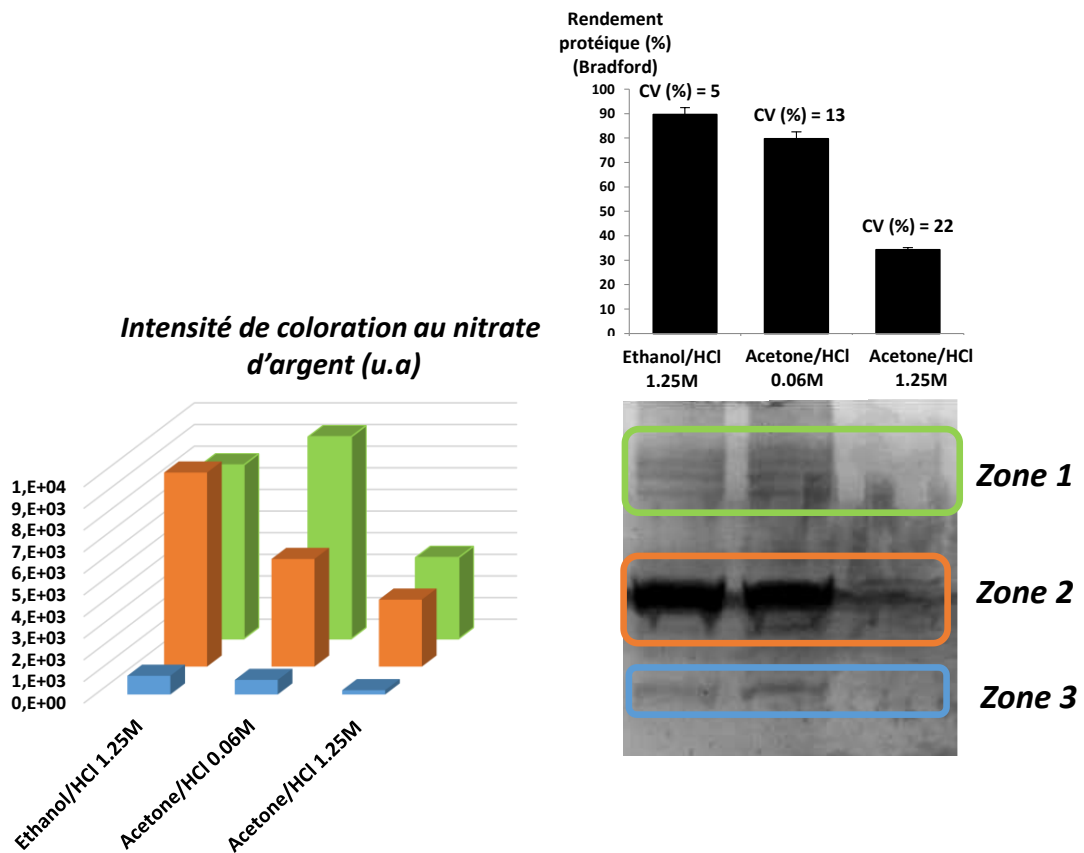


Figure 111. Electrophorèse 1D d'une solution de protéines standards (HSA, lysozyme, Rnase et exoglucanase-1) précipitée au TCA et lavée avec les différents solvants (éthanol/HCl 1,25 M, acétone/HCl 0,06 M et acétone/HCl 1,25 M). La révélation des bandes protéiques a été réalisée au nitrate d'argent. La quantité totale de protéines a également été déterminée par un dosage Bradford

Annexe 2: Spectres d'émission de la HSA, lysozyme et Rnase après précipitation au TCA et lavage à l'éthanol/HCl

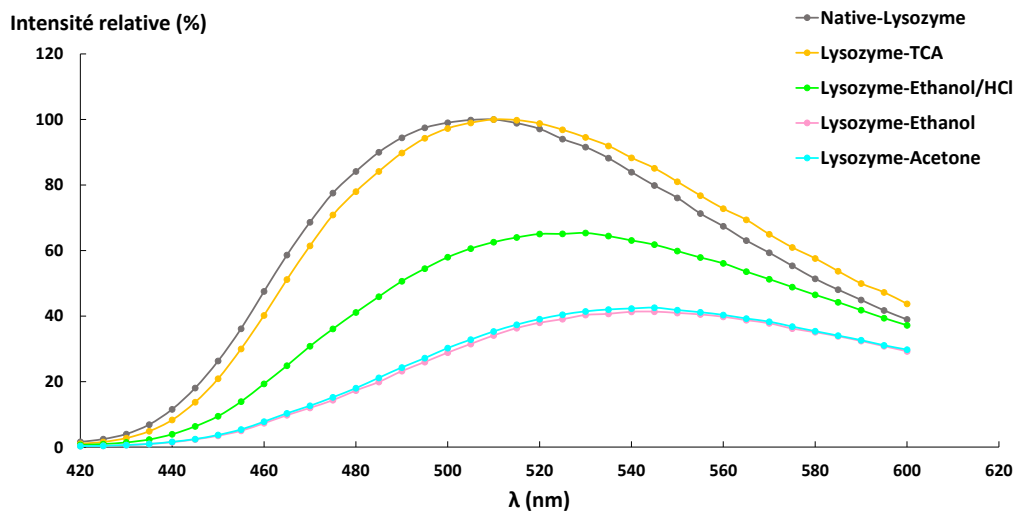


Figure 112. Spectre d'émission du lysozyme-ANS ($\lambda_{exc}=400\text{ nm}$) à différentes étapes de purification

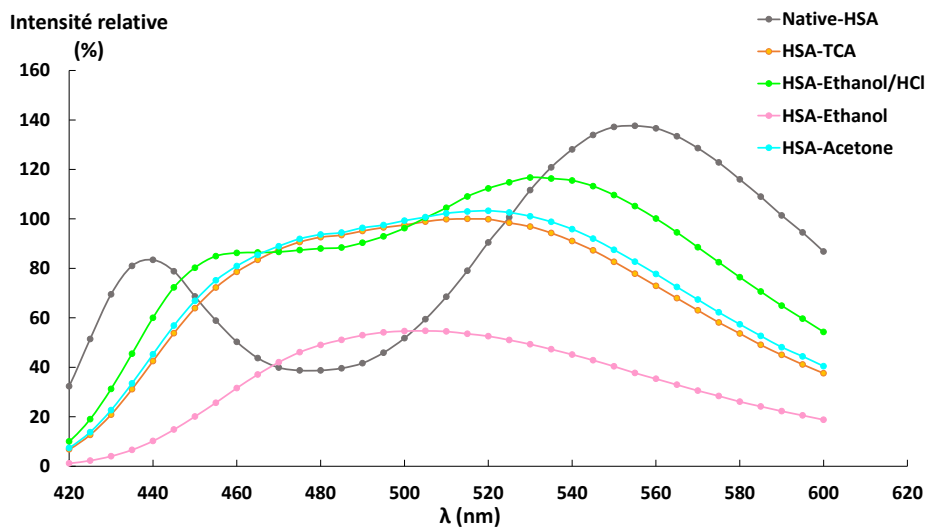


Figure 113. Spectre d'émission du HSA-ANS ($\lambda_{exc}=400\text{ nm}$) à différentes étapes de purification

Annexes

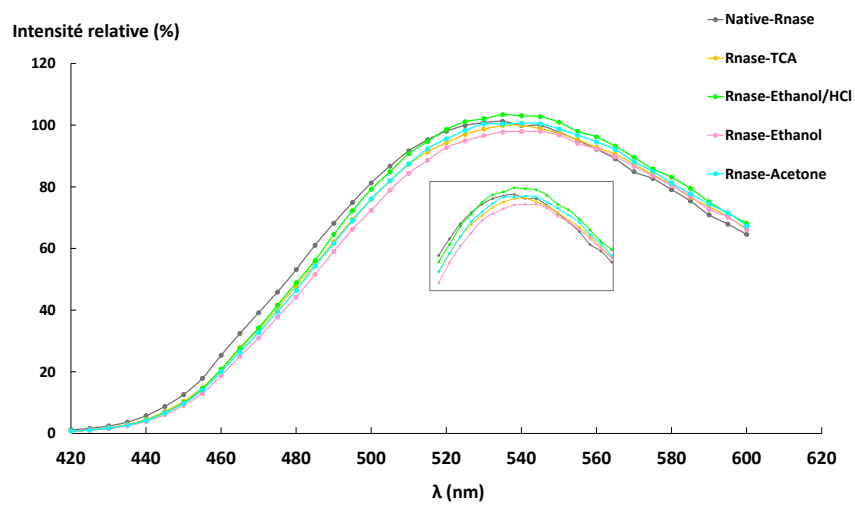


Figure 114. Spectre d'émission de la Rnase-ANS ($\lambda_{exc} = 400$ nm) à différentes étapes de purification

Annexe 3: Liste des peptides endogènes du biofilm quantifiés par Skyline avec les transitions PRM correspondantes

Tableau 16. Liste des peptides endogènes du biofilm quantifiés par Skyline avec les transitions PRM correspondantes

Protéine	Abréviation	Peptide	PRM	
			Précurseur (m/z)	Fragment (m/z)
R-phycoérythrine <i>Porphyra purpurea</i>	PHEA_PORP U	SVITTTISAADAAG R	717,3834	1134,5749
				1033,5273
			374,2146	
		FPSSDLESVQGNI QR	588,6235	715,3846
			621,2515	
			587,3260	
		NPGEAGDSQEK	566,2493	920,3956
				663,2944
				491,2460
C-phycoyanine- , <i>Synechococcus</i> <i>sp.</i>	PHCA1_SYNP 6	TPLTEAVAAADSQ GR	743,8784	1175,5651
				945,4748
			775,3693	
		FLSSTELQVAFGR	727,8855	1194,6113
			1107,5793	
			790,457	
C-phycoérythrine , <i>Synechocystis</i> <i>sp.</i>	PHEA_SYNY 1	TLGLPTAPYVEAL SFAR	602,6647	1152,6048
				793,4203
			664,3777	
		FPSTSDLESVQGSI QR	584,2917	688,3737
			635,2671	
			560,3151	
C-phycoerythrin , <i>Microchaete</i> <i>diplosiphon</i>	PHEA_MICDP	SVVTTVIAAADAA GR	701,3834	1116,6008
				815,437
			374,2146	
		ALGLPTAPYVEAL SFAR	592,6612	1152,6048
			793,4203	
			664,3777	

	FPSTSDLESVQGSIQR	584,2917	688,3737
			635,2671
			560,3151

Tableau 17. Liste des peptides endogènes du foie de souris quantifiés par Skyline avec les transitions PRM correspondantes

Protéine	Abréviati on	Peptide	PRM	
			Précurseur (<i>m/z</i>)	Fragment (<i>m/z</i>)
Carbamoyl- phosphate synthase	CPSM	TAVDSGIALLTNFQVTK	898,4844	950,5306
				837,4465
				736,3988
		VLGTSVESIMATEDR	804,4009	1051,4725
				722,3138
				591,2733
		AFAMTNQILVER	696,8688	972,5473
				516,3140
				403,2300
		GQNQPVLNITNR	677,3653	926,5418
				617,3365
				390,2096
		AADTIGYPVMIR	653,8448	835,4495
				615,3647
				472,2402
EPLFGISTGNIITGLAAGA K	644,0263	801,4829		
		688,3988		
		587,3511		
IALGIPLPEIK	582,3735	696,4291		
		355,2340		
		468,3180		
VMIGESIDEK	560,7814	890,4466		
		777,3625		
		231,1162		
SVGEVMAIGR	509,7711	832,4345		
		646,3705		
		547,3021		
Argininosucci nate synthase	ASSY	EQGYDVIAYLANIGQK	891,4571	977,5415
				743,4410
				630,3570
		FELTCYSLAPQIK	785,4027	1085,4972
				556,3453
				485,3082
QHGIPIVTPK	593,8508	921,5768		
		751,4713		

Annexes

				541,3344
		NQAPPGLYTK	544,7904	846,472
				775,4349
				314,1459
		YLLGTSLARPCAR	530,9643	657,8692
				601,3271
				277,1547
Protéine liant le sélénium	SBP2	GSFVLLDGETFEVK	770,8983	1037,515
				924,4309
				809,404
		EEIVYLPCIYR	727,871	984,4971
				821,4338
				708,3498
		LTGQIFLGGSIIVR	680,901	848,4989
				701,4304
				588,3464
		IYVVDVGSEPR	617,3273	957,5
				858,4316
				545,2678
		IFVWDWQR	575,2956	889,4315
				790,3631
				261,1598
		VIEASEIQAK	544,3033	875,4469
				746,4043
				675,3672
Glycéraldéhyde -3-phosphate déshydrogénase	G3P	VPTPNVSVVDLTCR	778,9087	1259,6412
				949,4771
				630,3243
		WGEAGAEYVVESTGVFT TMEK	764,3561	912,4495
				892,4123
				756,3597
		GAAQNIIPASTGAAK	685,3753	815,4621
				702,3781
				668,3726
		LISWYDNEYGYSNR	593,9373	1021,4625
				539,2572
				376,1939
Arginase-1	ARGI1	VMEETFSYLLGR	722,8607	1214,6052
				855,4723
				708,4039
		EGLYITEEIIYK	679,3479	1058,5405
				895,4771
				782,3931
		VSVVLGGDHSLAVGSIS GHAR	673,3641	866,9581
				817,4239
				760,8819
		SLEIIGAPFSK	581,3293	606,3246

Annexes

	556,3341
	478,266

Annexe 4: Extraction des protéines cibles à partir d'une matrice sable : effet d'un mélange de DMSO/Tris-HCl

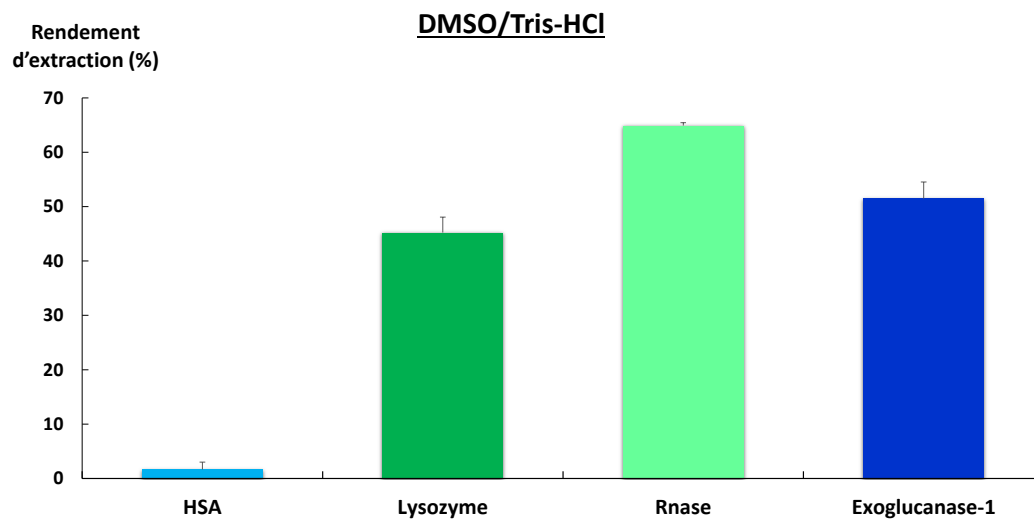


Figure 115. Extraction des protéines avec un mélange DMSO/Tris-HCl

Annexes

Annexe 5: Effet de la température sur l'extraction des protéines

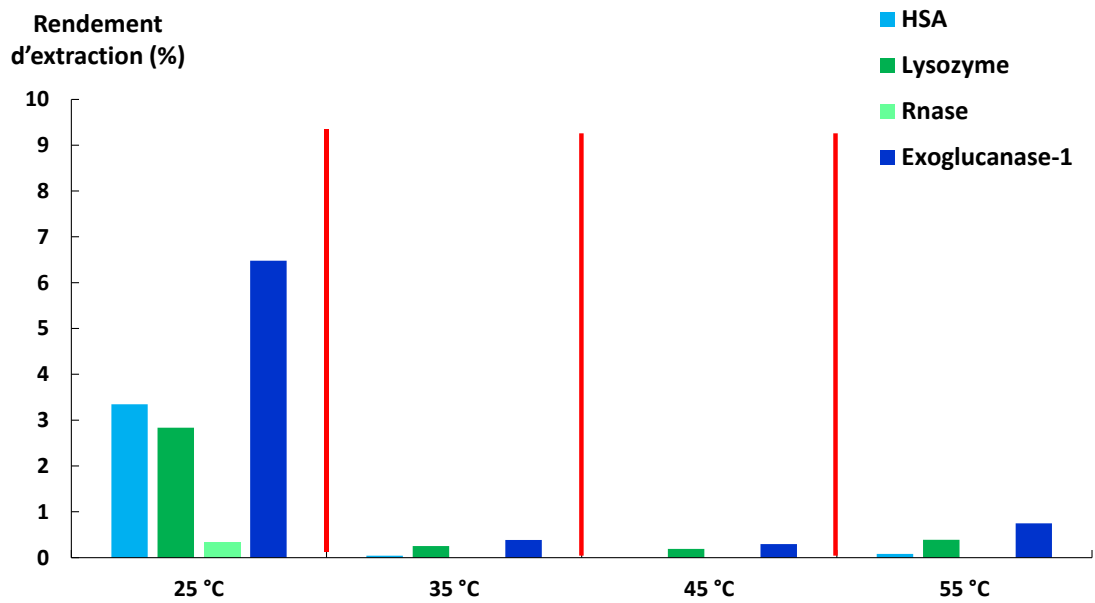


Figure 116. Effet de la température d'extraction (le temps d'extraction a été fixé à 5h00)

Annexes

Annexe 6: Effet du temps sur l'extraction des protéines

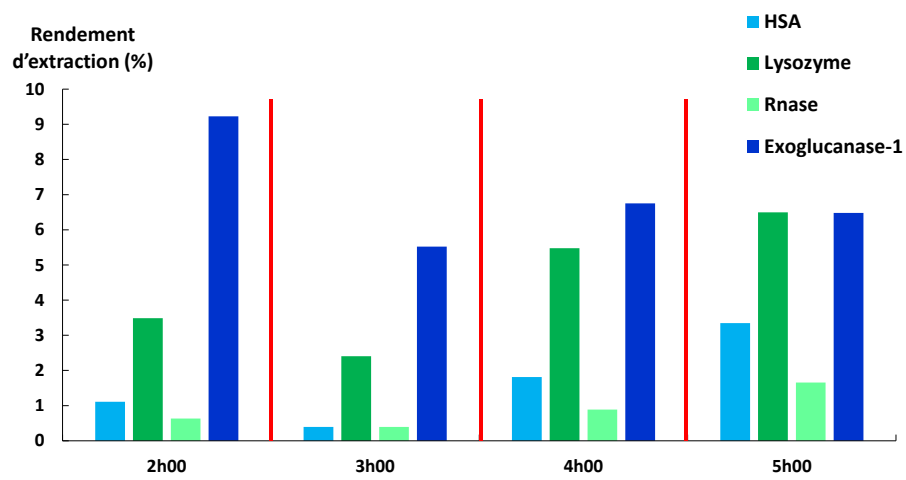


Figure 117. Effet du temps d'extraction (la température d'extraction a été fixée à 25 °C)

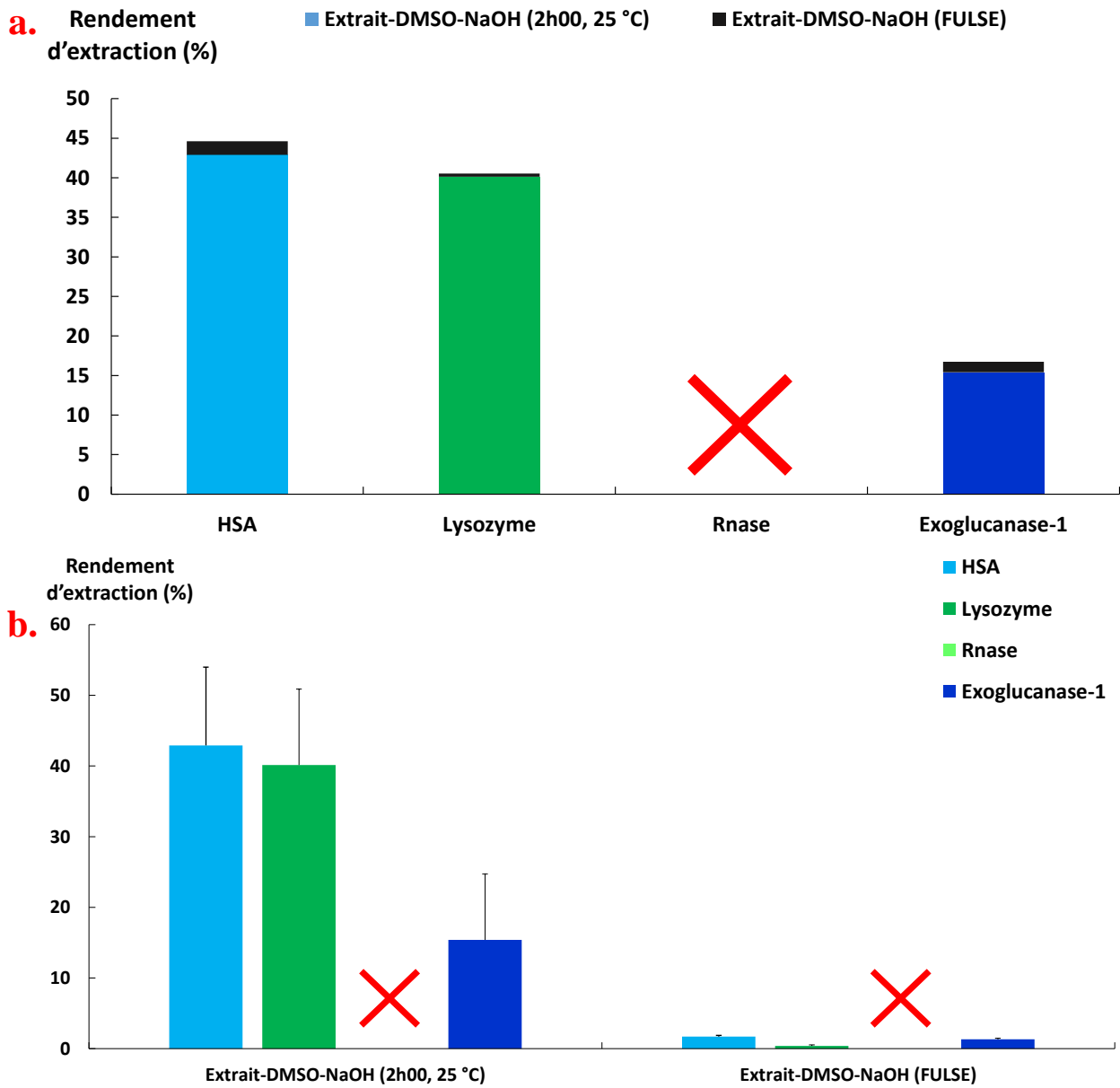


Figure 118. Effet d'une extraction assistée par ultrasons sur la récupération des protéines. Les deux histogrammes représentent les mêmes résultats, a. histogramme empilé ; b. histogramme mettant en évidence les barres d'erreurs pour chaque méthode

1. Analyse par diffraction aux rayons X (DRX)

Les rayons X sont des ondes électromagnétiques, qui lorsqu'ils interagissent avec le nuage électronique d'un atome de la matière, sont diffusés élastiquement induisant une réémission d'ondes électromagnétiques de même fréquence. Un réseau cristallin est caractérisé par la répétition triperiodique d'atomes ou de molécules. Lorsque les rayons X interagissent avec une matière, les rayons diffusés interfèrent entre eux. On parle alors d'interférences constructives lorsque les ondes diffusées par les atomes sont en phase et d'interférences destructrices lorsqu'elles sont en opposition de phase. Pour que la diffraction ait lieu, il faut que les ondes diffractées soient en phase (interférences constructives) donnant ainsi lieu à des pics de diffraction.

La caractérisation des phénomènes d'adsorption des protéines sur les argiles a été réalisée par diffraction aux rayons X. Cette technique de caractérisation non destructive fournit des informations pertinentes sur la structure cristalline des minéraux argileux.

1.2. Conditions expérimentales

Les expériences d'adsorption ont été réalisées sur la HSA et le lysozyme. Pour cela, une masse de 25 mg de chaque argile kaolinite et montmorillonite-sodique a été alors pesée puis solubilisée dans 1 mL d'une solution de HSA ou de lysozyme à 1 mg.mL^{-1} . L'échantillon a été homogénéisé pendant 2h00 à température ambiante dans l'agitateur thermostaté, puis centrifugé à 4000 g pendant 10 min à 20 °C (Ralla et al., 2010). Le surnageant a été ensuite éliminé puis analyser à 280 nm avec un spectromètre UV-visible. Le complexe argile-protéine a été repris dans 1 mL d'eau ultrapure puis déposé sur une lame de verre afin d'obtenir une préparation orientée pour l'étude des minéraux argileux à température ambiante. Un témoin a été aussi testé en remplaçant la solution de protéines par de l'eau ultrapure. Après séchage à l'air, une petite quantité de chaque échantillon a été broyée dans un mortier en agate. La poudre obtenue a été analysée par diffraction des DRX.

Les diffractogrammes ont été réalisés sur un diffractomètre BRUCKER D8 Advance, équipé d'un tube à anticathode au cuivre. Le détecteur est de type semi-conducteur solide permettant un comptage simultané sur un secteur angulaire de $3^\circ 2\theta$. Les conditions analytiques sont fixées à 40 kV et 40 mA avec un domaine angulaire d'analyse de $2,5$ à $35^\circ 2\theta$.

Annexes

1.3.Résultats et discussions

Les phénomènes d'adsorption des molécules organiques sur les minéraux argileux ont été longuement étudiés par de nombreux scientifiques. Ces dernières interviennent dans de nombreux processus de biodégradation ou de stabilisation de matières organiques de nature protéiques indispensables à la vie (enzymes par exemple) ou encore de protection de l'activité biologique de nombreuses communautés microbiennes. En effet, la teneur des minéraux argileux (exemple les phyllosilicates : kaolinite et montmorillonite) présents dans les milieux environnementaux tels que les sols, présente un impact majeur sur la biodégradation de ces molécules organiques qui peut être soit accélérée, soit ralentie. Ces processus sont liés principalement à des effets de protection et à la formation de complexe entre l'argile et la molécule organique. À titre d'exemple, la fixation des protéines enzymatiques ou non enzymatiques sur les minéraux argileux protège leurs activités contre les agressions environnementales externes, elle constitue aussi un réservoir important du carbone terrestre (Schmidt and Martínez, 2016). La présence des argiles protège la viabilité des certains microorganismes tels que le genre *Rhizobium* ou *Azobacter* et améliore ainsi leur résistance vis-à-vis des phénomènes de dessiccation ou d'élévation de température. L'adsorption aux argiles conserve aussi certaines protéines à caractère pathogène telles que les prions responsables des encéphalopathies spongiformes transmissibles (Wyckoff et al., 2013). Tous ces phénomènes protecteurs diminuent néanmoins de manière considérable la disponibilité des protéines et rendent difficiles leur extraction ainsi que leur analyse à partir des matrices environnementales, entraînent de ce fait, une sous-estimation des quantités réelles dans les sols.

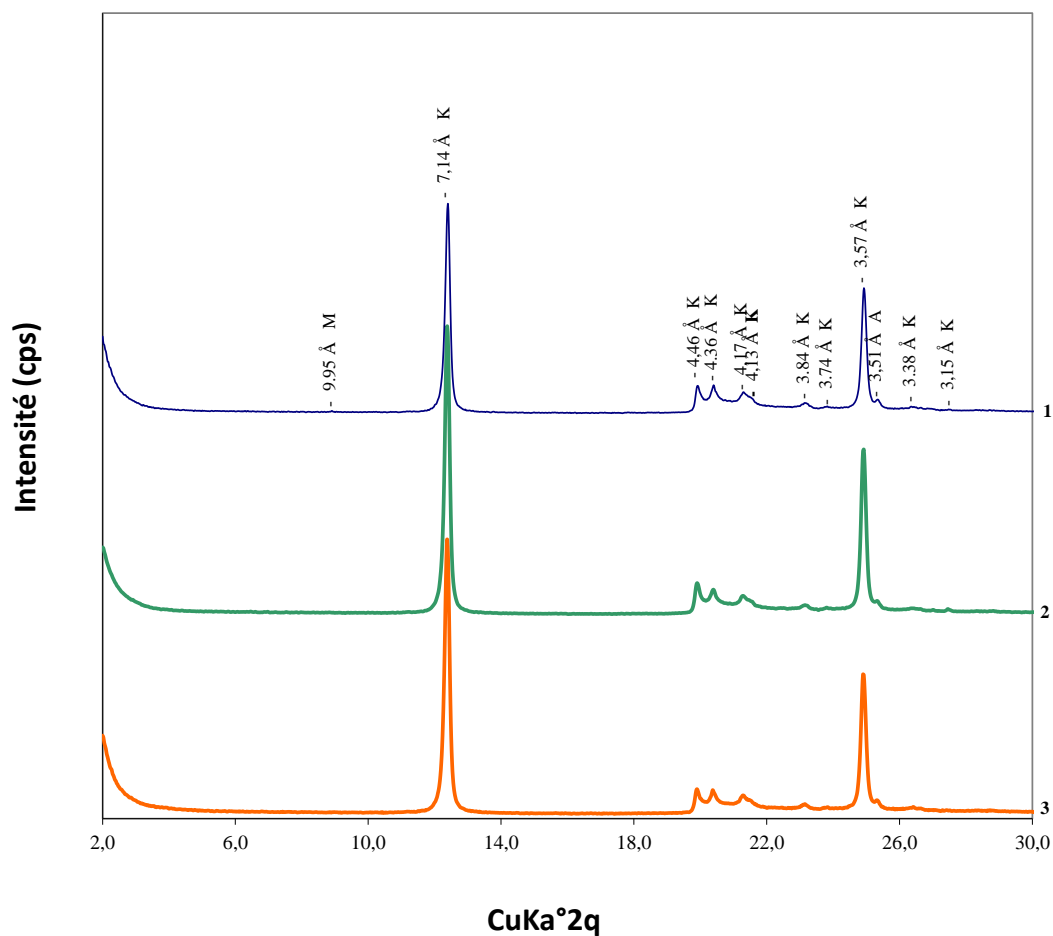
En science des sols, les argiles correspondent à la fraction granulométrique inférieure à 2 μm et sont principalement dominées par les phyllosilicates, silicates en feuillets. Ces minéraux appartenant au groupe des silicates sont caractérisés par la superposition de couches tétraédriques et de couches octaédriques. Celles-ci comportent des oxygènes ou des groupements hydroxyles à leur sommet et des cations différents (Al, Mg, Fe etc.) en leur centre (Caner, 2011). Dans notre étude nous avons étudié les phénomènes d'adsorption des protéines sur deux types d'argiles standards, la kaolinite (K) et la montmorillonite sodique (Na-M).

La kaolinite appartient au groupe des aluminosilicates où certains atomes de silicium ont été remplacés par des atomes d'aluminium. Il s'agit d'un feuillet non expansible après hydratation ou solvatation à l'éthylène glycol, et de ce fait, n'est pas caractérisé par la capacité d'expansion dans l'espace inter-foliaire. Les protéines auront tendance à s'adsorber à la surface de l'argile. Contrairement à la montmorillonite qui a un caractère gonflant et a donc la capacité d'accueillir

Annexes

les molécules polaires ou chargées au sein de l'espace inter-foliaire. La taille des protéines joue un rôle important dans les mécanismes d'adsorption. À titre d'exemple, le lysozyme s'intercale aisément entre les feuillets de la montmorillonite *via* des mécanismes d'échange cationique contrairement l'albumine ou la β -glucosidase, qui se fixent plutôt à la surface. Les mécanismes d'interaction sont principalement dus à des échanges cationiques, des interactions électrostatiques, des interactions hydrophiles ou hydrophobes, des liaisons hydrogènes etc. Ces interactions dépendent fortement du pH du milieu (Yu et al., 2013b).

Afin de mieux appréhender les mécanismes d'interactions entre les protéines et leur milieu au niveau moléculaire, nous avons étudié dans un premier temps, les phénomènes d'adsorption de protéines (la HSA et le lysozyme) utilisées comme modèles d'encapsulation ou de piégeage dans les matrices organo-minérales (Fiorito et al., 2008; Ralla et al., 2010; Schmidt and Martínez, 2016; Tomaszewski et al., 2011). Ces deux protéines ont été adsorbées sur deux types d'argiles standards, la kaolinite (K) et la montmorillonite sodique (Na-M). Les diffractogrammes relatifs aux expériences de sorption sur la K et la Na-M sont présentés, respectivement sur les figures 119 et 120.



Annexes

Figure 119. Diffractogrammes de l'échantillon Kaolinite préparé sur lame orientée séchée à l'air ; (1) témoin : après contact à l'eau ultrapure ; (2) en présence d'une solution de HSA à 1 mg.mL⁻¹ ; (3) en présence d'une solution de lysozyme à 1 mg.mL⁻¹. Minéraux : (A) Anatase, (K) Kaolinite, (M) Mica

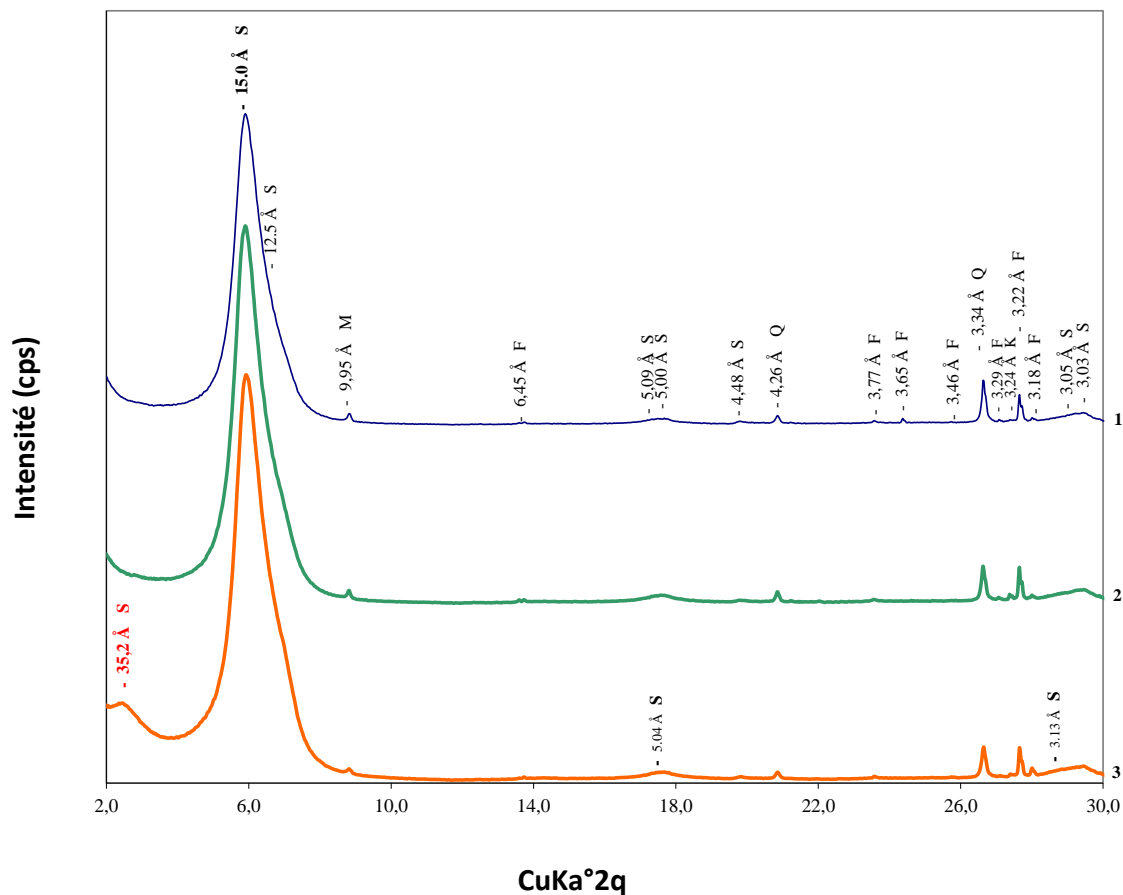


Figure 120. Diffractogrammes de l'échantillon montmorillonite préparé sur lame orientée séchée à l'air ; (1) témoin : après contact à l'eau ultrapure ; (2) en présence d'une solution de HSA à 1 mg.mL⁻¹ ; (3) en présence d'une solution de lysozyme à 1 mg.mL⁻¹. Minéraux : (F) Feldspath, (Q) Quartz, (M) Mica, (S) Smectite

L'étude de diffraction de rayons X a montré qu'en présence de kaolinite (Figure 1), aucune formation de complexe intercalaire kaolinite-protéine n'est mise en évidence. Toutefois, des phénomènes d'adsorption de surface au niveau des bordures des feuillets de la kaolinite peuvent être observés. Dans une étude menée par Barral et al. (2008), ont démontré que de la BSA (albumine sérique bovine), α -lactalbumine et β -lactalbumine interagissent *via* des interactions électrostatiques avec les groupements chimiques situés au niveau de la surface de la kaolinite (Barral et al., 2008). De même le lysozyme semble s'adsorber sur la surface de la kaolinite.

Annexes

Contrairement à kaolinite, la montmorillonite sodique est une smectite expansible, c'est-à-dire que des phénomènes de sorption peuvent se produire dans l'espace inter-foliaire ou au niveau de la surface de l'argile. Concernant la nature des cations inter-foliaires naturellement présents dans la montmorillonite, il ne s'agit pas d'une smectite purement sodique, contrairement à ce qui est indiqué sur le produit acheté, mais d'une smectite sodi-calcique pour laquelle on distingue nettement une réflexion à 15,0 Å (feuillets calciques à 2 couches d'eau) et une réflexion à 12,5 Å (feuillets sodiques à 1 couche d'eau) (Figure 2). D'après la figure 2, la HSA ne semble pas s'intercaler dans l'espace inter-foliaire, contrairement au lysozyme, comme l'atteste le diffractogramme (3), nous avons une apparition d'une réflexion vers 35,2 Å. Compte tenu de la dimension intrinsèque d'un feuillet de smectite (10,0 à 9,6 Å), le lysozyme semble occuper une épaisseur d'environ de 25 Å. Johnston et collaborateurs (2012), ont étudié les interactions entre une saponite sodique et le lysozyme, ils ont mis en évidence une intercalation de la protéine entre les feuillets de la Na-saponite. L'espacement est estimé à 44 Å, ce qui correspond à une épaisseur de 34,4 Å dans le cas où le lysozyme est orienté parallèlement à la surface basale de l'argile (Johnston et al., 2012). Ainsi l'espacement dans l'espace inter-foliaire va dépendre de l'orientation de la protéine au sein des feuillets. Par ailleurs, lorsqu'on suit le ratio des surfaces des réflexions à 15,0 Å et 12,5 Å entre la montmorillonite témoin et la montmorillonite-lysozyme, nous pouvons remarquer que l'apparition de la réflexion à 35,2 Å s'accompagne de la diminution de celle à 12,5 Å (près de 25% de son intensité). Nous pouvons donc se demander si la sorption n'aurait pas été meilleure avec une montmorillonite purement sodique.

Références bibliographiques

Références bibliographiques

- Adoubel, A.A., Guenu, S., Elfakir, C., Dreux, M., 2000. Separation of underivatized small peptides on a porous graphitic carbon column by ion-pair chromatography and evaporative light scattering detection. *J. Liq. Chromatogr. Relat. Technol.* 23, 2433–2446. <https://doi.org/10.1081/JLC-100100499>
- Aebersold, R., Mann, M., 2016. Mass-spectrometric exploration of proteome structure and function. *Nature* 537, 347–355. <https://doi.org/10.1038/nature19949>
- Ahn, Y.H., Kim, J.Y., Yoo, J.S., 2015. Quantitative mass spectrometric analysis of glycoproteins combined with enrichment methods. *Mass spectrom. Rev.* 34, 148–165. <https://doi.org/10.1002/mas.21428>
- Albani JR, J.R., 2001. Absorption et fluorescence, Edition Tec Doc Lavoisier. ed, Edition Tec Doc Lavoisier.
- Alushin, G.M., Lander, G.C., Kellogg, E.H., Zhang, R., Baker, D., Nogales, E., 2014. High-resolution microtubule structures reveal the structural transitions in $\alpha\beta$ -tubulin upon GTP hydrolysis. *Cell* 157, 1117–1129. <https://doi.org/10.1016/j.cell.2014.03.053>
- Anderson, M., Bocharova, O.V., Makarava, N., Breydo, L., Salnikov, V.V., Baskakov, I.V., 2006. Polymorphism and ultrastructural organization of prion protein amyloid fibrils: an insight from high resolution atomic force microscopy. *J. Mol. Biol.* 358, 580–596. <https://doi.org/10.1016/j.jmb.2006.02.007>
- Anderson, N.L., Anderson, N.G., 1998. Proteome and proteomics: New technologies, new concepts, and new words. *ELECTROPHORESIS* 19, 1853–1861. <https://doi.org/10.1002/elps.1150191103>
- Apostol, M.I., Perry, K., Surewicz, W.K., 2013. Crystal structure of a human prion protein fragment reveals a motif for oligomer formation. *J. Am. Chem. Soc.* 135, 10202–10205. <https://doi.org/10.1021/ja403001q>
- Araújo, J.E., Oliveira, E., Kouvonen, P., Corthals, G.L., Lodeiro, C., Santos, H.M., Capelo, J.L., 2014. A journey through PROTEOSONICS. *Talanta* 121, 71–80. <https://doi.org/10.1016/j.talanta.2013.12.054>
- Arca-Ramos, A., Ammann, E.M., Gasser, C.A., Nastold, P., Eibes, G., Feijoo, G., Lema, J.M., Moreira, M.T., Corvini, P.F.-X., 2016. Assessing the use of nanoimmobilized laccases to remove micropollutants from wastewater. *Environ. Sci. Pollut. Res.* 23, 3217–3228. <https://doi.org/10.1007/s11356-015-5564-6>
- Aresta, A., Calvano, C.D., Palmisano, F., Zambonin, C.G., Monaco, A., Tommasi, S., Pilato, B., Paradiso, A., 2008. Impact of sample preparation in peptide/protein profiling in human serum by MALDI-TOF mass spectrometry. *J. Pharm. Biomed. Anal.* 46, 157–164. <https://doi.org/10.1016/j.jpba.2007.10.015>
- Arnal, I., Wade, R.H., 1995. How does taxol stabilize microtubules? *Curr. Biol.* 5, 900–908. [https://doi.org/10.1016/S0960-9822\(95\)00180-1](https://doi.org/10.1016/S0960-9822(95)00180-1)
- Ashton, L., Blanch, E.W., 2010. pH-induced conformational transitions in α -lactalbumin investigated with two-dimensional Raman correlation variance plots and moving windows. *J. Mol. Struct.* 974, 132–138.
- Aslebagh, R., Pfeffer, B.A., Fliesler, S.J., Darie, C.C., 2016. Mass spectrometry-based proteomics of oxidative stress: Identification of 4-hydroxy-2-nonenal (HNE) adducts of amino acids using lysozyme and bovine serum albumin as model proteins. *Electrophoresis* 37, 2615–2623. <https://doi.org/10.1002/elps.201600134>
- Baker, E.N., Hubbard, R.E., 1984. Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* 44, 97–179. [https://doi.org/10.1016/0079-6107\(84\)90007-5](https://doi.org/10.1016/0079-6107(84)90007-5)
- Bala, M., Gupta, S., Pasha, M.A., 2000. Angiotensin-converting enzyme assay optimization: influence of various buffers and their concentrations. *Clin. Biochem.* 33, 687–689.
- Bannister, A.J., Kouzarides, T., 2011. Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395. <https://doi.org/10.1038/cr.2011.22>

Références bibliographiques

- Barral, S., Villa-García, M.A., Rendueles, M., Díaz, M., 2008. Interactions between whey proteins and kaolinite surfaces. *Acta Mater.* 56, 2784–2790. <https://doi.org/10.1016/j.actamat.2008.02.009>
- Basak, T., Tanwar, V.S., Bhardwaj, G., Bhardwaj, N., Ahmad, S., Garg, G., V, S., Karthikeyan, G., Seth, S., Sengupta, S., 2016. Plasma proteomic analysis of stable coronary artery disease indicates impairment of reverse cholesterol pathway. *Sci. Rep.* 6, 28042. <https://doi.org/10.1038/srep28042>
- Bastida, F., Hernández, T., García, C., 2014. Metaproteomics of soils from semiarid environment: Functional and phylogenetic information obtained with different protein extraction methods. *J. Proteomics* 101, 31–42. <https://doi.org/10.1016/j.jprot.2014.02.006>
- Bastida, F., Moreno, J.L., Nicolás, C., Hernández, T., García, C., 2009. Soil metaproteomics: a review of an emerging environmental science. Significance, methodology and perspectives. *Eur. J. Soil Sci.* 60, 845–859.
- Bastida, F., Selevsek, N., Torres, I.F., Hernández, T., García, C., 2015. Soil restoration with organic amendments: linking cellular functionality and ecosystem processes. *Sci. Rep.* 5, 15550. <https://doi.org/10.1038/srep15550>
- Bateman, N.W., Goulding, S.P., Shulman, N.J., Gadok, A.K., Szumlinski, K.K., MacCoss, M.J., Wu, C.C., 2014. Maximizing peptide identification events in proteomic workflows using data-dependent acquisition (DDA). *Mol. Cell. Proteomics MCP* 13, 329–338. <https://doi.org/10.1074/mcp.M112.026500>
- Becher, F., Duriez, E., Volland, H., Tabet, J.C., Ezan, E., 2007. Detection of Functional Ricin by Immunoaffinity and Liquid Chromatography–Tandem Mass Spectrometry. *Anal. Chem.* 79, 659–665. <https://doi.org/10.1021/ac061498b>
- Benndorf, D., Balcke, G.U., Harms, H., von Bergen, M., 2007. Functional metaproteome analysis of protein extracts from contaminated soil and groundwater. *ISME J* 1, 224–234.
- Benndorf, D., Vogt, C., Jehmlich, N., Schmidt, Y., Thomas, H., Woffendin, G., Shevchenko, A., Richnow, H.-H., von Bergen, M., 2009. Improving protein extraction and separation methods for investigating the metaproteome of anaerobic benzene communities within sediments. *Biodegradation* 20, 737–750. <https://doi.org/10.1007/s10532-009-9261-3>
- Beretov, J., Wasinger, V.C., Millar, E.K.A., Schwartz, P., Graham, P.H., Li, Y., 2015. Proteomic Analysis of Urine to Identify Breast Cancer Biomarker Candidates Using a Label-Free LC-MS/MS Approach. *PLoS ONE* 10. <https://doi.org/10.1371/journal.pone.0141876>
- Berg, J.M., Tymoczko, J.L., Stryer, L., Berg, J.M., Tymoczko, J.L., Stryer, L., 2012. *Biochemistry*, 7th ed. W H Freeman.
- Bertrand, M., 2009. Origine exogène des acides aminés prébiotiques. Orléans.
- Bertrand, M., Chabin, A., Brack, A., Cottin, H., Chaput, D., Westall, F., 2012. The PROCESS experiment: exposure of amino acids in the EXPOSE-E experiment on the international space station and in laboratory simulations. *Astrobiology* 12, 426–435. <https://doi.org/10.1089/ast.2011.0755>
- Biomarkers Definitions Working Group, 2001. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clin. Pharmacol. Ther.* 69, 89–95. <https://doi.org/10.1067/mcp.2001.113989>
- Blackstock, W.P., Weir, M.P., 1999. Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol.* 17, 121–127. [https://doi.org/10.1016/S0167-7799\(98\)01245-1](https://doi.org/10.1016/S0167-7799(98)01245-1)
- Blonder, J., Chan, K.C., Issaq, H.J., Veenstra, T.D., 2007. Identification of membrane proteins from mammalian cell/tissue using methanol-facilitated solubilization and tryptic digestion coupled with 2D-LC-MS/MS. *Nat Protoc.* 1, 2784–2790. <https://doi.org/10.1038/nprot.2006.359>

Références bibliographiques

- Boix, C., Ibáñez, M., Sancho, J.V., Rambla, J., Aranda, J.L., Ballester, S., Hernández, F., 2015. Fast determination of 40 drugs in water using large volume direct injection liquid chromatography-tandem mass spectrometry. *Talanta* 131, 719–727. <https://doi.org/10.1016/j.talanta.2014.08.005>
- Borozan, S.Z., Stojanović, S.Đ., 2013. Halogen bonding in complexes of proteins and non-natural amino acids. *Comput. Biol. Chem.* 47, 231–239. <https://doi.org/10.1016/j.compbiolchem.2013.10.002>
- Bowles, J.E., 1984. *Physical and geotechnical properties of soils*. McGraw-Hill.
- Brändén, C.-I., Tooze, J., 1999. *Introduction to Protein Structure*. Taylor & Francis.
- Bruderer, R., Bernhardt, O.M., Gandhi, T., Miladinović, S.M., Cheng, L.-Y., Messner, S., Ehrenberger, T., Zanotelli, V., Butscheid, Y., Escher, C., Vitek, O., Rinner, O., Reiter, L., 2015. Extending the Limits of Quantitative Proteome Profiling with Data-Independent Acquisition and Application to Acetaminophen-Treated Three-Dimensional Liver Microtissues. *Mol. Cell. Proteomics MCP* 14, 1400–1410. <https://doi.org/10.1074/mcp.M114.044305>
- Bruins, A.P., 1998. Mechanistic aspects of electrospray ionization. *J. Chromatogr. A* 794, 345–357. [https://doi.org/10.1016/S0021-9673\(97\)01110-2](https://doi.org/10.1016/S0021-9673(97)01110-2)
- Brundin, P., Melki, R., Kopito, R., 2010. Prion-like transmission of protein aggregates in neurodegenerative diseases. *Nat. Rev. Mol. Cell Biol.* 11, 301–307. <https://doi.org/10.1038/nrm2873>
- Butler, H.J., Ashton, L., Bird, B., Cinque, G., Curtis, K., Dorney, J., Esmonde-White, K., Fullwood, N.J., Gardner, B., Martin-Hirsch, P.L., Walsh, M.J., McAinsh, M.R., Stone, N., Martin, F.L., 2016. Using Raman spectroscopy to characterize biological materials. *Nat Protoc.* 11, 664–687.
- Callesen, A.K., Madsen, J.S., Vach, W., Kruse, T.A., Mogensen, O., Jensen, O.N., 2009. Serum protein profiling by solid phase extraction and mass spectrometry: A future diagnostics tool? *PROTEOMICS* 9, 1428–1441. <https://doi.org/10.1002/pmic.200800382>
- Cañas, B., Piñeiro, C., Calvo, E., López-Ferrer, D., Gallardo, J.M., 2007. Trends in sample preparation for classical and second generation proteomics. *Adv. Sample Prep. II* 1153, 235–258. <https://doi.org/10.1016/j.chroma.2007.01.045>
- Capriotti, A.L., Cavaliere, C., Cavazzini, A., Gasparrini, F., Pierri, G., Piovesana, S., Laganà, A., 2017. A multidimensional liquid chromatography–tandem mass spectrometry platform to improve protein identification in high-throughput shotgun proteomics. *J. Chromatogr. A, High-Resolution Separations of Proteins and Peptides* 1498, 176–182. <https://doi.org/10.1016/j.chroma.2017.03.032>
- Carvalho, V., Cerca, N., Vilanova, M., Vitorino, R., 2015. Proteomic profile of dormancy within *Staphylococcus epidermidis* biofilms using iTRAQ and label-free strategies. *Appl. Microbiol. Biotechnol.* 99, 2751–2762. <https://doi.org/10.1007/s00253-015-6434-3>
- Castillo, M.J., Reynolds, K.J., Gomes, A., Fenselau, C., Yao, X., 2014. Quantitative Protein Analysis Using Enzymatic [18O]Water Labeling. *Curr. Protoc. Protein Sci. Editor. Board John E Coligan Al* 76, 23.4.1–23.4.9. <https://doi.org/10.1002/0471140864.ps2304s76>
- Catherman, A.D., Skinner, O.S., Kelleher, N.L., 2014. Top Down proteomics: facts and perspectives. *Biochem. Biophys. Res. Commun.* 445, 683–693. <https://doi.org/10.1016/j.bbrc.2014.02.041>
- Cech, N.B., Enke, C.G., 2001. Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrom. Rev.* 20, 362–387. <https://doi.org/10.1002/mas.10008>
- Célérier, J., 2008. Caractérisation moléculaire et dynamique de la matière organique de compost (déchets verts, biodéchets) dans un sol. Université de Poitiers.

Références bibliographiques

- Chaimbault, P., Petritis, K., Elfakir, C., Dreux, M., 2000. Ion-pair chromatography on a porous graphitic carbon stationary phase for the analysis of twenty underivatized protein amino acids. *J. Chromatogr. A* 870, 245–254. [https://doi.org/10.1016/S0021-9673\(99\)00863-8](https://doi.org/10.1016/S0021-9673(99)00863-8)
- Chang, N., Hen, S.J., Klibanov, A.M., 1991. Protein separation and purification in neat dimethyl sulfoxide. *Biochem. Biophys. Res. Commun.* 176, 1462–1468.
- Chaturongakul, S., Ounjai, P., 2014. Phage–host interplay: examples from tailed phages and Gram-negative bacterial pathogens. *Front. Microbiol.* 5. <https://doi.org/10.3389/fmicb.2014.00442>
- Chemat, F., Rombaut, N., Sicaire, A.-G., Meullemiestre, A., Fabiano-Tixier, A.-S., Abert-Vian, M., 2017. Ultrasound assisted extraction of food and natural products. Mechanisms, techniques, combinations, protocols and applications. A review. *Ultrason. Sonochem.* 34, 540–560. <https://doi.org/10.1016/j.ultsonch.2016.06.035>
- Chen, E.I., Cociorva, D., Norris, J.L., Yates, J.R., 2007. Optimization of Mass Spectrometry Compatible Surfactants for Shotgun Proteomics. *J. Proteome Res.* 6, 2529–2538. <https://doi.org/10.1021/pr060682a>
- Chen, M., Wang, Y., Zhao, Y., Wang, L., Gong, J., Wu, L., Gao, X., Yang, Z., Qian, L., 2009. Dynamic proteomic and metabonomic analysis reveal dysfunction and subclinical injury in rat liver during restraint stress. *Biochim. Biophys. Acta BBA - Proteins Proteomics* 1794, 1751–1765. <https://doi.org/10.1016/j.bbapap.2009.08.012>
- Chen, W., Westerhoff, P., Leenheer, J.A., Booksh, K., 2003. Fluorescence excitation-emission matrix regional integration to quantify spectra for dissolved organic matter. *Environ. Sci. Technol.* 37, 5701–5710.
- Chen, Y.-Y., Lin, S.-Y., Yeh, Y.-Y., Hsiao, H.-H., Wu, C.-Y., Chen, S.-T., Wang, A.H.-J., 2005. A modified protein precipitation procedure for efficient removal of albumin from serum. *Electrophoresis* 26, 2117–2127.
- Chevalier, F., 2010. Standard Dyes for Total Protein Staining in Gel-Based Proteomic Analysis. *Materials* 3, 4784–4792. <https://doi.org/10.3390/ma3104784>
- Chevalier, F., Centeno, D., Rofidal, V., Tauzin, M., Martin, O., Sommerer, N., Rossignol, M., 2006. Different Impact of Staining Procedures Using Visible Stains and Fluorescent Dyes for Large-Scale Investigation of Proteomes by MALDI-TOF Mass Spectrometry. *J. Proteome Res.* 5, 512–520. <https://doi.org/10.1021/pr050194n>
- Chevallet, M., Lescuyer, P., Diemer, H., van Dorsselaer, A., Leize-Wagner, E., Rabilloud, T., 2006. Alterations of the mitochondrial proteome caused by the absence of mitochondrial DNA: A proteomic view. *Electrophoresis* 27, 1574–1583. <https://doi.org/10.1002/elps.200500704>
- Chevallet, M., Santoni, V., Poinas, A., Rouquié, D., Fuchs, A., Kieffer, S., Kieffer, S., Lunardi, J., Garin, J., Rabilloud, T., 1998. New zwitterionic detergents improve the analysis of membrane proteins by two-dimensional electrophoresis. *ELECTROPHORESIS* 19, 1901–1909. <https://doi.org/10.1002/elps.1150191108>
- Chin, E.T., Papac, D.I., 1999. The Use of a Porous Graphitic Carbon Column for Desalting Hydrophilic Peptides prior to Matrix-Assisted Laser Desorption/ Ionization Time-of-Flight Mass Spectrometry. *Anal. Biochem.* 273, 179–185. <https://doi.org/10.1006/abio.1999.4242>
- Chong, Y., Kleinhammes, A., Tang, P., Xu, Y., Wu, Y., 2015. Dominant Alcohol-Protein Interaction via Hydration-Enabled Enthalpy-Driven Binding Mechanism. *J. Phys. Chem. B* 119, 5367–5375. <https://doi.org/10.1021/acs.jpcc.5b00378>
- Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K.L., Tom, L.M., Brodie, E.L., Hettich, R.L., 2010. Direct Cellular Lysis/Protein Extraction Protocol for Soil Metaproteomics. *J. Proteome Res.* 9, 6615–6622. <https://doi.org/10.1021/pr100787q>

Références bibliographiques

- Cobb, N.J., Sönnichsen, F.D., Mchaourab, H., Surewicz, W.K., 2007. Molecular architecture of human prion protein amyloid: A parallel, in-register β -structure. *Proc. Natl. Acad. Sci. U. S. A.* 104, 18946–18951. <https://doi.org/10.1073/pnas.0706522104>
- Cottrell, J.S., 2011. Protein identification using MS/MS data. *J. Proteomics* 74, 1842–1851. <https://doi.org/10.1016/j.jprot.2011.05.014>
- Crowell, A.M.J., MacLellan, D.L., Doucette, A.A., 2015. A two-stage spin cartridge for integrated protein precipitation, digestion and SDS removal in a comparative bottom-up proteomics workflow. *J. Proteomics, Protein dynamics in health and disease* 118, 140–150. <https://doi.org/10.1016/j.jprot.2014.09.030>
- Crowell, A.M.J., Wall, M.J., Doucette, A.A., 2013. Maximizing recovery of water-soluble proteins through acetone precipitation. *Anal. Chim. Acta* 796, 48–54. <https://doi.org/10.1016/j.aca.2013.08.005>
- Cui, Q., Ni, X., Zeng, L., Tu, Z., Li, J., Sun, K., Chen, X., Li, X., 2017. Optimization of Protein Extraction and Decoloration Conditions for Tea Residues. *Hortic. Plant J.* <https://doi.org/10.1016/j.hpj.2017.06.003>
- David, C., Foley, S., Mavon, C., Enescu, M., 2008. Reductive unfolding of serum albumins uncovered by Raman spectroscopy. *Biopolymers* 89, 623–634. <https://doi.org/10.1002/bip.20972>
- Depledge, M.H., Fossi, M.C., 1994. The role of biomarkers in environmental assessment (2). *Invertebrates. Ecotoxicology* 3, 161–172. <https://doi.org/10.1007/BF00117081>
- Di Girolamo, F., Lante, I., Muraca, M., Putignani, L., 2013. The Role of Mass Spectrometry in the “Omics” Era. *Curr. Org. Chem.* 17, 2891–2905. <https://doi.org/10.2174/1385272817888131118162725>
- Di Palma, S., Hennrich, M.L., Heck, A.J.R., Mohammed, S., 2012. Recent advances in peptide separation by multidimensional liquid chromatography for proteome analysis. *J. Proteomics* 75, 3791–3813. <https://doi.org/10.1016/j.jprot.2012.04.033>
- Dick, W., Tabatabai, M., 1993. Significance and potential uses of soil enzymes.
- Dill, K.A., 1990. Dominant forces in protein folding. *Biochemistry (Mosc.)* 29, 7133–7155. <https://doi.org/10.1021/bi00483a001>
- Dobson, C.M., Šali, A., Karplus, M., 1998. Protein Folding: A Perspective from Theory and Experiment. *Angew. Chem. Int. Ed.* 37, 868–893. [https://doi.org/10.1002/\(SICI\)1521-3773\(19980420\)37:7<868::AID-ANIE868>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1521-3773(19980420)37:7<868::AID-ANIE868>3.0.CO;2-H)
- Dolgikh, D.A., Gilmanshin, R.I., Brazhnikov, E.V., Bychkova, V.E., Semisotnov, G.V., Venyaminov SYu, null, Ptitsyn, O.B., 1981. Alpha-Lactalbumin: compact state with fluctuating tertiary structure? *FEBS Lett.* 136, 311–315.
- Dong, H.-P., Wang, D.-Z., Dai, M., Hong, H.-S., 2010. Characterization of particulate organic matter in the water column of the South China Sea using a shotgun proteomic approach. *Limnol. Oceanogr.* 55, 1565–1578. <https://doi.org/10.4319/lo.2010.55.4.1565>
- dos Remedios, C.G., Chhabra, D., Kekic, M., Dedova, I.V., Tsubakihara, M., Berry, D.A., Nosworthy, N.J., 2003. Actin binding proteins: regulation of cytoskeletal microfilaments. *Physiol. Rev.* 83, 433–473. <https://doi.org/10.1152/physrev.00026.2002>
- Doucette, A.A., Vieira, D.B., Orton, D.J., Wall, M.J., 2014. Resolubilization of Precipitated Intact Membrane Proteins with Cold Formic Acid for Analysis by Mass Spectrometry. *J. Proteome Res.* 13, 6001–6012. <https://doi.org/10.1021/pr500864a>
- Douglass, K.A., Venter, A.R., 2013. Protein analysis by desorption electrospray ionization mass spectrometry and related methods. *J. Mass Spectrom.* 48, 553–560. <https://doi.org/10.1002/jms.3206>
- Dwyer, D.S., Bradley, R.J., 2000. Chemical properties of alcohols and their protein binding sites. *Cell. Mol. Life Sci. CMLS* 57, 265–275. <https://doi.org/10.1007/PL00000689>

Références bibliographiques

- Egertson, J.D., MacLean, B., Johnson, R., Xuan, Y., MacCoss, M.J., 2015. Multiplexed peptide analysis using data-independent acquisition and Skyline. *Nat. Protoc.* 10, 887–903. <https://doi.org/10.1038/nprot.2015.055>
- Eisenberg, D., Jucker, M., 2012. The amyloid state of proteins in human diseases. *Cell* 148, 1188–1203. <https://doi.org/10.1016/j.cell.2012.02.022>
- Elkordy, A.A., Forbes, R.T., Barry, B.W., 2008. Study of protein conformational stability and integrity using calorimetry and FT-Raman spectroscopy correlated with enzymatic activity. *Eur. J. Pharm. Sci. Off. J. Eur. Fed. Pharm. Sci.* 33, 177–190. <https://doi.org/10.1016/j.ejps.2007.11.002>
- Faure, G., Bornot, A., de Brevern, A.G., 2008. Protein contacts, inter-residue interactions and side-chain modelling. *Recent Adv. Complete Genome Anal.* 90, 626–639. <https://doi.org/10.1016/j.biochi.2007.11.007>
- Feist, P., Hummon, A.B., 2015. Proteomic Challenges: Sample Preparation Techniques for Microgram-Quantity Protein Analysis from Biological Samples. *Int. J. Mol. Sci.* 16, 3537–3563.
- Fennema, O.R., 1996. *Food Chemistry*, Third Edition. CRC Press.
- Ferguson, A.D., Ködding, J., Walker, G., Bös, C., Coulton, J.W., Diederichs, K., Braun, V., Welte, W., 2001. Active transport of an antibiotic rifamycin derivative by the outer-membrane protein FhuA. *Struct. Lond. Engl.* 1993 9, 707–716.
- Fernández-Cisnal, R., García-Sevillano, M.A., Gómez-Ariza, J.L., Pueyo, C., López-Barea, J., Abril, N., 2017. 2D-DIGE as a proteomic biomarker discovery tool in environmental studies with *Procambarus clarkii*. *Sci. Total Environ.* 584–585, 813–827. <https://doi.org/10.1016/j.scitotenv.2017.01.125>
- Fernandez-Gomez, F.-J., Jumeau, F., Derisbourg, M., Burnouf, S., Tran, H., Eddarkaoui, S., Obriot, H., Dutoit-Lefevre, V., Deramecourt, V., Mitchell, V., Lefranc, D., Hamdane, M., Blum, D., Buée, L., Buée-Scherrer, V., Sergeant, N., 2014. Consensus Brain-derived Protein, Extraction Protocol for the Study of Human and Murine Brain Proteome Using Both 2D-DIGE and Mini 2DE Immunoblotting. *J. Vis. Exp. JoVE.* <https://doi.org/10.3791/51339>
- Fic, E., Kedracka-Krok, S., Jankowska, U., Pirog, A., Dziejzicka-Wasylewska, M., 2010. Comparison of protein precipitation methods for various rat brain structures prior to proteomic analysis. *ELECTROPHORESIS* 31, 3573–3579. <https://doi.org/10.1002/elps.201000197>
- Fijneman, R.J.A., de Wit, M., Pourghasian, M., Piersma, S.R., Pham, T.V., Warmoes, M.O., Lavaei, M., Piso, C., Smit, F., Delis-van Diemen, P.M., van Turenhout, S.T., Terhaar sive Droste, J.S., Mulder, C.J.J., Blankenstein, M.A., Robanus-Maandag, E.C., Smits, R., Fodde, R., van Hinsbergh, V.W.M., Meijer, G.A., Jimenez, C.R., 2012. Proximal fluid proteome profiling of mouse colon tumors reveals biomarkers for early diagnosis of human colorectal cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 18, 2613–2624. <https://doi.org/10.1158/1078-0432.CCR-11-1937>
- Fiorito, T.M., Icoz, I., Stotzky, G., 2008. Adsorption and binding of the transgenic plant proteins, human serum albumin, β -glucuronidase, and Cry3Bb1, on montmorillonite and kaolinite: Microbial utilization and enzymatic activity of free and clay-bound proteins. *Appl. Clay Sci.* 39, 142–150. <https://doi.org/10.1016/j.clay.2007.07.006>
- Fortin, T., Salvador, A., Charrier, J.P., Lenz, C., Lacoux, X., Morla, A., Choquet-Kastylevsky, G., Lemoine, J., 2009. Clinical Quantitation of Prostate-specific Antigen Biomarker in the Low Nanogram/Milliliter Range by Conventional Bore Liquid Chromatography-Tandem Mass Spectrometry (Multiple Reaction Monitoring) Coupling and Correlation with ELISA Tests. *Mol. Cell. Proteomics MCP* 8, 1006–1015. <https://doi.org/10.1074/mcp.M800238-MCP200>
- Frahm, J.L., Howard, B.E., Heber, S., Muddiman, D.C., 2006. Accessible proteomics space and its implications for peak capacity for zero-, one- and two-dimensional separations

Références bibliographiques

- coupled with FT-ICR and TOF mass spectrometry. *J. Mass Spectrom.* JMS 41, 281–288. <https://doi.org/10.1002/jms.1024>
- Frantzi, M., Bhat, A., Latosinska, A., 2014. Clinical proteomic biomarkers: relevant issues on study design & technical considerations in biomarker development. *Clin. Transl. Med.* 3, 7. <https://doi.org/10.1186/2001-1326-3-7>
- Friedman, M., 1999. Chemistry, Nutrition, and Microbiology of d-Amino Acids. *J. Agric. Food Chem.* 47, 3457–3479. <https://doi.org/10.1021/jf990080u>
- Fu, P.-Q., Liu, C.-Q., Wu, F.-C., 2005. [Three-dimensional excitation emission matrix fluorescence spectroscopic characterization of dissolved organic matter]. *Guang Pu Xue Yu Guang Pu Fen Xi Guang Pu* 25, 2024–2028.
- Fuchs, S.A., Berger, R., Klomp, L.W.J., de Koning, T.J., 2005. d-Amino acids in the central nervous system in health and disease. *Mol. Genet. Metab.* 85, 168–180. <https://doi.org/10.1016/j.ymgme.2005.03.003>
- Fujii, Noriko, Kaji, Y., Fujii, Norihiko, 2011. d-Amino acids in aged proteins: Analysis and biological relevance. *Anal. Biol. Relev. -Amino Acids Relat. Compd.* 879, 3141–3147. <https://doi.org/10.1016/j.jchromb.2011.05.051>
- Gallien, S., Duriez, E., Demeure, K., Domon, B., 2013. Selectivity of LC-MS/MS analysis: Implication for proteomics experiments. *J. Proteomics, Special Issue: From protein structures to clinical applications* 81, 148–158. <https://doi.org/10.1016/j.jprot.2012.11.005>
- Gama, M.R., Collins, C.H., Bottoli, C.B.G., 2013. Nano-Liquid Chromatography in Pharmaceutical and Biomedical Research. *J. Chromatogr. Sci.* 51, 694–703. <https://doi.org/10.1093/chromsci/bmt023>
- García-Otero, N., Barciela-Alonso, M.C., Moreda-Piñeiro, A., Bermejo-Barrera, P., 2013. Study of extraction procedures for protein analysis in plankton samples by OFFGEL electrophoresis hyphenated with Lab-on-a-chip technology. *Talanta* 115, 631–641.
- Georgia, B.-M., Xavier, R., 2014. *Culture de cellules animales (3^e éd.)*. Lavoisier.
- Ghisaidoobe, A.B.T., Chung, S.J., 2014. Intrinsic Tryptophan Fluorescence in the Detection and Analysis of Proteins: A Focus on Förster Resonance Energy Transfer Techniques. *Int. J. Mol. Sci.* 15, 22518–22538. <https://doi.org/10.3390/ijms151222518>
- Ghuman, J., Zunszain, P.A., Petitpas, I., Bhattacharya, A.A., Otagiri, M., Curry, S., 2005. Structural basis of the drug-binding specificity of human serum albumin. *J. Mol. Biol.* 353, 38–52. <https://doi.org/10.1016/j.jmb.2005.07.075>
- Gillet, L.C., Leitner, A., Aebersold, R., 2016. Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. *Annu. Rev. Anal. Chem. Palo Alto Calif* 9, 449–472. <https://doi.org/10.1146/annurev-anchem-071015-041535>
- Gobat, J.-M., Aragno, M., Matthey, W., 2010. *Le sol vivant: bases de pédologie, biologie des sols*. PPUR Presses polytechniques.
- Godfrey, M.A.J., 1998. Immunoaffinity extraction in veterinary residue analysis—a regulatory viewpoint†‡. *Analyst* 123, 2501–2506. <https://doi.org/10.1039/A805162C>
- Goedert, M., 2015. NEURODEGENERATION. Alzheimer's and Parkinson's diseases: The prion concept in relation to assembled A β , tau, and α -synuclein. *Science* 349, 1255555. <https://doi.org/10.1126/science.1255555>
- Goers, J., Permyakov, S.E., Permyakov, E.A., Uversky, V.N., Fink, A.L., 2002. Conformational prerequisites for alpha-lactalbumin fibrillation. *Biochemistry (Mosc.)* 41, 12546–12551.
- Golizeh, M., Schneider, C., Ohlund, L.B., Sleno, L., 2015. Multidimensional LC–MS/MS analysis of liver proteins in rat, mouse and human microsomal and S9 fractions. *EuPA Open Proteomics* 6, 16–27. <https://doi.org/10.1016/j.euprot.2015.01.003>
- Gong, Y., Wang, N., Wu, F., Cass, C.E., Damaraju, S., Mackey, J.R., Li, L., 2008. Proteome profile of human breast cancer tissue generated by LC-ESI-MS/MS combined with

Références bibliographiques

- sequential protein precipitation and solubilization. *J. Proteome Res.* 7, 3583–3590. <https://doi.org/10.1021/pr800229j>
- González-Antuña, A., Domínguez-Romero, J.C., García-Reyes, J.F., Rodríguez-González, P., Centineo, G., García Alonso, J.I., Molina-Díaz, A., 2013. Overcoming matrix effects in electrospray: Quantitation of β -agonists in complex matrices by isotope dilution liquid chromatography–mass spectrometry using singly ^{13}C -labeled analogues. *J. Chromatogr. A* 1288, 40–47. <https://doi.org/10.1016/j.chroma.2013.02.074>
- Granger, J., Siddiqui, J., Copeland, S., Remick, D., 2005. Albumin depletion of human plasma also removes low abundance proteins including the cytokines. *PROTEOMICS* 5, 4713–4718. <https://doi.org/10.1002/pmic.200401331>
- Greenfield, L.G., 1972. The nature of the organic nitrogen of soils. *Plant Soil* 36, 191–198. <https://doi.org/10.1007/BF01373469>
- Gross, M., 2000. Proteins that convert from alpha helix to beta sheet: implications for folding and disease. *Curr. Protein Pept. Sci.* 1, 339–347.
- Gruber, K.A., Stein, S., Brink, L., Radhakrishnan, A., Udenfriend, S., 1976. Fluorometric assay of vasopressin and oxytocin: a general approach to the assay of peptides in tissues. *Proc. Natl. Acad. Sci. U. S. A.* 73, 1314–1318.
- Gumber, S., Taylor, D.L., Whittington, R.J., 2007. Protein extraction from *Mycobacterium avium* subsp. *paratuberculosis*: Comparison of methods for analysis by sodium dodecyl sulphate polyacrylamide gel electrophoresis, native PAGE and surface enhanced laser desorption/ionization time of flight mass spectrometry. *J. Microbiol. Methods* 68, 115–127. <https://doi.org/10.1016/j.mimet.2006.07.003>
- Gumustas, M., Kurbanoglu, S., Uslu, B., Ozkan, S.A., 2013. UPLC versus HPLC on Drug Analysis: Advantageous, Applications and Their Validation Parameters. *Chromatographia* 76, 1365–1427. <https://doi.org/10.1007/s10337-013-2477-8>
- Hagen, J. von, 2011. *Proteomics Sample Preparation*. John Wiley & Sons.
- Hahne, H., Pachl, F., Ruprecht, B., Maier, S.K., Klaeger, S., Helm, D., Médard, G., Wilm, M., Lemeer, S., Kuster, B., 2013. DMSO enhances electrospray response, boosting sensitivity of proteomic experiments. *Nat. Methods* 10, 989–991. <https://doi.org/10.1038/nmeth.2610>
- Hall, Q., Cannon, M.C., 2002. The Cell Wall Hydroxyproline-Rich Glycoprotein RSH Is Essential for Normal Embryo Development in *Arabidopsis*. *Plant Cell* 14, 1161–1172. <https://doi.org/10.1105/tpc.010477>
- Hao, P., Ren, Y., Datta, A., Tam, J.P., Sze, S.K., 2015. Evaluation of the Effect of Trypsin Digestion Buffers on Artificial Deamidation. *J. Proteome Res.* 14, 1308–1314. <https://doi.org/10.1021/pr500903b>
- Hao, R., Adoligbe, C., Jiang, B., Zhao, X., Gui, L., Qu, K., Wu, S., Zan, L., 2015. An Optimized Trichloroacetic Acid/Acetone Precipitation Method for Two-Dimensional Gel Electrophoresis Analysis of Qinchuan Cattle *Longissimus Dorsi* Muscle Containing High Proportion of Marbling. *PloS One* 10, e0124723. <https://doi.org/10.1371/journal.pone.0124723>
- Haslene-Hox, H., Oveland, E., Berg, K.C., Kolmannskog, O., Woie, K., Salvesen, H.B., Tenstad, O., Wiig, H., 2011. A New Method for Isolation of Interstitial Fluid from Human Solid Tumors Applied to Proteomic Analysis of Ovarian Carcinoma Tissue. *PLoS ONE* 6. <https://doi.org/10.1371/journal.pone.0019217>
- Hédoux, A., Krenzlin, S., Paccou, L., Guinet, Y., Flament, M.-P., Siepmann, J., 2010. Influence of urea and guanidine hydrochloride on lysozyme stability and thermal denaturation; a correlation between activity, protein dynamics and conformational changes. *Phys. Chem. Chem. Phys.* PCCP 12, 13189–13196. <https://doi.org/10.1039/c0cp00602e>
- Hedoux, A., Paccou, L., Achir, S., Guinet, Y., 2013. Mechanism of protein stabilization by trehalose during freeze-drying analyzed by in situ micro-Raman spectroscopy. *J. Pharm. Sci.* 102, 2484–2494. <https://doi.org/10.1002/jps.23638>

Références bibliographiques

- Hédoux, A., Willart, J.-F., Paccou, L., Guinet, Y., Affouard, F., Lerbret, A., Descamps, M., 2009. Thermostabilization mechanism of bovine serum albumin by trehalose. *J. Phys. Chem. B* 113, 6119–6126. <https://doi.org/10.1021/jp900330r>
- Hempstead, P.D., Yewdall, S.J., Fernie, A.R., Lawson, D.M., Artymiuk, P.J., Rice, D.W., Ford, G.C., Harrison, P.M., 1997. Comparison of the three-dimensional structures of recombinant human H and horse L ferritins at high resolution. Edited by R. Huber. *J. Mol. Biol.* 268, 424–448. <https://doi.org/10.1006/jmbi.1997.0970>
- Hennion, M.-C., Pichon, V., 2003. Immuno-based sample preparation for trace analysis. *J. Chromatogr. A, A Century of Chromatography 1903-2003* 1000, 29–52. [https://doi.org/10.1016/S0021-9673\(03\)00529-6](https://doi.org/10.1016/S0021-9673(03)00529-6)
- Henry, R.J., Sobel, C., Segalove, M., 1956. Turbidimetric Determination of Proteins with Sulfosalicylic and Trichloroacetic Acids. *Proc. Soc. Exp. Biol. Med.* 92, 748–751. <https://doi.org/10.3181/00379727-92-22601>
- Herbert, B., 1999. Advances in protein solubilisation for two-dimensional electrophoresis. *Electrophoresis* 20, 660–663. [https://doi.org/10.1002/\(SICI\)1522-2683\(19990101\)20:4/5<660::AID-ELPS660>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1522-2683(19990101)20:4/5<660::AID-ELPS660>3.0.CO;2-Q)
- Hildonen, S., Halvorsen, T.G., Reubsæet, L., 2014. Why less is more when generating tryptic peptides in bottom-up proteomics. *Proteomics* 14, 2031–2041. <https://doi.org/10.1002/pmic.201300479>
- Hirsch, J., Hansen, K.C., Burlingame, A.L., Matthay, M.A., 2004. Proteomics: current techniques and potential applications to lung disease. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* 287, L1–L23. <https://doi.org/10.1152/ajplung.00301.2003>
- Hochstrasser, D.F., Sanchez, J.-C., Appel, R.D., 2002. Proteomics and its trends facing nature's complexity. *Proteomics* 2, 807–812. [https://doi.org/10.1002/1615-9861\(200207\)2:7<807::AID-PROT807>3.0.CO;2-4](https://doi.org/10.1002/1615-9861(200207)2:7<807::AID-PROT807>3.0.CO;2-4)
- Hol, W.G., 1985. The role of the alpha-helix dipole in protein function and structure. *Prog. Biophys. Mol. Biol.* 45, 149–195.
- Holčapek, M., Jirásko, R., Lisa, M., 2012. Recent developments in liquid chromatography–mass spectrometry and related techniques. *J. Chromatogr. A, Mass Spectrometry: Innovation and Application. Part VII* 1259, 3–15. <https://doi.org/10.1016/j.chroma.2012.08.072>
- Hoofnagle, A.N., Becker, J.O., Wener, M.H., Heinecke, J.W., 2008. Quantification of Thyroglobulin, a Low-abundance Serum Protein, by Immunoaffinity Peptide Enrichment and Tandem Mass Spectrometry. *Clin. Chem.* 54, 1796–1804. <https://doi.org/10.1373/clinchem.2008.109652>
- Horie, K., Sato, Y., Kimura, T., Nakamura, T., Ishihama, Y., Oda, Y., Ikegami, T., Tanaka, N., 2012. Estimation and optimization of the peak capacity of one-dimensional gradient high performance liquid chromatography using a long monolithic silica capillary column. *J. Chromatogr. A, High-Performance Columns and their Operations: The Speed, Efficiency and Peak Capacity Obtainable with High-Performance Columns under Various Operation Conditions* 1228, 283–291. <https://doi.org/10.1016/j.chroma.2011.12.088>
- Hu, Q., Noll, R.J., Li, H., Makarov, A., Hardman, M., Graham Cooks, R., 2005. The Orbitrap: a new mass spectrometer. *J. Mass Spectrom.* 40, 430–443. <https://doi.org/10.1002/jms.856>
- Huang, H.-J., Chen, W.-Y., Wu, J.-H., 2014. Total Protein Extraction for Metaproteomics Analysis of Methane Producing Biofilm: The Effects of Detergents. *Int. J. Mol. Sci.* 15, 10169–10184. <https://doi.org/10.3390/ijms150610169>
- Hultman, J., Waldrop, M.P., Mackelprang, R., David, M.M., McFarland, J., Blazewicz, S.J., Harden, J., Turetsky, M.R., McGuire, A.D., Shah, M.B., VerBerkmoes, N.C., Lee, L.H., Mavrommatis, K., Jansson, J.K., 2015. Multi-omics of permafrost, active layer and

Références bibliographiques

- thermocarst bog soil microbiomes. *Nature* 521, 208–212. <https://doi.org/10.1038/nature14238>
- Hustoft, H.K., Malerod, H., Wilson, S.R., Reubsæet, L., Lundanes, E., Greibrokk, T., 2012. A Critical Review of Trypsin Digestion for LC-MS Based Proteomics. <https://doi.org/10.5772/29326>
- Hüttenhain, R., Malmström, J., Picotti, P., Aebersold, R., 2009. Perspectives of targeted mass spectrometry for protein biomarker verification. *Curr. Opin. Chem. Biol.* 13, 518–525. <https://doi.org/10.1016/j.cbpa.2009.09.014>
- Ifa, D.R., Wu, C., Ouyang, Z., Cooks, R.G., 2010. Desorption electrospray ionization and other ambient ionization methods: current progress and preview. *Analyst* 135, 669–681. <https://doi.org/10.1039/B925257F>
- Isaacson, T., Damasceno, C.M.B., Saravanan, R.S., He, Y., Catala, C., Saladie, M., Rose, J.K.C., 2006. Sample extraction techniques for enhanced proteomic analysis of plant tissues. *Nat Protoc.* 1, 769–774. <https://doi.org/10.1038/nprot.2006.102>
- Jacobs, D.I., van Rijssen, M.S., van der Heijden, R., Verpoorte, R., 2001. Sequential solubilization of proteins precipitated with trichloroacetic acid in acetone from cultured *Catharanthus roseus* cells yields 52% more spots after two-dimensional electrophoresis. *Proteomics* 1, 1345–1350. [https://doi.org/10.1002/1615-9861\(200111\)1:11<1345::AID-PROT1345>3.0.CO;2-F](https://doi.org/10.1002/1615-9861(200111)1:11<1345::AID-PROT1345>3.0.CO;2-F)
- James, P., 1997. Protein identification in the post-genome era: the rapid rise of proteomics [WWW Document]. *Q. Rev. Biophys.* URL [/core/journals/quarterly-reviews-of-biophysics/article/protein-identification-in-the-postgenome-era-the-rapid-rise-of-proteomics/F72E4F1C6FF1D481E88752465C14313E](http://core/journals/quarterly-reviews-of-biophysics/article/protein-identification-in-the-postgenome-era-the-rapid-rise-of-proteomics/F72E4F1C6FF1D481E88752465C14313E) (accessed 5.26.17).
- Jan, M.T., Roberts, P., Tonheim, S.K., Jones, D.L., 2009. Protein breakdown represents a major bottleneck in nitrogen cycling in grassland soils. *Soil Biol. Biochem.* 41, 2272–2282. <https://doi.org/10.1016/j.soilbio.2009.08.013>
- Jez, J.M., 2017. Revisiting protein structure, function, and evolution in the genomic era. *Struct. New Insecticidal Proteins Regul. Chall. Commer.* 142, 11–15. <https://doi.org/10.1016/j.jip.2016.07.013>
- Jiang, L., He, L., Fountoulakis, M., 2004. Comparison of protein precipitation methods for sample preparation prior to proteomic analysis. *J. Chromatogr. A* 1023, 317–320. <https://doi.org/10.1016/j.chroma.2003.10.029>
- Jiang, S., Zhang, Z., Li, L., 2015. A One-Step Preparation Method of Monolithic Enzyme Reactor for Highly Efficient Sample Preparation Coupled to Mass Spectrometry-Based Proteomics Studies. *J. Chromatogr. A* 1412, 75–81. <https://doi.org/10.1016/j.chroma.2015.07.121>
- Johnson, D., Boyes, B., Fields, T., Kopkin, R., Orlando, R., 2013a. Optimization of Data-Dependent Acquisition Parameters for Coupling High-Speed Separations with LC-MS/MS for Protein Identifications. *J. Biomol. Tech. JBT* 24, 62–72. <https://doi.org/10.7171/jbt.13-2402-003>
- Johnson, D., Boyes, B., Orlando, R., 2013b. The Use of Ammonium Formate as a Mobile-Phase Modifier for LC-MS/MS Analysis of Tryptic Digests. *J. Biomol. Tech. JBT* 24, 187–197. <https://doi.org/10.7171/jbt.13-2404-005>
- Johnston, C.T., Premachandra, G.S., Szabo, T., Lok, J., Schoonheydt, R.A., 2012. Interaction of Biological Molecules with Clay Minerals: A Combined Spectroscopic and Sorption Study of Lysozyme on Saponite. *Langmuir* 28, 611–619. <https://doi.org/10.1021/la203161n>
- Jordan, T., Eads, J.C., Spiro, T.G., 1995. Secondary and tertiary structure of the A-state of cytochrome c from resonance Raman spectroscopy. *Protein Sci. Publ. Protein Soc.* 4, 716–728.

Références bibliographiques

- Kadam, S.U., Tiwari, B.K., Álvarez, C., O'Donnell, C.P., 2015. Ultrasound applications for the extraction, identification and delivery of food proteins and bioactive peptides. *Trends Food Sci. Technol.* 46, 60–67. <https://doi.org/10.1016/j.tifs.2015.07.012>
- Kanerva, S., Smolander, A., Kitunen, V., Ketola, R.A., Kotiaho, T., 2013. Comparison of extractants and applicability of MALDI–TOF–MS in the analysis of soil proteinaceous material from different types of soil. *Org. Geochem.* 56, 1–9. <https://doi.org/10.1016/j.orggeochem.2012.12.003>
- Karas, M., Hillenkamp, F., 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* 60, 2299–2301. <https://doi.org/10.1021/ac00171a028>
- Karpievitch, Y.V., Polpitiya, A.D., Anderson, G.A., Smith, R.D., Dabney, A.R., 2010. Liquid Chromatography Mass Spectrometry–Based Proteomics: Biological and Technological Aspects. *Ann. Appl. Stat.* 4, 1797–1823. <https://doi.org/10.1214/10-AOAS341>
- Kauzmann, W., 1959. Some Factors in the Interpretation of Protein Denaturation. *Adv. Protein Chem.* 14, 1–63. [https://doi.org/10.1016/S0065-3233\(08\)60608-7](https://doi.org/10.1016/S0065-3233(08)60608-7)
- Keiblinger, K.M., Fuchs, S., Zechmeister-Boltenstern, S., Riedel, K., 2016. Soil and leaf litter metaproteomics—a brief guideline from sampling to understanding. *FEMS Microbiol. Ecol.* 92, fiw180. <https://doi.org/10.1093/femsec/fiw180>
- Keiblinger, K.M., Wilhartitz, I.C., Schneider, T., Roschitzki, B., Schmid, E., Eberl, L., Riedel, K., Zechmeister-Boltenstern, S., 2012. Soil metaproteomics – Comparative evaluation of protein extraction protocols. *Soil Biol. Biochem.* 54, 14–24. <https://doi.org/10.1016/j.soilbio.2012.05.014>
- Kelly, S.T., Senyo Opong, W., Zydney, A.L., 1993. The influence of protein aggregates on the fouling of microfiltration membranes during stirred cell filtration. *J. Membr. Sci.* 80, 175–187. [https://doi.org/10.1016/0376-7388\(93\)85142-J](https://doi.org/10.1016/0376-7388(93)85142-J)
- Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H., Phillips, D.C., 1958. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* 181, 662–666.
- Kentish, S., Ashokkumar, M., 2011. The Physical and Chemical Effects of Ultrasound, in: *Ultrasound Technologies for Food and Bioprocessing*, Food Engineering Series. Springer, New York, NY, pp. 1–12. https://doi.org/10.1007/978-1-4419-7472-3_1
- Keshishian, H., Addona, T., Burgess, M., Kuhn, E., Carr, S.A., 2007. Quantitative, Multiplexed Assays for Low Abundance Proteins in Plasma by Targeted Mass Spectrometry and Stable Isotope Dilution. *Mol. Cell. Proteomics MCP* 6, 2212–2229. <https://doi.org/10.1074/mcp.M700354-MCP200>
- Khabiri, M., Minofar, B., Brezovský, J., Damborský, J., Ettrich, R., 2013. Interaction of organic solvents with protein structures at protein–solvent interface. *J. Mol. Model.* 19, 4701–4711. <https://doi.org/10.1007/s00894-012-1507-z>
- Kim, H.-J., Lin, D., Lee, H.-J., Li, M., Liebler, D.C., 2016. Quantitative Profiling of Protein Tyrosine Kinases in Human Cancer Cell Lines by Multiplexed Parallel Reaction Monitoring Assays. *Mol. Cell. Proteomics MCP* 15, 682–691. <https://doi.org/10.1074/mcp.O115.056713>
- Kim, M.-S., Pinto, S.M., Getnet, D., Nirujogi, R.S., Manda, S.S., Chaerkady, R., Madugundu, A.K., Kelkar, D.S., Isserlin, R., Jain, S., Thomas, J.K., Muthusamy, B., Leal-Rojas, P., Kumar, P., Sahasrabudhe, N.A., Balakrishnan, L., Advani, J., George, B., Renuse, S., Selvan, L.D.N., Patil, A.H., Nanjappa, V., Radhakrishnan, A., Prasad, S., Subbannayya, T., Raju, R., Kumar, M., Sreenivasamurthy, S.K., Marimuthu, A., Sathe, G.J., Chavan, S., Datta, K.K., Subbannayya, Y., Sahu, A., Yelamanchi, S.D., Jayaram, S., Rajagopalan, P., Sharma, J., Murthy, K.R., Syed, N., Goel, R., Khan, A.A., Ahmad, S., Dey, G., Mudgal, K., Chatterjee, A., Huang, T.-C., Zhong, J., Wu, X., Shaw, P.G.,

Références bibliographiques

- Freed, D., Zahari, M.S., Mukherjee, K.K., Shankar, S., Mahadevan, A., Lam, H., Mitchell, C.J., Shankar, S.K., Satishchandra, P., Schroeder, J.T., Sirdeshmukh, R., Maitra, A., Leach, S.D., Drake, C.G., Halushka, M.K., Prasad, T.S.K., Hruban, R.H., Kerr, C.L., Bader, G.D., Iacobuzio-Donahue, C.A., Gowda, H., Pandey, A., 2014. A draft map of the human proteome. *Nature* 509, 575–581. <https://doi.org/10.1038/nature13302>
- Kiss, S., Drăgan-Bularda, M., Rădulescu, D., 1975. Biological Significance of Enzymes Accumulated in Soil. *Adv. Agron.* 27, 25–87. [https://doi.org/10.1016/S0065-2113\(08\)70007-5](https://doi.org/10.1016/S0065-2113(08)70007-5)
- Kito, K., Ito, T., 2008. Mass Spectrometry-Based Approaches Toward Absolute Quantitative Proteomics. *Curr. Genomics* 9, 263–274. <https://doi.org/10.2174/138920208784533647>
- Kleiner, M., Wentrup, C., Lott, C., Teeling, H., Wetzel, S., Young, J., Chang, Y.-J., Shah, M., VerBerkmoes, N.C., Zarzycki, J., Fuchs, G., Markert, S., Hempel, K., Voigt, B., Becher, D., Liebeke, M., Lalk, M., Albrecht, D., Hecker, M., Schweder, T., Dubilier, N., 2012. Metaproteomics of a gutless marine worm and its symbiotic microbial community reveal unusual pathways for carbon and energy use. *Proc. Natl. Acad. Sci.* 109, E1173–E1182. <https://doi.org/10.1073/pnas.1121198109>
- Kneussel, M., Wagner, W., 2013. Myosin motors at neuronal synapses: drivers of membrane transport and actin dynamics. *Nat. Rev. Neurosci.* 14, 233–247. <https://doi.org/10.1038/nrn3445>
- Koppen, V., Jones, R., Bockx, M., Cuyckens, F., 2014. High volume injections of biological samples for sensitive metabolite profiling and quantitation. *J. Chromatogr. A* 1372, 102–109. <https://doi.org/10.1016/j.chroma.2014.10.087>
- KORNER, A., DEBRO, J.R., 1956. Solubility of Albumin in Alcohol after Precipitation by Trichloroacetic Acid: a Simplified Procedure for Separation of Albumin. *Nature* 178, 1067–1067. <https://doi.org/10.1038/1781067a0>
- Kuang, H., Wang, W., Xu, L., Ma, W., Liu, L., Wang, L., Xu, C., 2013. Monoclonal Antibody-Based Sandwich ELISA for the Detection of Staphylococcal Enterotoxin A. *Int. J. Environ. Res. Public Health* 10, 1598–1608. <https://doi.org/10.3390/ijerph10041598>
- Kuhn, E., Wu, J., Karl, J., Liao, H., Zolg, W., Guild, B., 2004. Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and ¹³C-labeled peptide standards. *PROTEOMICS* 4, 1175–1186. <https://doi.org/10.1002/pmic.200300670>
- Kumar, T.K., Subbiah, V., Ramakrishna, T., Pandit, M.W., 1994. Trichloroacetic acid-induced unfolding of bovine pancreatic ribonuclease. Existence of molten globule-like state. *J. Biol. Chem.* 269, 12620–12625.
- Kurouski, D., Van Duyne, R.P., Lednev, I.K., 2015. Exploring the structure and formation mechanism of amyloid fibrils by Raman spectroscopy: a review. *Analyst* 140, 4967–4980.
- Lacerda, C.M.R., Reardon, K.F., 2009. Environmental proteomics: applications of proteome profiling in environmental microbiology and biotechnology. *Brief. Funct. Genomic. Proteomic.* 8, 75–87. <https://doi.org/10.1093/bfpg/elp005>
- Lacorte, S., Fernandez-Alba, A.R., 2006. Time of flight mass spectrometry applied to the liquid chromatographic analysis of pesticides in water and food. *Mass Spectrom. Rev.* 25, 866–880. <https://doi.org/10.1002/mas.20094>
- Lakowicz, J.R., 2006. Fluorophores, in: *Principles of Fluorescence Spectroscopy*. Springer, Boston, MA, pp. 63–95. https://doi.org/10.1007/978-0-387-46312-4_3
- Lange, V., Picotti, P., Domon, B., Aebersold, R., 2008. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* 4, 222. <https://doi.org/10.1038/msb.2008.61>
- Larsen, M.R., Graham, M.E., Robinson, P.J., Roepstorff, P., 2004. Improved Detection of Hydrophilic Phosphopeptides Using Graphite Powder Microcolumns and Mass

Références bibliographiques

- Spectrometry Evidence for in Vivo Doubly Phosphorylated Dynamin I and Dynamin III. *Mol. Cell. Proteomics* 3, 456–465. <https://doi.org/10.1074/mcp.M300105-MCP200>
- Law, H.C.H., Kong, R.P.W., Szeto, S.S.W., Zhao, Y., Zhang, Z., Wang, Y., Li, G., Quan, Q., Lee, S.M.Y., Lam, H.C., Chu, I.K., 2015. A versatile reversed phase-strong cation exchange-reversed phase (RP–SCX–RP) multidimensional liquid chromatography platform for qualitative and quantitative shotgun proteomics. *Analyst* 140, 1237–1252. <https://doi.org/10.1039/C4AN01893A>
- Leary, D.H., Hervey, W.J., Li, R.W., Deschamps, J.R., Kusterbeck, A.W., Vora, G.J., 2012. Method development for metaproteomic analyses of marine biofilms. *Anal. Chem.* 84, 4006–4013. <https://doi.org/10.1021/ac203315n>
- Lebert, D., Louwagie, M., Goetze, S., Picard, G., Ossola, R., Duquesne, C., Basler, K., Ferro, M., Rinner, O., Aebersold, R., Garin, J., Mouz, N., Brunner, E., Brun, V., 2015. DIGESTIF: A Universal Quality Standard for the Control of Bottom-Up Proteomics Experiments. *J. Proteome Res.* 14, 787–803. <https://doi.org/10.1021/pr500834z>
- Lesser, G.J., Rose, G.D., 1990. Hydrophobicity of amino acid subgroups in proteins. *Proteins Struct. Funct. Bioinforma.* 8, 6–13. <https://doi.org/10.1002/prot.340080104>
- Lewandrowski, U., Sickmann, A., 2010. Online Dual Gradient Reversed-Phase/Porous Graphitized Carbon nanoHPLC for Proteomic Applications. *Anal. Chem.* 82, 5391–5396. <https://doi.org/10.1021/ac100853w>
- Li, J., Zhou, L., Wang, H., Yan, H., Li, N., Zhai, R., Jiao, F., Hao, F., Jin, Z., Tian, F., Peng, B., Zhang, Y., Qian, X., 2015. A new sample preparation method for the absolute quantitation of a target proteome using ¹⁸O labeling combined with multiple reaction monitoring mass spectrometry. *Analyst* 140, 1281–1290. <https://doi.org/10.1039/C4AN02092H>
- Lim, L.W., Okouchi, Y., Takeuchi, T., 2007. On-line preconcentration of trace carcinogenic polycyclic aromatic hydrocarbons (PAHs) in microcolumn liquid chromatography via large volume injection. *Talanta, Special Issue on China-Japan-Korea Environmental Analysis* 72, 1600–1608. <https://doi.org/10.1016/j.talanta.2007.02.033>
- Lim, W.K., Rösgen, J., Englander, S.W., 2009. Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group. *Proc. Natl. Acad. Sci. U. S. A.* 106, 2595–2600. <https://doi.org/10.1073/pnas.0812588106>
- Lin, Y., Liu, H., Liu, Z., Liu, Y., He, Q., Chen, P., Wang, X., Liang, S., 2013. Development and evaluation of an entirely solution-based combinative sample preparation method for membrane proteomics. *Anal. Biochem.* 432, 41–48. <https://doi.org/10.1016/j.ab.2012.09.023>
- Liu, J., Wan, Y., Zhao, Z., Chen, H., 2013. Determination of the content of rosmarinic acid by HPLC and analytical comparison of volatile constituents by GC-MS in different parts of *Perilla frutescens* (L.) Britt. *Chem. Cent. J.* 7, 61. <https://doi.org/10.1186/1752-153X-7-61>
- Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., Darnell, J., 2000. Hierarchical Structure of Proteins.
- Lowik, D.W.P.M., Leunissen, E.H.P., van den Heuvel, M., Hansen, M.B., van Hest, J.C.M., 2010. Stimulus responsive peptide based materials. *Chem. Soc. Rev.* 39, 3394–3412. <https://doi.org/10.1039/B914342B>
- Lund, H., Løvstetten, K., Paus, E., Halvorsen, T.G., Reubsæet, L., 2012. Immuno-MS based targeted proteomics: highly specific, sensitive, and reproducible human chorionic gonadotropin determination for clinical diagnostics and doping analysis. *Anal. Chem.* 84, 7926–7932. <https://doi.org/10.1021/ac301418f>
- M. Lesk, A., 2000. Introduction to Protein Architecture-The Structural Biology of Proteins.
- Ma, X., Benson, C.H., McKenzie, D., Aiken, J.M., Pedersen, J.A., 2007. Adsorption of Pathogenic Prion Protein to Quartz Sand. *Environ. Sci. Technol.* 41, 2324–2330. <https://doi.org/10.1021/es062122i>

Références bibliographiques

- 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., 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinforma. Oxf. Engl.* 26, 966–968. <https://doi.org/10.1093/bioinformatics/btq054>
- Magdeldin, S., Enany, S., Yoshida, Y., Xu, B., Zhang, Y., Zureena, Z., Lokamani, I., Yaoita, E., Yamamoto, T., 2014a. Basics and recent advances of two dimensional-polyacrylamide gel electrophoresis. *Clin. Proteomics* 11, 16. <https://doi.org/10.1186/1559-0275-11-16>
- Magdeldin, S., Moresco, J.J., Yamamoto, T., Yates, J.R., 2014b. Off-Line Multidimensional Liquid Chromatography and Auto Sampling Result in Sample Loss in LC/LC–MS/MS. *J. Proteome Res.* 13, 3826–3836. <https://doi.org/10.1021/pr500530e>
- Magdeldin, S., Yoshida, Y., Li, H., Maeda, Y., Yokoyama, M., Enany, S., Zhang, Y., Xu, B., Fujinaka, H., Yaoita, E., Sasaki, S., Yamamoto, T., 2012. Murine colon proteome and characterization of the protein pathways. *BioData Min.* 5, 11. <https://doi.org/10.1186/1756-0381-5-11>
- Makarov, A., 2000. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal. Chem.* 72, 1156–1162. <https://doi.org/10.1021/ac991131p>
- Makarov, A., Denisov, E., Lange, O., 2009. Performance evaluation of a high-field orbitrap mass analyzer. *J. Am. Soc. Mass Spectrom.* 20, 1391–1396. <https://doi.org/10.1016/j.jasms.2009.01.005>
- Makarov, D.E., Plaxco, K.W., 2003. The topomer search model: A simple, quantitative theory of two-state protein folding kinetics. *Protein Sci. Publ. Protein Soc.* 12, 17–26.
- Mallick, P., Schirle, M., Chen, S.S., Flory, M.R., Lee, H., Martin, D., Ranish, J., Raught, B., Schmitt, R., Werner, T., Kuster, B., Aebersold, R., 2007. Computational prediction of proteotypic peptides for quantitative proteomics. *Nat. Biotechnol.* 25, 125–131. <https://doi.org/10.1038/nbt1275>
- Manadas, B.J., Vougas, K., Fountoulakis, M., Duarte, C.B., 2006. Sample sonication after trichloroacetic acid precipitation increases protein recovery from cultured hippocampal neurons, and improves resolution and reproducibility in two-dimensional gel electrophoresis. *Electrophoresis* 27, 1825–1831.
- Marchin, K.L., Berrie, C.L., 2003. Conformational Changes in the Plasma Protein Fibrinogen upon Adsorption to Graphite and Mica Investigated by Atomic Force Microscopy. *Langmuir* 19, 9883–9888. <https://doi.org/10.1021/la035127r>
- Marikovsky, M., Ziv, V., Nevo, N., Harris-Cerruti, C., Mahler, O., 2003. Cu/Zn Superoxide Dismutase Plays Important Role in Immune Response. *J. Immunol.* 170, 2993–3001. <https://doi.org/10.4049/jimmunol.170.6.2993>
- Masciandaro, G., Macci, C., Doni, S., Maserti, B.E., Leo, A.C.-B., Ceccanti, B., Wellington, E., 2008. Comparison of extraction methods for recovery of extracellular β -glucosidase in two different forest soils. *Soil Biol. Biochem., Special Section: Enzymes in the Environment* 40, 2156–2161. <https://doi.org/10.1016/j.soilbio.2008.05.001>
- Matafora, V., Corno, A., Ciliberto, A., Bachi, A., 2017. Missing Value Monitoring Enhances the Robustness in Proteomics Quantitation. *J. Proteome Res.* 16, 1719–1727. <https://doi.org/10.1021/acs.jproteome.6b01056>
- Matulis, D., Lovrien, R., 1998. 1-Anilino-8-naphthalene sulfonate anion-protein binding depends primarily on ion pair formation. *Biophys. J.* 74, 422–429.
- Medzihradzky, K.F., 2005. In-Solution Digestion of Proteins for Mass Spectrometry, in: *Methods in Enzymology*. Academic Press, pp. 50–65. [https://doi.org/10.1016/S0076-6879\(05\)05003-2](https://doi.org/10.1016/S0076-6879(05)05003-2)
- Mesquita, R.O., Soares, E. de A., Barros, E.G. de, Loureiro, M.E., 2012. Method optimization for proteomic analysis of soybean leaf: improvements in identification of new and low-abundance proteins. *Genet. Mol. Biol.* 35, 353–361.

Références bibliographiques

- Michalski, A., Damoc, E., Hauschild, J.-P., Lange, O., Wieghaus, A., Makarov, A., Nagaraj, N., Cox, J., Mann, M., Horning, S., 2011. Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer. *Mol. Cell. Proteomics MCP* 10. <https://doi.org/10.1074/mcp.M111.011015>
- Mikhailov, V.A., Griffiths, R.L., Cooper, H.J., 2016. Liquid extraction surface analysis for native mass spectrometry: Protein complexes and ligand binding. *Int. J. Mass Spectrom.* <https://doi.org/10.1016/j.ijms.2016.09.011>
- Min, L., Choe, L.H., Lee, K.H., 2015. Improved protease digestion conditions for membrane protein detection. *Electrophoresis* 36, 1690–1698. <https://doi.org/10.1002/elps.201400579>
- Minshull, T.C., Cole, J., Dockrell, D.H., Read, R.C., Dickman, M.J., 2016. Analysis of histone post translational modifications in primary monocyte derived macrophages using reverse phase×reverse phase chromatography in conjunction with porous graphitic carbon stationary phase. *J. Chromatogr. A* 1453, 43–53. <https://doi.org/10.1016/j.chroma.2016.05.025>
- Mitulović, G., 2015. New HPLC Techniques for Proteomics Analysis: A Short Overview of Latest Developments. *J. Liq. Chromatogr. Relat. Technol.* 38, 390–403. <https://doi.org/10.1080/10826076.2014.941266>
- Monteoliva, L., Albar, J.P., 2004. Differential proteomics: an overview of gel and non-gel based approaches. *Brief. Funct. Genomic. Proteomic.* 3, 220–239.
- Morris, R.M., Nunn, B.L., Frazar, C., Goodlett, D.R., Ting, Y.S., Rocap, G., 2010. Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J* 4, 673–685.
- Motoyama, A., Venable, J.D., Ruse, C.I., Yates, J.R., 2006. Automated Ultra-High-Pressure Multidimensional Protein Identification Technology (UHP-MudPIT) for Improved Peptide Identification of Proteomic Samples. *Anal. Chem.* 78, 5109–5118. <https://doi.org/10.1021/ac060354u>
- Mukhtar, E., Adhami, V.M., Mukhtar, H., 2014. Targeting Microtubules by Natural Agents for Cancer Therapy. *Mol. Cancer Ther.* 13, 275–284. <https://doi.org/10.1158/1535-7163.MCT-13-0791>
- Müntener, T., Häussinger, D., Selenko, P., Theillet, F.-X., 2016. In-Cell Protein Structures from 2D NMR Experiments. *J. Phys. Chem. Lett.* 7, 2821–2825. <https://doi.org/10.1021/acs.jpcclett.6b01074>
- Murase, A., Yoneda, M., Ueno, R., Yonebayashi, K., 2003. Isolation of extracellular protein from greenhouse soil. *Soil Biol. Biochem.* 35, 733–736. [https://doi.org/10.1016/S0038-0717\(03\)00087-7](https://doi.org/10.1016/S0038-0717(03)00087-7)
- Muthu, M., Vimala, A., Mendoza, O.H., Gopal, J., 2016. Tracing the voyage of SELDI-TOF MS in cancer biomarker discovery and its current depreciation trend – need for resurrection? *TrAC Trends Anal. Chem.* 76, 95–101. <https://doi.org/10.1016/j.trac.2015.10.004>
- Myers, J.K., Pace, C.N., 1996. Hydrogen bonding stabilizes globular proteins. *Biophys. J.* 71, 2033–2039.
- Nägele, E., Vollmer, M., Hörth, P., Vad, C., 2004. 2D-LC/MS techniques for the identification of proteins in highly complex mixtures. *Expert Rev. Proteomics* 1, 37–46. <https://doi.org/10.1586/14789450.1.1.37>
- Nandakumar, M.P., Cheung, A., Marten, M.R., 2006. Proteomic analysis of extracellular proteins from *Escherichia coli* W3110. *J. Proteome Res.* 5, 1155–1161. <https://doi.org/10.1021/pr050401j>
- Nandakumar, M.P., Shen, J., Raman, B., Marten, M.R., 2003. Solubilization of trichloroacetic acid (TCA) precipitated microbial proteins via NaOH for two-dimensional electrophoresis. *J. Proteome Res.* 2, 89–93.

Références bibliographiques

- Nannipieri, P., Sequi, P., Fusi, P., 1996. Chapter 7 - Humus and Enzyme Activity, in: Piccolo, A. (Ed.), *Humic Substances in Terrestrial Ecosystems*. Elsevier Science B.V., Amsterdam, pp. 293–328. <https://doi.org/10.1016/B978-044481516-3/50008-6>
- Naru, J., Aggarwal, R., Mohanty, A.K., Singh, U., Bansal, D., Kakkar, N., Agnihotri, N., 2017. Identification of differentially expressed proteins in retinoblastoma tumors using mass spectrometry-based comparative proteomic approach. *J. Proteomics* 159, 77–91. <https://doi.org/10.1016/j.jprot.2017.02.006>
- Nelson, R.W., Nedelkov, D., Tubbs, K.A., Kiernan, U.A., 2004. Quantitative Mass Spectrometric Immunoassay of Insulin Like Growth Factor 1. *J. Proteome Res.* 3, 851–855. <https://doi.org/10.1021/pr0499388>
- Nestola, M., Friedrich, R., Bluhme, P., Schmidt, T.C., 2015. Universal Route to Polycyclic Aromatic Hydrocarbon Analysis in Foodstuff: Two-Dimensional Heart-Cut Liquid Chromatography–Gas Chromatography–Mass Spectrometry. *Anal. Chem.* 87, 6195–6203. <https://doi.org/10.1021/acs.analchem.5b00825>
- Neuhoff, V., Arold, N., Taube, D., Ehrhardt, W., 1988. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *ELECTROPHORESIS* 9, 255–262. <https://doi.org/10.1002/elps.1150090603>
- Nick Pace, C., Scholtz, J.M., Grimsley, G.R., 2014. Forces stabilizing proteins. *FEBS Lett.* 588, 2177–2184. <https://doi.org/10.1016/j.febslet.2014.05.006>
- Nicora, C.D., Anderson, B.J., Callister, S.J., Norbeck, A.D., Purvine, S.O., Jansson, J.K., Mason, O.U., David, M.M., Jurelevicius, D., Smith, R.D., Lipton, M.S., 2013. Amino acid treatment enhances protein recovery from sediment and soils for metaproteomic studies. *PROTEOMICS* 13, 2776–2785. <https://doi.org/10.1002/pmic.201300003>
- Nie, J., An, L., Miao, K., Hou, Z., Yu, Y., Tan, K., Sui, L., He, S., Liu, Q., Lei, X., Wu, Z., Tian, J., 2013. Comparative Analysis of Dynamic Proteomic Profiles between in Vivo and in Vitro Produced Mouse Embryos during Postimplantation Period. *J. Proteome Res.* 12, 3843–3856. <https://doi.org/10.1021/pr301044b>
- Nikkanen, L., Rintamäki, E., 2014. Thioredoxin-dependent regulatory networks in chloroplasts under fluctuating light conditions. *Phil Trans R Soc B* 369, 20130224. <https://doi.org/10.1098/rstb.2013.0224>
- Nishii, I., Kataoka, M., Goto, Y., 1995. Thermodynamic Stability of the Molten Globule States of Apomyoglobin. *J. Mol. Biol.* 250, 223–238. <https://doi.org/10.1006/jmbi.1995.0373>
- Noguera, M.E., Vazquez, D.S., Ferrer-Sueta, G., Agudelo, W.A., Howard, E., Rasia, R.M., Manta, B., Cousido-Siah, A., Mitschler, A., Podjarny, A., Santos, J., 2017. Structural variability of *E. coli* thioredoxin captured in the crystal structures of single-point mutants. *Sci. Rep.* 7. <https://doi.org/10.1038/srep42343>
- O’Connell, C.B., Khodjakov, A.L., 2007. Cooperative mechanisms of mitotic spindle formation. *J. Cell Sci.* 120, 1717–1722. <https://doi.org/10.1242/jcs.03442>
- O’Farrell, P.H., 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250, 4007–4021.
- Oshokoya, O.O., JiJi, R.D., 2015. “Parallel factor analysis of multi-excitation ultraviolet resonance Raman spectra for protein secondary structure determination.” *Anal. Chim. Acta* 892, 59–68. <https://doi.org/10.1016/j.aca.2015.08.035>
- Owusu-Apenten, R., 2004. *Introduction to Food Chemistry*. CRC Press.
- Pace, C.N., 2009. Energetics of protein hydrogen bonds. *Nat. Struct. Mol. Biol.* 16, 681–682. <https://doi.org/10.1038/nsmb0709-681>
- Paizs, B., Suhai, S., 2005. Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.* 24, 508–548. <https://doi.org/10.1002/mas.20024>
- Pampaloni, F., Florin, E.-L., 2008. Microtubule architecture: inspiration for novel carbon nanotube-based biomimetic materials. *Trends Biotechnol.* 26, 302–310. <https://doi.org/10.1016/j.tibtech.2008.03.002>

Références bibliographiques

- Pan, Y., Cheng, K., Mao, J., Liu, F., Liu, J., Ye, M., Zou, H., 2014. Quantitative proteomics reveals the kinetics of trypsin-catalyzed protein digestion. *Anal. Bioanal. Chem.* 406, 6247–6256. <https://doi.org/10.1007/s00216-014-8071-6>
- Panuwet, P., Hunter, R.E., D'Souza, P.E., Chen, X., Radford, S.A., Cohen, J.R., Marder, M.E., Kartavenka, K., Ryan, P.B., Barr, D.B., 2016. Biological Matrix Effects in Quantitative Tandem Mass Spectrometry-Based Analytical Methods: Advancing Biomonitoring. *Crit. Rev. Anal. Chem.* 46, 93–105. <https://doi.org/10.1080/10408347.2014.980775>
- Pauling, L., Corey, R.B., Branson, H.R., 1951. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci.* 37, 205–211. <https://doi.org/10.1073/pnas.37.4.205>
- Peng, B., Guo, C., Guan, H., Liu, S., Sun, M.-Z., 2014. Annexin A5 as a potential marker in tumors. *Clin. Chim. Acta Int. J. Clin. Chem.* 427, 42–48. <https://doi.org/10.1016/j.cca.2013.09.048>
- Perkins, D.N., Pappin, D.J.C., Creasy, D.M., Cottrell, J.S., 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *ELECTROPHORESIS* 20, 3551–3567. [https://doi.org/10.1002/\(SICI\)1522-2683\(19991201\)20:18<3551::AID-ELPS3551>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2)
- Phoon, P.Y., Narsimhan, G., San Martin-Gonzalez, M.F., 2013. Effect of Thermal Behavior of β -Lactoglobulin on the Oxidative Stability of Menhaden Oil-in-Water Emulsions. *J. Agric. Food Chem.* 61, 1954–1967. <https://doi.org/10.1021/jf304834n>
- Płotka-Wasyłka, J., Szczepańska, N., de la Guardia, M., Namieśnik, J., 2015. Miniaturized solid-phase extraction techniques. *TrAC Trends Anal. Chem.* 73, 19–38. <https://doi.org/10.1016/j.trac.2015.04.026>
- Ponnuswamy, P.K., Gromiha, M.M., 1994. On the Conformational Stability of Folded Proteins. *J. Theor. Biol.* 166, 63–74. <https://doi.org/10.1006/jtbi.1994.1005>
- Ponnuswamy, P.K., Muthusamy, R., Manavalan, P., 1982. Amino acid composition and thermal stability of proteins. *Int. J. Biol. Macromol.* 4, 186–190. [https://doi.org/10.1016/0141-8130\(82\)90049-6](https://doi.org/10.1016/0141-8130(82)90049-6)
- Popot, J.-L., 2010. Amphipols, Nanodiscs, and Fluorinated Surfactants: Three Nonconventional Approaches to Studying Membrane Proteins in Aqueous Solutions. *Annu. Rev. Biochem.* 79, 737–775. <https://doi.org/10.1146/annurev.biochem.052208.114057>
- Pottiez, G., Yang, L., Stewart, T., Song, N., Aro, P., Galasko, D.R., Quinn, J.F., Peskind, E.R., Shi, M., Zhang, J., 2017. Mass-Spectrometry-Based Method To Quantify in Parallel Tau and Amyloid β 1–42 in CSF for the Diagnosis of Alzheimer's Disease. *J. Proteome Res.* 16, 1228–1238. <https://doi.org/10.1021/acs.jproteome.6b00829>
- Pragya, P., Kaur, G., Ali, S.A., Bhatla, S., Rawat, P., Lule, V., Kumar, S., Mohanty, A.K., Behare, P., 2017. High-resolution mass spectrometry-based global proteomic analysis of probiotic strains *Lactobacillus fermentum* NCDC 400 and RS2. *J. Proteomics* 152, 121–130. <https://doi.org/10.1016/j.jprot.2016.10.016>
- Prieto, J.H., Koncarevic, S., Park, S.K., Iii, J.Y., Becker, K., 2008. Large-Scale Differential Proteome Analysis in *Plasmodium falciparum* under Drug Treatment. *PLOS ONE* 3, e4098. <https://doi.org/10.1371/journal.pone.0004098>
- Ptitsyn, O.B., 1995. Molten globule and protein folding. *Adv. Protein Chem.* 47, 83–229.
- Qadeer, A., Rabbani, G., Zaidi, N., Ahmad, E., Khan, J.M., Khan, R.H., 2012. 1-Anilino-8-Naphthalene Sulfonate (ANS) Is Not a Desirable Probe for Determining the Molten Globule State of Chymopapain. *PLoS ONE* 7, e50633. <https://doi.org/10.1371/journal.pone.0050633>
- Qian, C., Hettich, R.L., 2017. Optimized Extraction Method To Remove Humic Acid Interferences from Soil Samples Prior to Microbial Proteome Measurements. *J. Proteome Res.* <https://doi.org/10.1021/acs.jproteome.7b00103>
- Quan, L., Wei, D., Jiang, X., Liu, Y., Li, Z., Li, N., Li, K., Liu, F., Lai, L., 2008. Resurveying the Tris buffer solution: the specific interaction between

Références bibliographiques

- tris(hydroxymethyl)aminomethane and lysozyme. *Anal. Biochem.* 378, 144–150. <https://doi.org/10.1016/j.ab.2008.04.006>
- Rabilloud, T., Adessi, C., Giraudel, A., Lunardi, J., 2007. Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 18, 307–316. <https://doi.org/10.1002/elps.1150180303>
- Rabilloud, T., Lelong, C., 2011. Two-dimensional gel electrophoresis in proteomics: A tutorial. *J. Proteomics* 74, 1829–1841. <https://doi.org/10.1016/j.jprot.2011.05.040>
- Radzicka, A., Wolfenden, R., 1995. A proficient enzyme. *Science* 267, 90–93.
- Raines, R. T., 1998. Ribonuclease A. *Chem Rev* 98. <https://doi.org/10.1021/cr960427h>
- Raines, Ronald T., 1998. Ribonuclease A. *Chem. Rev.* 98, 1045–1066. <https://doi.org/10.1021/cr960427h>
- Rajalingam, D., Loftis, C., Xu, J.J., Kumar, T.K.S., 2009. Trichloroacetic acid-induced protein precipitation involves the reversible association of a stable partially structured intermediate. *Protein Sci. Publ. Protein Soc.* 18, 980–993. <https://doi.org/10.1002/pro.108>
- Ralla, K., Sohling, U., Riechers, D., Kasper, C., Ruf, F., Scheper, T., 2010. Adsorption and separation of proteins by a smectitic clay mineral. *Bioprocess Biosyst. Eng.* 33, 847–861. <https://doi.org/10.1007/s00449-010-0408-8>
- Raman, C.V., Krishnan, K.S., 1928. A new class of spectra due to secondary radiation. Part I. *Indian J. Phys.* 2, 399–419.
- Rao, Q., Hsieh, Y.-H.P., 2014. Enhanced immunodetection of bovine central nervous tissue using an improved extraction method. *Food Control* 46, 282–290.
- Rauniyar, N., 2015. Parallel Reaction Monitoring: A Targeted Experiment Performed Using High Resolution and High Mass Accuracy Mass Spectrometry. *Int. J. Mol. Sci.* 16, 28566–28581. <https://doi.org/10.3390/ijms161226120>
- Renard, B.Y., Xu, B., Kirchner, M., Zickmann, F., Winter, D., Korten, S., Brattig, N.W., Tzur, A., Hamprecht, F.A., Steen, H., 2012. Overcoming Species Boundaries in Peptide Identification with Bayesian Information Criterion-driven Error-tolerant Peptide Search (BICEPS). *Mol. Cell. Proteomics MCP* 11. <https://doi.org/10.1074/mcp.M111.014167>
- Reynolds, J.A., Tanford, C., 1970. Binding of dodecyl sulfate to proteins at high binding ratios. Possible implications for the state of proteins in biological membranes. *Proc. Natl. Acad. Sci. U. S. A.* 66, 1002–1007.
- Richardson, M.R., Liu, S., Ringham, H.N., Chan, V., Witzmann, F.A., 2008. Sample complexity reduction for two-dimensional electrophoresis using solution isoelectric focusing prefractionation. *Electrophoresis* 29, 2637–2644. <https://doi.org/10.1002/elps.200700707>
- Rillig, M.C., 2004. Arbuscular mycorrhizae, glomalin, and soil aggregation. *Can. J. Soil Sci.* 84, 355–363. <https://doi.org/10.4141/S04-003>
- Rinalducci, S., Roepstorff, P., Zolla, L., 2009. De novo sequence analysis and intact mass measurements for characterization of phycocyanin subunit isoforms from the blue-green alga *Aphanizomenon flos-aquae*. *J. Mass Spectrom.* *JMS* 44, 503–515. <https://doi.org/10.1002/jms.1526>
- Risler, T., 2009. Cytoskeleton Cytoskeleton and Cell Motility, in: Ph. D, R.A.M. (Ed.), *Encyclopedia of Complexity and Systems Science*. Springer New York, pp. 1738–1774. https://doi.org/10.1007/978-0-387-30440-3_112
- Roberts, P., Jones, D.L., 2008. Critical evaluation of methods for determining total protein in soil solution. *Spec. Sect. Funct. Microb. Ecol. Mol. Approaches Microb. Ecol. Microb. Habitats 18th World Congr. Soil Sci.* 40, 1485–1495. <https://doi.org/10.1016/j.soilbio.2008.01.001>
- Rodkey, F.L., 1965. Direct Spectrophotometric Determination of Albumin in Human Serum. *Clin. Chem.* 11, 478.

Références bibliographiques

- Rogers, C.A., Stockham, P.C., Nash, C.M., Martin, S.M., Kostakis, C., Lenehan, C.E., 2016. An alternative approach for assessment of liquid chromatography-mass spectrometry matrix effects using auto-sampler programmed co-injection. *Anal. Bioanal. Chem.* 408, 2009–2017. <https://doi.org/10.1007/s00216-015-9278-x>
- Ronsein, G.E., Pamir, N., von Haller, P.D., Kim, D.S., Oda, M.N., Jarvik, G.P., Vaisar, T., Heinecke, J.W., 2015. Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. *J. Proteomics* 0, 388–399. <https://doi.org/10.1016/j.jprot.2014.10.017>
- Rose, G.D., Wolfenden, R., 1993. Hydrogen Bonding, Hydrophobicity, Packing, and Protein Folding. *Annu. Rev. Biophys. Biomol. Struct.* 22, 381–415. <https://doi.org/10.1146/annurev.bb.22.060193.002121>
- Ruhl, S., Berlenbach, P., Langenfelder, S., Hörl, D., Lehn, N., Hiller, K.-A., Schmalz, G., Durchschlag, H., 2011. Integrity of Proteins in Human Saliva after Sterilization by Gamma Irradiation. *Appl. Environ. Microbiol.* 77, 749–755. <https://doi.org/10.1128/AEM.01374-10>
- Ruprecht, J.J., Hellowell, A.M., Harding, M., Crichton, P.G., McCoy, A.J., Kunji, E.R.S., 2014. Structures of yeast mitochondrial ADP/ATP carriers support a domain-based alternating-access transport mechanism. *Proc. Natl. Acad. Sci. U. S. A.* 111, E426–E434. <https://doi.org/10.1073/pnas.1320692111>
- Russo, D.A., Couto, N., Beckerman, A.P., Pandhal, J., 2016. A Metaproteomic Analysis of the Response of a Freshwater Microbial Community under Nutrient Enrichment. *Front. Microbiol.* 7, 1172. <https://doi.org/10.3389/fmicb.2016.01172>
- Rygula, A., Majzner, K., Marzec, K.M., Kaczor, A., Pilarczyk, M., Baranska, M., 2013. Raman spectroscopy of proteins: a review. *J. Raman Spectrosc.* 1061.
- Saleh, I.A., Vinatoru, M., Mason, T.J., Abdel-Azim, N.S., Aboutabl, E.A., Hammouda, F.M., 2016. A possible general mechanism for ultrasound-assisted extraction (UAE) suggested from the results of UAE of chlorogenic acid from *Cynara scolymus* L. (artichoke) leaves. *Ultrason. Sonochem.* 31, 330–336. <https://doi.org/10.1016/j.ultsonch.2016.01.002>
- Sandra, K., Steenbeke, M., Vandenheede, I., Vanhoenacker, G., Sandra, P., 2017. The versatility of heart-cutting and comprehensive two-dimensional liquid chromatography in monoclonal antibody clone selection. *J. Chromatogr. A.* <https://doi.org/10.1016/j.chroma.2017.06.052>
- Sane, S.U., Cramer, S.M., Przybycien, T.M., 1999. A Holistic Approach to Protein Secondary Structure Characterization Using Amide I Band Raman Spectroscopy. *Anal. Biochem.* 269, 255–272. <https://doi.org/10.1006/abio.1999.4034>
- Sano, S., Tagami, S., Hashimoto, Y., Yoshizawa-Kumagaye, K., Tsunemi, M., Okochi, M., Tomonaga, T., 2014. Absolute quantitation of low abundance plasma APL1 β peptides at sub-fmol/mL Level by SRM/MRM without immunoaffinity enrichment. *J. Proteome Res.* 13, 1012–1020. <https://doi.org/10.1021/pr4010103>
- Sarrut, M., Crétier, G., Heinisch, S., 2014. Theoretical and practical interest in UHPLC technology for 2D-LC. *TrAC Trends Anal. Chem., UHPLC: Where Are We 10 Years After Its Commercial Introduction?* 63, 104–112. <https://doi.org/10.1016/j.trac.2014.08.005>
- Sarsby, J., Griffiths, R.L., Race, A.M., Bunch, J., Randall, E.C., Creese, A.J., Cooper, H.J., 2015. Liquid Extraction Surface Analysis Mass Spectrometry Coupled with Field Asymmetric Waveform Ion Mobility Spectrometry for Analysis of Intact Proteins from Biological Substrates. *Anal. Chem.* 87, 6794–6800. <https://doi.org/10.1021/acs.analchem.5b01151>
- Scheerlinck, E., Dhaenens, M., Van Soom, A., Peelman, L., De Sutter, P., Van Steendam, K., Deforce, D., 2015. Minimizing technical variation during sample preparation prior

Références bibliographiques

- to label-free quantitative mass spectrometry. *Anal. Biochem.* 490, 14–19. <https://doi.org/10.1016/j.ab.2015.08.018>
- Schiess, R., Wollscheid, B., Aebersold, R., 2009. Targeted proteomic strategy for clinical biomarker discovery. *Mol. Oncol.* 3, 33–44. <https://doi.org/10.1016/j.molonc.2008.12.001>
- Schiffmann, C., Hansen, R., Baumann, S., Kublik, A., Nielsen, P.H., Adrian, L., von Bergen, M., Jehmlich, N., Seifert, J., 2014. Comparison of targeted peptide quantification assays for reductive dehalogenases by selective reaction monitoring (SRM) and precursor reaction monitoring (PRM). *Anal. Bioanal. Chem.* 406, 283–291. <https://doi.org/10.1007/s00216-013-7451-7>
- Schlosser, G., Kačer, P., Kuzma, M., Szilágyi, Z., Sorrentino, A., Manzo, C., Pizzano, R., Malorni, L., Pocsfalvi, G., 2007. Coupling Immunomagnetic Separation on Magnetic Beads with Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry for Detection of Staphylococcal Enterotoxin B. *Appl. Environ. Microbiol.* 73, 6945–6952. <https://doi.org/10.1128/AEM.01136-07>
- Schmidt, M.P., Martínez, C.E., 2016. Kinetic and Conformational Insights of Protein Adsorption onto Montmorillonite Revealed Using in Situ ATR-FTIR/2D-COS. *Langmuir* 32, 7719–7729. <https://doi.org/10.1021/acs.langmuir.6b00786>
- Schultz, C.P., 2000. Illuminating folding intermediates. *Nat. Struct. Mol. Biol.* 7, 7–10. <https://doi.org/10.1038/71197>
- Schulze, W.X., Gleixner, G., Kaiser, K., Guggenberger, G., Mann, M., Schulze, E.-D., 2005. A proteomic fingerprint of dissolved organic carbon and of soil particles. *Oecologia* 142, 335–343. <https://doi.org/10.1007/s00442-004-1698-9>
- Scigelova, M., Makarov, A., 2006. Orbitrap Mass Analyzer – Overview and Applications in Proteomics. *PROTEOMICS* 6, 16–21. <https://doi.org/10.1002/pmic.200600528>
- Seddon, A.M., Curnow, P., Booth, P.J., 2004. Membrane proteins, lipids and detergents: not just a soap opera. *Biochim. Biophys. Acta* 1666, 105–117. <https://doi.org/10.1016/j.bbamem.2004.04.011>
- Semisotnov, G.V., Rodionova, N.A., Razgulyaev, O.I., Uversky, V.N., Gripas', A.F., Gilmanshin, R.I., 1991. Study of the “molten globule” intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* 31, 119–128. <https://doi.org/10.1002/bip.360310111>
- Shao, Q., Fan, Y., Yang, L., Gao, Y.Q., 2012. From protein denaturant to protectant: comparative molecular dynamics study of alcohol/protein interactions. *J. Chem. Phys.* 136, 115101. <https://doi.org/10.1063/1.3692801>
- Shen, Y., Zhang, R., Moore, R.J., Kim, J., Metz, T.O., Hixson, K.K., Zhao, R., Livesay, E.A., Udseth, H.R., Smith, R.D., 2005. Automated 20 kpsi RPLC-MS and MS/MS with Chromatographic Peak Capacities of 1000–1500 and Capabilities in Proteomics and Metabolomics. *Anal. Chem.* 77, 3090–3100. <https://doi.org/10.1021/ac0483062>
- Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., Mann, M., 2007. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc.* 1, 2856–2860. <https://doi.org/10.1038/nprot.2006.468>
- Shi, J., Chen, Y., Xu, Y., Ji, D., Chen, C., Xie, C., 2017. Differential Proteomic Analysis by iTRAQ Reveals the Mechanism of *Pyropia haitanensis* Responding to High Temperature Stress. *Sci. Rep.* 7. <https://doi.org/10.1038/srep44734>
- Shi, R., Kumar, C., Zougman, A., Zhang, Y., Podtelejnikov, A., Cox, J., Wiśniewski, J.R., Mann, M., 2007. Analysis of the mouse liver proteome using advanced mass spectrometry. *J. Proteome Res.* 6, 2963–2972. <https://doi.org/10.1021/pr0605668>
- Shi, T., Song, E., Nie, S., Rodland, K.D., Liu, T., Qian, W.-J., Smith, R.D., 2016. Advances in targeted proteomics and applications to biomedical research. *Proteomics* 16, 2160–2182. <https://doi.org/10.1002/pmic.201500449>

Références bibliographiques

- Simoneit, B.R.T., 2002. Molecular indicators (biomarkers) of past life. *Anat. Rec.* 268, 186–195. <https://doi.org/10.1002/ar.10153>
- Singh, S.M., Sharma, A., Upadhyay, A.K., Singh, A., Garg, L.C., Panda, A.K., 2012. Solubilization of inclusion body proteins using n-propanol and its refolding into bioactive form. *Protein Expr. Purif.* 81, 75–82. <https://doi.org/10.1016/j.pep.2011.09.004>
- Singleton, I., Merrington, G., Colvan, S., Delahunty, J.S., 2003. The potential of soil protein-based methods to indicate metal contamination. *Appl. Soil Ecol.* 23, 25–32. [https://doi.org/10.1016/S0929-1393\(03\)00004-0](https://doi.org/10.1016/S0929-1393(03)00004-0)
- Sinnaeve, B.A., Van Bocxlaer, J.F., 2004. Evaluation of nano-liquid chromatography-tandem mass spectrometry in a column switching setup for the absolute quantification of peptides in the picomolar range. *J. Chromatogr. A* 1058, 113–119.
- Sivaraman, T., Kumar, T.K., Jayaraman, G., Yu, C., 1997. The mechanism of 2,2,2-trichloroacetic acid-induced protein precipitation. *J. Protein Chem.* 16, 291–297.
- Šlechtová, T., Gilar, M., Kalíková, K., Moore, S.M., Jorgenson, J.W., Tesařová, E., 2017. Performance comparison of three trypsin columns used in liquid chromatography. *J. Chromatogr. A* 1490, 126–132. <https://doi.org/10.1016/j.chroma.2017.02.024>
- Smith, K.T., Workman, J.L., 2009. Introducing the acetylome. *Nat. Biotechnol.* 27, 917–919. <https://doi.org/10.1038/nbt1009-917>
- Song, S., Hooiveld, G.J., Zhang, W., Li, M., Zhao, F., Zhu, J., Xu, X., Muller, M., Li, C., Zhou, G., 2016. Comparative Proteomics Provides Insights into Metabolic Responses in Rat Liver to Isolated Soy and Meat Proteins. *J. Proteome Res.* 15, 1135–1142. <https://doi.org/10.1021/acs.jproteome.5b00922>
- Song, Y., Zhang, H., Wang, G., Shen, Z., 2012a. DMSO, an Organic Cleanup Solvent for TCA/Acetone-Precipitated Proteins, Improves 2-DE Protein Analysis of Rice Roots. *Plant Mol. Biol. Report.* 30, 1204–1209.
- Song, Y., Zhang, H., Wang, G., Shen, Z., 2012b. DMSO, an Organic Cleanup Solvent for TCA/Acetone-Precipitated Proteins, Improves 2-DE Protein Analysis of Rice Roots. *Plant Mol. Biol. Report.* 30, 1204–1209.
- Starke, R., Kermer, R., Ullmann-Zeunert, L., Baldwin, I.T., Seifert, J., Bastida, F., von Bergen, M., Jehmlich, N., 2016. Bacteria dominate the short-term assimilation of plant-derived N in soil. *Soil Biol. Biochem.* 96, 30–38. <https://doi.org/10.1016/j.soilbio.2016.01.009>
- Stathopoulos, P.B., Scholz, G.A., Hwang, Y.-M., Rumfeldt, J.A.O., Lepock, J.R., Meiering, E.M., 2004. Sonication of proteins causes formation of aggregates that resemble amyloid. *Protein Sci. Publ. Protein Soc.* 13, 3017–3027. <https://doi.org/10.1110/ps.04831804>
- Stephan, S., Jakob, C., Hippler, J., Schmitz, O.J., 2016. A novel four-dimensional analytical approach for analysis of complex samples. *Anal. Bioanal. Chem.* 408, 3751–3759. <https://doi.org/10.1007/s00216-016-9460-9>
- Stevenson, F., 1994. *Humus Chemistry: Genesis, Composition, Reactions*, 2nd Edition, 2nd Edition. ed. John Wiley & Sons.
- Sun, W., Li, F., Wu, S., Wang, X., Zheng, D., Wang, J., Gao, Y., 2005. Human urine proteome analysis by three separation approaches. *Proteomics* 5, 4994–5001. <https://doi.org/10.1002/pmic.200401334>
- Taha, M., Lee, M.-J., 2010. Interactions of TRIS [tris(hydroxymethyl)aminomethane] and related buffers with peptide backbone: thermodynamic characterization. *Phys. Chem. Chem. Phys.* PCCP 12, 12840–12850. <https://doi.org/10.1039/c0cp00253d>
- Taylor, P., Nielsen, P.A., Trelle, M.B., Hørning, O.B., Andersen, M.B., Vorm, O., Moran, M.F., Kislinger, T., 2009. Automated 2D Peptide Separation on a 1D Nano-LC-MS System. *J. Proteome Res.* 8, 1610–1616. <https://doi.org/10.1021/pr800986c>

Références bibliographiques

- Tejaswi Naidu, K., Prakash Prabhu, N., 2011. Protein–Surfactant Interaction: Sodium Dodecyl Sulfate-Induced Unfolding of Ribonuclease A. *J. Phys. Chem. B* 115, 14760–14767. <https://doi.org/10.1021/jp2062496>
- Thongboonkerd, V., Mungdee, S., Chiangjong, W., 2009. Should urine pH be adjusted prior to gel-based proteome analysis? *J. Proteome Res.* 8, 3206–3211. <https://doi.org/10.1021/pr900127x>
- Tjernberg, A., Markova, N., Griffiths, W.J., Hallén, D., 2006. DMSO-related effects in protein characterization. *J. Biomol. Screen.* 11, 131–137. <https://doi.org/10.1177/1087057105284218>
- Toby, T.K., Fornelli, L., Kelleher, N.L., 2016. Progress in Top-Down Proteomics and the Analysis of Proteoforms. *Annu. Rev. Anal. Chem. Palo Alto Calif* 9, 499–519. <https://doi.org/10.1146/annurev-anchem-071015-041550>
- Tomaszewski, J.E., Schwarzenbach, R.P., Sander, M., 2011. Protein Encapsulation by Humic Substances. *Environ. Sci. Technol.* 45, 6003–6010. <https://doi.org/10.1021/es200663h>
- Trufelli, H., Palma, P., Famigliani, G., Cappiello, A., 2011. An overview of matrix effects in liquid chromatography–mass spectrometry. *Mass Spectrom. Rev.* 30, 491–509. <https://doi.org/10.1002/mas.20298>
- Tuma, R., 2005. Raman spectroscopy of proteins: from peptides to large assemblies. *J. Raman Spectrosc.* 36, 307–319.
- Uchida, Y., Tachikawa, M., Obuchi, W., Hoshi, Y., Tomioka, Y., Ohtsuki, S., Terasaki, T., 2013. A study protocol for quantitative targeted absolute proteomics (QTAP) by LC-MS/MS: application for inter-strain differences in protein expression levels of transporters, receptors, claudin-5, and marker proteins at the blood–brain barrier in ddY, FVB, and C57BL/6J mice. *Fluids Barriers CNS* 10, 21. <https://doi.org/10.1186/2045-8118-10-21>
- Uversky, V.N., Dunker, A.K., 2010. Understanding Protein Non-Folding. *Biochim. Biophys. Acta* 1804, 1231–1264. <https://doi.org/10.1016/j.bbapap.2010.01.017>
- Valeja, S.G., Xiu, L., Gregorich, Z.R., Guner, H., Jin, S., Ge, Y., 2015. Three Dimensional Liquid Chromatography Coupling Ion Exchange Chromatography/Hydrophobic Interaction Chromatography/Reverse Phase Chromatography for Effective Protein Separation in Top-Down Proteomics. *Anal. Chem.* 87, 5363–5371. <https://doi.org/10.1021/acs.analchem.5b00657>
- Varzakas, T., Tzia, C., 2014. *Food Engineering Handbook: Food Process Engineering*. CRC Press.
- Vertommen, A., Panis, B., Swennen, R., Carpentier, S.C., 2010. Evaluation of chloroform/methanol extraction to facilitate the study of membrane proteins of non-model plants. *Planta* 231, 1113–1125. <https://doi.org/10.1007/s00425-010-1121-1>
- Voet, D., Voet, J.G., 2016. *Biochimie. De Boeck Supérieur*.
- Vonk, R.J., Gargano, A.F.G., Davydova, E., Dekker, H.L., Eeltink, S., de Koning, L.J., Schoenmakers, P.J., 2015. Comprehensive Two-Dimensional Liquid Chromatography with Stationary-Phase-Assisted Modulation Coupled to High-Resolution Mass Spectrometry Applied to Proteome Analysis of *Saccharomyces cerevisiae*. *Anal. Chem.* 87, 5387–5394. <https://doi.org/10.1021/acs.analchem.5b00708>
- Waller, L.N., Shores, K., Knapp, D.R., 2008. Shotgun Proteomic Analysis of Cerebrospinal Fluid Using Off-Gel Electrophoresis as the First Dimension Separation. *J. Proteome Res.* 7, 4577–4584. <https://doi.org/10.1021/pr8001518>
- Walsh, C.T., O'Brien, R.V., Khosla, C., 2013. Nonproteinogenic Amino Acid Building Blocks for Nonribosomal Peptide and Hybrid Polyketide Scaffolds. *Angew. Chem. Int. Ed Engl.* 52, 7098–7124. <https://doi.org/10.1002/anie.201208344>
- Wang, D.-Z., Xie, Z.-X., Zhang, S.-F., 2014. Marine metaproteomics: Current status and future directions. *Spec. Issue Trends Microb. Proteomics* 97, 27–35. <https://doi.org/10.1016/j.jprot.2013.08.024>

Références bibliographiques

- Wang, F., Chen, R., Zhu, J., Sun, D., Song, C., Wu, Y., Ye, M., Wang, L., Zou, H., 2010. A Fully Automated System with Online Sample Loading, Isotope Dimethyl Labeling and Multidimensional Separation for High-Throughput Quantitative Proteome Analysis. *Anal. Chem.* 82, 3007–3015. <https://doi.org/10.1021/ac100075y>
- Wang, H.-B., Zhang, Z.-X., Li, H., He, H.-B., Fang, C.-X., Zhang, A.-J., Li, Q.-S., Chen, R.-S., Guo, X.-K., Lin, H.-F., Wu, L.-K., Lin, S., Chen, T., Lin, R.-Y., Peng, X.-X., Lin, W.-X., 2011. Characterization of metaproteomics in crop rhizospheric soil. *J. Proteome Res.* 10, 932–940. <https://doi.org/10.1021/pr100981r>
- Wang, L., Weller, C.L., 2006. Recent advances in extraction of nutraceuticals from plants. *Trends Food Sci. Technol.* 17, 300–312. <https://doi.org/10.1016/j.tifs.2005.12.004>
- Wang, S., Qiao, L., Shi, X., Hu, C., Kong, H., Xu, G., 2015. On-line stop-flow two-dimensional liquid chromatography–mass spectrometry method for the separation and identification of triterpenoid saponins from ginseng extract. *Anal. Bioanal. Chem.* 407, 331–341. <https://doi.org/10.1007/s00216-014-8219-4>
- Wang, S., Shi, X., Xu, G., 2017. Online Three Dimensional Liquid Chromatography/Mass Spectrometry Method for the Separation of Complex Samples. *Anal. Chem.* 89, 1433–1438. <https://doi.org/10.1021/acs.analchem.6b04401>
- Wessel, D., Flügge, U.I., 1984. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* 138, 141–143. [https://doi.org/10.1016/0003-2697\(84\)90782-6](https://doi.org/10.1016/0003-2697(84)90782-6)
- Wilkins, M.R., Sanchez, J.C., Gooley, A.A., Appel, R.D., Humphery-Smith, I., Hochstrasser, D.F., Williams, K.L., 1996. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol. Genet. Eng. Rev.* 13, 19–50.
- Williams, M.A., Taylor, E.B., Mula, H.P., 2010. Metaproteomic characterization of a soil microbial community following carbon amendment. *Soil Biol. Biochem.* 42, 1148–1156. <https://doi.org/10.1016/j.soilbio.2010.03.021>
- Wilm, M., 2011. Principles of Electrospray Ionization. *Mol. Cell. Proteomics MCP* 10. <https://doi.org/10.1074/mcp.M111.009407>
- Wilmes, P., Bond, P.L., 2004. The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. *Environ. Microbiol.* 6, 911–920.
- Wilson, I.D., Nicholson, J.K., Castro-Perez, J., Granger, J.H., Johnson, K.A., Smith, B.W., Plumb, R.S., 2005. High Resolution “Ultra Performance” Liquid Chromatography Coupled to oa-TOF Mass Spectrometry as a Tool for Differential Metabolic Pathway Profiling in Functional Genomic Studies. *J. Proteome Res.* 4, 591–598. <https://doi.org/10.1021/pr049769r>
- Wilson, S.R., Vehus, T., Berg, H.S., Lundanes, E., 2015. Nano-LC in proteomics: recent advances and approaches. *Bioanalysis* 7, 1799–1815. <https://doi.org/10.4155/bio.15.92>
- Witzel, K., Shahzad, M., Matros, A., Mock, H.-P., Mühling, K.H., 2011. Comparative evaluation of extraction methods for apoplastic proteins from maize leaves. *Plant Methods* 7, 48. <https://doi.org/10.1186/1746-4811-7-48>
- Wu, L.-Z., Sheng, Y.-B., Xie, J.-B., Wang, W., 2008. Photoexcitation of tryptophan groups induced reduction of disulfide bonds in hen egg white lysozyme. *J. Mol. Struct.* 882, 101–106. <https://doi.org/10.1016/j.molstruc.2007.09.016>
- Wu, Q., Yuan, H., Zhang, L., Zhang, Y., 2012. Recent advances on multidimensional liquid chromatography–mass spectrometry for proteomics: From qualitative to quantitative analysis—A review. *Anal. Chim. Acta* 731, 1–10. <https://doi.org/10.1016/j.aca.2012.04.010>
- Wu, X., Xiong, E., Wang, W., Scali, M., Cresti, M., 2014. Universal sample preparation method integrating trichloroacetic acid/acetone precipitation with phenol extraction for crop proteomic analysis. *Nat Protoc.* 9, 362–374.

Références bibliographiques

- Wu, Y., Teng, Y., Li, Z., Liao, X., Luo, Y., 2008. Potential role of polycyclic aromatic hydrocarbons (PAHs) oxidation by fungal laccase in the remediation of an aged contaminated soil. *Soil Biol. Biochem.* 40, 789–796. <https://doi.org/10.1016/j.soilbio.2007.10.013>
- Wyckoff, A.C., Lockwood, K.L., Meyerett-Reid, C., Michel, B.A., Bender, H., VerCauteren, K.C., Zabel, M.D., 2013. Estimating Prion Adsorption Capacity of Soil by BioAssay of Subtracted Infectivity from Complex Solutions (BASICS). *PLoS ONE* 8. <https://doi.org/10.1371/journal.pone.0058630>
- Xiao, H., Zhang, Y., Kim, Y., Kim, S., Kim, J.J., Kim, K.M., Yoshizawa, J., Fan, L.-Y., Cao, C.-X., Wong, D.T.W., 2016. Differential Proteomic Analysis of Human Saliva using Tandem Mass Tags Quantification for Gastric Cancer Detection. *Sci. Rep.* 6, 22165. <https://doi.org/10.1038/srep22165>
- Xiao, Y., Ma, B., McElheny, D., Parthasarathy, S., Long, F., Hoshi, M., Nussinov, R., Ishii, Y., 2015. A β (1-42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. *Nat. Struct. Mol. Biol.* 22, 499–505. <https://doi.org/10.1038/nsmb.2991>
- Xu, Z., Xie, Q., Zhou, H.-M., 2003. Trichloroacetic acid-induced molten globule state of aminoacylase from pig kidney. *J. Protein Chem.* 22, 669–675.
- Yamashita, M., Fenn, J.B., 1984. Electrospray ion source. Another variation on the free-jet theme. *J. Phys. Chem.* 88, 4451–4459. <https://doi.org/10.1021/j150664a002>
- Yao, Y., Dong, J., Dong, M., Liu, F., Wang, Y., Mao, J., Ye, M., Zou, H., 2017. An immobilized titanium (IV) ion affinity chromatography adsorbent for solid phase extraction of phosphopeptides for phosphoproteome analysis. *J. Chromatogr. A, High-Resolution Separations of Proteins and Peptides* 1498, 22–28. <https://doi.org/10.1016/j.chroma.2017.03.026>
- Ye, F., Smith, P.B., Wu, C., Chiu, D.T., 2013. Ultrasensitive Detection of Proteins on Western Blots with Semiconducting Polymer Dots. *Macromol. Rapid Commun.* 34. <https://doi.org/10.1002/marc.201200809>
- Ye, X., Li, L., 2012. Microwave-assisted protein solubilization for mass spectrometry-based shotgun proteome analysis. *Anal. Chem.* 84, 6181–6191. <https://doi.org/10.1021/ac301169q>
- Yeh, C.T., Han, C.M., Lo, S.Y., Ou, J.H., Fan, K.D., Sheen, I.S., Chu, C.M., Liaw, Y.F., 1994. Early detection of anti-HC_c antibody in acute hepatitis C virus (HCV) by western blot (immunoblot) using a recombinant HCV core protein fragment. *J. Clin. Microbiol.* 32, 2235–2241.
- Yeung, Y.-G., Stanley, E.R., 2010. Rapid Detergent Removal From Peptide Samples With Ethyl Acetate For Mass Spectrometry Analysis. *Curr. Protoc. Protein Sci.* Editor. Board John E Coligan Al CHAPTER, Unit-16.12. <https://doi.org/10.1002/0471140864.ps1612s59>
- Yokoo, H., Kondo, T., Fujii, K., Yamada, T., Todo, S., Hirohashi, S., 2004. Proteomic signature corresponding to alpha fetoprotein expression in liver cancer cells. *Hepatology* 40, 609–617. <https://doi.org/10.1002/hep.20372>
- Yu, W.H., Li, N., Tong, D.S., Zhou, C.H., Lin, C.X. (Cynthia), Xu, C.Y., 2013a. Adsorption of proteins and nucleic acids on clay minerals and their interactions: A review. *Appl. Clay Sci.* 80–81, 443–452. <https://doi.org/10.1016/j.clay.2013.06.003>
- Yu, W.H., Li, N., Tong, D.S., Zhou, C.H., Lin, C.X. (Cynthia), Xu, C.Y., 2013b. Adsorption of proteins and nucleic acids on clay minerals and their interactions: A review. *Appl. Clay Sci.* 80–81, 443–452. <https://doi.org/10.1016/j.clay.2013.06.003>
- Zamanian-Daryoush, M., DiDonato, J.A., 2015. Apolipoprotein A-I and Cancer. *Front. Pharmacol.* 6. <https://doi.org/10.3389/fphar.2015.00265>
- Zamanian-Daryoush, M., Lindner, D., Tallant, T.C., Wang, Z., Buffa, J., Klipfell, E., Parker, Y., Hatala, D., Parsons-Wingerter, P., Rayman, P., Yusufshaq, M.S.S., Fisher, E.A.,

Références bibliographiques

- Smith, J.D., Finke, J., DiDonato, J.A., Hazen, S.L., 2013. The Cardioprotective Protein Apolipoprotein A1 Promotes Potent Anti-tumorigenic Effects♦. *J. Biol. Chem.* 288, 21237–21252. <https://doi.org/10.1074/jbc.M113.468967>
- Zampieri, E., Chiapello, M., Daghino, S., Bonfante, P., Mello, A., 2016. Soil metaproteomics reveals an inter-kingdom stress response to the presence of black truffles. *Sci. Rep.* 6, 25773. <https://doi.org/10.1038/srep25773>
- Zhang, E., Chen, X., Liang, X., 2011. Resolubilization of TCA precipitated plant proteins for 2-D electrophoresis. *Electrophoresis* 32, 696–698. <https://doi.org/10.1002/elps.201000557>
- Zhang, X., Fang, A., Riley, C.P., Wang, M., Regnier, F.E., Buck, C., 2010. Multi-dimensional Liquid Chromatography in Proteomics. *Anal. Chim. Acta* 664, 101–113. <https://doi.org/10.1016/j.aca.2010.02.001>
- Zhang, Y., Bottinelli, D., Lisacek, F., Luban, J., Strambio-De-Castillia, C., Varesio, E., Hopfgartner, G., 2015. Optimization of human dendritic cell sample preparation for mass spectrometry-based proteomic studies. *Anal. Biochem.* 484, 40–50. <https://doi.org/10.1016/j.ab.2015.05.007>
- Zhao, J., Qiu, W., Simeone, D.M., Lubman, D.M., 2007. N-linked glycosylation profiling of pancreatic cancer serum using capillary liquid phase separation coupled with mass spectrometric analysis. *J. Proteome Res.* 6, 1126–1138. <https://doi.org/10.1021/pr0604458>
- Zhao, Y., Szeto, S.S.W., Kong, R.P.W., Law, C.H., Li, G., Quan, Q., Zhang, Z., Wang, Y., Chu, I.K., 2014. Online Two-Dimensional Porous Graphitic Carbon/Reversed Phase Liquid Chromatography Platform Applied to Shotgun Proteomics and Glycoproteomics. *Anal. Chem.* 86, 12172–12179. <https://doi.org/10.1021/ac503254t>
- Zhao, Y.-Y., Lin, R.-C., 2014. UPLC–MSE application in disease biomarker discovery: The discoveries in proteomics to metabolomics. *Chem. Biol. Interact.* 215, 7–16. <https://doi.org/10.1016/j.cbi.2014.02.014>
- Zhou, F., Lu, Y., Ficarro, S.B., Adelmant, G., Jiang, W., Luckey, C.J., Marto, J.A., 2013. Genome-scale proteome quantification by DEEP SEQ mass spectrometry. *Nat. Commun.* 4. <https://doi.org/10.1038/ncomms3171>
- Zhou, J.-Y., Dann, G.P., Shi, T., Wang, L., Gao, X., Su, D., Nicora, C.D., Shukla, A.K., Moore, R.J., Liu, T., Camp II, D.G., Smith, R.D., Qian, W.-J., 2012. Simple Sodium Dodecyl Sulfate-Assisted Sample Preparation Method for LC-MS-Based Proteomics Applications. *Anal. Chem.* 84, 2862–2867. <https://doi.org/10.1021/ac203394r>
- Zink, J., Wyrobnik, T., Prinz, T., Schmid, M., 2016. Physical, Chemical and Biochemical Modifications of Protein-Based Films and Coatings: An Extensive Review. *Int. J. Mol. Sci.* 17. <https://doi.org/10.3390/ijms17091376>
- Zubarev, R.A., Makarov, A., 2013. Orbitrap Mass Spectrometry. *Anal. Chem.* 85, 5288–5296. <https://doi.org/10.1021/ac4001223>

Contributions scientifiques

Publications

B.Renoux, F.Raes, T.Legigan, E.Péraudeau; **B.Eddhif**, P.Poinot, I.Tranoy-Opalinski, J.Alsarraf, S.Lerondel, A.Lepape, J.Clarhaut, S.Papot. **Targeting the Tumour Microenvironment with Enzyme-Responsive Drug Delivery System for Efficient Therapy of Breast and Pancreatic Cancers.** (En ligne dans *Chemical Science* IF: 8.668).

B.Eddhif, A.Allavena, S.Liu, T.Ribette, N.Abou Mrad, T.Chiavassa, L.Le Sergeant d'Hendecourt, R.Sternberg, G.Danger, C.Geffroy, P.Poinot. **Development of Liquid Chromatography High Resolution Mass Spectrometry Strategies for the Screening of Complex Organic Matter: Application to Astrophysical Simulated Materials.** (En ligne dans *Talanta* IF: 4.712).

B.Eddhif, J.Lange, N.Guignard, Y.Bateneau, J.Clarhaut, S.Papot, C.Geffroy, P.Poinot. **Study of a Novel Agent for TCA Precipitated Proteins Washing - Comprehensive Insights into the Role of Ethanol/HCl on Molten Globule State by Multi-Spectroscopic Analyses.** (En ligne dans *Journal of Proteomic* IF: 4.173).

Communications orales

Balkis Eddhif, Claude Geffroy, Nathalie Karpel, Pauline Poinot. Recherche de biomarqueurs protéiques : un défi analytique pour un diagnostic environnemental ciblé. Congrès national des chercheurs francophones en géochimie organique (FROG) du 7 au 8 juillet 2016 à Poitiers, France.

Balkis Eddhif, Claude Geffroy, Nathalie Karpel, Pauline Poinot. Approche analytique simplifiée appliquée à des biomarqueurs protéiques. Colloque des doctorants 3^{ème} année le 15 février 2017 à Poitiers, France.

Communications affichées

Balkis Eddhif, Claude Geffroy, Laurent Lemée Pauline Poinot. Development of an analytical approach for identification and quantification of protein biomarkers in complex environmental matrices. European conference on analytical chemistry « EUROANALYSIS » du 6 au 10 septembre, Bordeaux, France.

Balkis Eddhif, Claude Geffroy, Laurent Lemée Pauline Poinot. Characterization and quantification of protein biomarkers in complex environmental matrices: optimization of the analytical approach. 27th International Meeting on Organic Geochemistry « IMOG » du 13 au 18 septembre à Prague, Czech Republic.

Balkis Eddhif, Claude Geffroy, Pauline Poinot. Comprehensive Protocol for Bottom–up Mass Spectrometry-based Proteomics: Towards a dual-Step Approach for Protein Purification and Solubilization. Journées Françaises de Spectrométrie de masse JFSM du 27 au 30 septembre 2016, Bordeaux, France.

Balkis Eddhif, Claude Geffroy, Pauline Poinot. Comprehensive Protocol for Bottom–up Mass Spectrometry-based Proteomics: Towards a dual-Step Approach for Protein Purification and Solubilization. Congrès de l'Association Francophone des Sciences Séparatives du 28 au 30 mars 2017, Paris, France.

Résumés

Titre : Développement d'une approche analytique non ciblée pour l'étude des protéines dans les milieux complexes, environnementaux et biologiques.

RÉSUMÉ

Grâce aux récents progrès en terme d'instrumentation analytique, la protéomique, en tant que science qui étudie le protéome d'un organisme ou d'un milieu, a connu un véritable tournant et a permis d'étendre le champ des connaissances sur le fonctionnement du vivant dans son ensemble (structure, fonction, métabolisme, dynamisme). Néanmoins, l'étude des protéomes représente un challenge pour de nombreux biologistes, chimistes et biochimistes, en raison notamment de la complexité des échantillons étudiés. De nombreux protocoles analytiques ont d'ores et déjà été développés. Cependant, dans leur ensemble, ces stratégies sont relativement longues et multi-étapes et le plus souvent ciblées sur une protéine donnée.

Dans ce contexte, ce travail de thèse a pour objectif de simplifier les protocoles expérimentaux existants afin de limiter les pertes en protéines et ainsi amplifier leurs signaux au cours d'une analyse « *bottom-up* » par LC-HRMS (analyse « hors gel »). Chaque étape du processus a été décortiquée et optimisée. Dans un premier temps, les paramètres UPLC-HRMS/MS ont été optimisés afin d'améliorer la détection et la quantification des protéines présentes à des concentrations très variables dans les milieux étudiés. Ensuite, une approche de purification simplifiée qui repose sur une seule et unique étape de lavage et solubilisation des protéines a été mise au point. La démonstration de son efficacité « chimique » et « biologique » a ensuite été réalisée *via* une étude mécanistique au cours de laquelle les changements de conformation des protéines ont été étudiés à chacune des étapes de purification proposées. Enfin, certains paramètres influençant l'extraction des protéines à partir de ces mêmes matrices ont été étudiés afin de proposer à terme un protocole d'extraction à la carte compatible avec une analyse directe par LC-HRMS/MS.

Mots clés : protéines, analyse « *bottom up* », LC-HRMS/MS, purification, extraction, étude mécanistique.

Title: Development of an untargeted analytical approach for the identification and quantification of proteins in complex biological and environmental matrices.

ABSTRACT

Recent advances in proteomics have been spurred by the rapid development of hybrid and/or high-resolution mass analysers (HRMS/MS). These powerful instrumentations have led to significant improvements in « *bottom-up* » approach and have enabled to deepen our knowledge on the functionality of biological systems (structure, function, metabolism, dynamic, etc).

Despite their high sensibilities, the potential of such instruments could be significantly lessened by an imperfect sample pre-treatment. In this context, current sample pretreatments follow multi-steps experimental workflows, which alternatively lead to low recoveries of proteins.

In this line, this study aims at developing a simple and versatile strategy in order to reduce protein losses and enhance their detection in gel-free LC-MS analysis. First, an analytical method based on liquid chromatography and tandem mass spectrometry was developed to detect and quantify complex peptides mixture. Secondly, a universal, simple and fast purification approach was designed with the aim to purify protein extracts in only one-step. For this purpose, the molecular reactivity, dynamics and conformational changes of proteins at each development step were comprehensively investigated with a set of spectroscopic techniques, in order to select the best strategy. Finally, different factors influencing extraction of proteins were investigated with the goal in the long term to propose an on-demand extraction protocol for direct analysis of proteins by LC-HRMS/MS.

Key words: proteins, « *bottom-up* » analysis, LC-HRMS/MS, purification, extraction, mechanistic study.